# Comparison of Distance Discrepancies in Georeferenced and Non-Georeferenced Terrestrial LiDAR Models

**Gustavo Maldonado, Ph.D., P.E. and Marcel Maghiar, Ph.D.**
Georgia Southern University
Statesboro, Georgia

**Ananya Augustine**
Arcadis U.S., Inc.
Atlanta, Georgia

The objective of this study is twofold: (a) To investigate the magnitude of errors introduced by a typical georeferencing procedure when it is applied to a relatively large (five acres), non-georeferenced, but still accurate point-cloud model. The modeled area is a commercial site, and its model was generated via static terrestrial LiDAR. (b) To complete practical, hands-on service-learning activities that benefit not only members of the local community, but also graduate and undergraduate students whose learning is enhanced through experiential practical opportunities offered by our engineering and construction programs. The resulting 3D virtual models were donated to the property owners and serve as virtual surveying tools for the design of future modifications or additions to the involved buildings, parking lots, and landscaped areas. Distance comparisons between the non-georeference and the georeference models were completed. Their resulting statistical analysis indicated that even an accurate georeferencing approach introduced discrepancies in distances when they were measured from a virtual georeferenced 3D point-cloud model or from a non-georeferenced one. The comparison involved a total of 1016 distances from each of the two involved models. Those distances ranged from 4 to 700 ft. The Standard Deviation and Root Mean Square values of all discrepancies in distances were equal to 0.037 ft (1.1 cm). The resulting georeferenced model did not produce distance discrepancies larger than 0.111 ft (3.4 cm) with respect to the non-georeferenced one. This conclusion is valid for the case presented in this study and could be extrapolated to other cases with similarly accurate models.

**Key Words:** Georeferencing, Terrestrial LiDAR, Scanning, Accuracy

## Introduction, Objectives and Literature Review

The goal of this work was to complete a relatively large, service-learning, research project with graduate and undergraduate students in the Civil Engineering (CE), Construction Engineering (ConE) and Construction Management (CM) programs at Georgia Southern University. It consisted of two main objectives: (a) Collection of spatial data to analyze discrepancies in distances measured in two different resulting point-cloud models, a georeferenced (G) model and a non-georeferenced (NG) model. The

purpose of this first objective was to investigate the effect of georeferencing in distances measured within final 3D georeferenced point-cloud models. Numerous pairs of points are identified in both models (one point on each of them) and distances between them are measured from each model, i.e., from the G and the NG models. These corresponding distances were then compared to determine their discrepancies and calculate their associated statistics. (b) The generation of virtual 3-D point-cloud models of existing conditions at a commercial site (El Sombrero Restaurant) at 879 Buckhead Dr., in Statesboro, GA. This site covers an area of approximately five acres. It has two buildings, two parking lots, a large grassy landscape area with a small pond for catching rainwater. The resulting models are to be used as virtual surveying tools to assist the property owner in the design of future additions and modifications of the existing structures and parking/garden/lawn areas. Figure 1 shows an aerial view, extracted from Google Maps, of the modeled five-acre site.



Figure 1. Five-acre site modelled in this project (extracted from Google Maps)

Accuracies associated with surveying instruments and its procedures are usually made available by their manufacturers. However, it is known that those accuracies not always coincide with the ones obtained in the field, where multiple diverse conditions may exist (Fan, 2015). Additionally, after certain post-processing techniques, such as georeferencing, accuracy could also be degraded.

This article focuses on errors introduced after georeferencing virtual 3D point-cloud models. The georeferencing process consists of registering the model into a selected system of reference. This system could be defined by three or more fixed ground points with known coordinates in the selected system. It is also necessary that virtual images of the ground points exist in the point-cloud model. These ground points must have been previously scanned and acquired into the point cloud. After that acquisition, the authors have two sets of ground points. One is the real set, with real coordinates in the selected system of reference and the other is a virtual set within the point cloud. This second set possesses coordinates in the system used by the non-georeferenced point cloud. It is expected that both sets have the same relative positions and distances between their points. If that were the case, the NG model could be translated and rotated to make its virtual set of ground points coincide with the real set. However, most likely there will not be a perfect geometric match between both sets. So, the point cloud will rotate and translate to attain the least square error derived from this matching. After this least erroneous matching is attained, the virtual set of points acquire coordinates close to those originally contained by the real set of points. In other words, this introduces a small error in the now georeferenced virtual point cloud.

This work attempts to quantify this error for a given relatively large case resulting in the virtual 3D model of an area covering approximately 5 acres and containing two buildings.

Several analyses of accuracies in 3D point-cloud models have already being completed by the first two authors (Maldonado et. al., 2015; 2018, 2020) and numerous other investigators. In 2008, an article on the evaluation of 3D laser scanning for highway construction applications was published (Slattery et al., 2008). It concluded that "Preliminary results indicate it is feasible to use laser scanning technology to accurately map terrain prior to road construction." A study by Newsome (2016) presented and analysis of discrepancies between measurements obtained from the point cloud versus those attained with more accurate field instruments. The vertical accuracy of generated digital terrain models was assessed by Salach et al. (2018). However, there have only been a few studies analyzing georeferencing and registration errors. A particular work considered errors in target-based georeferencing and registration via terrestrial laser scanning (Fan et al., 2015). It indicated that registration and georeferencing errors presented by scanner manufacturer's software are incompetent measures of the actual registration and georeferencing errors in terrestrial laser scanning. This emphasizes the need to study and analyze errors generated by georeferencing processes.

## Instruments and Methodology

Two key instruments were employed in this project. One is a terrestrial laser-based twelve-second scanner, the Leica Geosystems ScanStation C10. This device was employed to capture spatial and color coordinate data of the target geometry. The other instrument is a highly accurate one-second robotic total-station device, TCRP 1201+, from the same manufacturer. The latter was employed to stablish accurate ground points, via closed traverse procedures, for georeferencing purposes. A previous article, by two of the current authors, includes a table comparing the main characteristics and capabilities of both instruments (Maldonado et al., 2018). The employed methodology consisted of the following steps:

1. _Benchmarks_: Seven strategic ground points were materialized with steel nails within the selected site. Their accurate relative positions were determined by completing two four-sided closed-traverse procedures. Each polygon shared a common vertex with the other. The robotic TCRP 1201+ instrument was employed for measuring required side lengths and horizontal angles. The coordinates of the involved ground points were obtained in a horizontally shifted Georgia East State Plane Coordinate System (SPCS), which uses US Survey Foot as its main unit of length. The horizontal shift was applied to reduce from 9 to 6 the number of significant digits assigned to the horizontal coordinates. This shift does not affect the accuracy of the present analysis because three digits after the decimal point were still considered in the coordinate values of all acquired points. The elevations of each fixed ground point (nail) were obtained via a regular, but accurate, differential leveling approach.
2. _Scanning_: The C10 instrument was used to scan the entire area, approximately 5 acres. This required to complete 64 overlapping individual scans, each at medium resolution (points separated by 1 mm at 10 m from the scanner). Each scan collected a minimum of three fixed points (target points) common to their neighboring scans. For this, two different white spherical targets, 6-inch diameter and 9-inch diameter, were employed. The larger ones were especially useful when the targets were located at longer distances from the scanner. The exterior of all existing buildings, and other surfaces, such as parking, grassy, and pond areas, were successfully captured by the 64 scans.
3. _Registration_: The referred common target points were employed to stitch (register) adjacent scans to each other via an automatic target-based registration approach, using Leica's Cyclone software package. Since several redundant targets were collected, this allowed for disabling a few selected ones, those causing larger registration errors. After attaining a maximum global registration error

of 0.033 ft (1 cm), no further targets were disabled, and the registration was considered successful. At this point a final 3D point-cloud model of the whole area was attained. This was the NG model. Its global system of reference was an arbitrary one. It coincided with the local system of reference of the first scan in the 64-scan list. It had a vertical z axis (elevation) and two orthogonal horizontal axes, x and y. The horizontal x axis coincided with the horizontal direction of the scanning telescope of the first scan in the list, at the instant in which that scan was started. The fully stitched (registered) model is observed below in figure 2, where the numerous common target points are indicated in yellow.
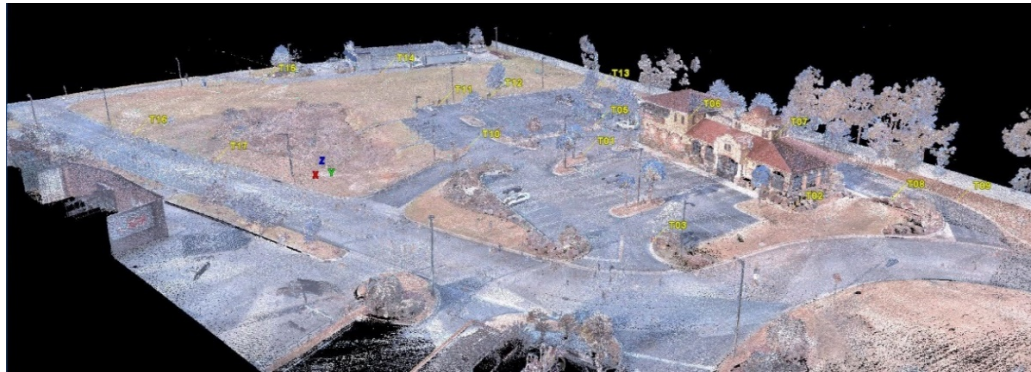


Figure 2. Non-georeferenced virtual 3D point-cloud model of the entire commercial site.

4. *Georeferencing*: Most vertices of two traverses were employed in this procedure. The North traverse has vertices T01, T02, T09 and T06. The South traverse has vertices T01, T12, T15 and T16. Only the coordinates of 5 of these 7 vertices were employed (T01, T02, T09, T06, T15 and T16) to georeference the NG model and attain the final G model in the above mentioned shifted SPCS. This was done using Leica's Cyclone software. It simply consisted of adding one extra scan file containing only the coordinates of those five points in the shifted SPCS. This file was positioned as the first one in the list containing all other 64 scans. Consequently, its system of referenced became the dominant one and was adopted as the global system of reference in the final G model. During this process, the software reported a mean absolute error of 0.021 ft (0.64 cm) with a minimum error of 0.011 ft (0.34 cm) at T09 and a maximum error of 0.031 ft (0.94 cm) at T01. That is, this georeferencing process did not introduce an error larger than the one attained when registering all 64 scans, 0.033 ft (1 cm).

5. *Data Post-Processing*: A total of n=135 sample points were selected in the NG model and their corresponding x-y-z coordinates (in the arbitrary system) were recorded in a Microsoft Excel spreadsheet. Similarly, the same 135 sample points were selected in the G model and their coordinates (in the shifted SPCS) were recorded in the same spreadsheet. The number 135 was selected since a minimum of 100 sample points (Sheskin, 2011) are to be considered to attain a 99% confidence in the calculated standard deviation, SD, of position discrepancies within the following confidence interval: [0.84*SD, 1.22*SD]. Even though 100 points were the target sample size, 135 points were originally collected to assure that 100 of them will still be considered if a few were to be discarded as outliers due to human errors while students were completing these tasks for their first time. Within these 135 points, eight were selected, almost randomly, as center points (CPs). Only two considerations were made to select them. First, they were to be distributed along different areas of the whole 5-acre project site. Second, they were easily identified in both, the G and NG models. These CPs were points #11, #25, #43, #58, #86, #127, #129, and #131. Then,

distances were measured from each CP to the remaining 134 points. That is, a total of 1072 pairs of comparable distances were extracted from the G and NG models (i.e., 1072 distances from each of them). However, as expected, it was soon evident that some of the original 135 points were not properly collected and they were producing outlying data. Seven points were clearly identified as outliers and were removed from the original set. So, the sample points were reduced to n=128 and the number of calculated distances from each CP was 127. Accordingly, the total number of comparable distance pairs was reduced to 1016. All statistics of discrepancies in distances from both models were calculated with and without outliers and are summarily presented in the Results section of this article. Statistics were determined locally, using the distances associated to each CP. This resulted in 8 local sets. Additionally, statistics were calculated globally, employing all available distances, following two different approaches: (1) considering all distances, and (2) averaging the local statistics associated to each of the eight CPs.
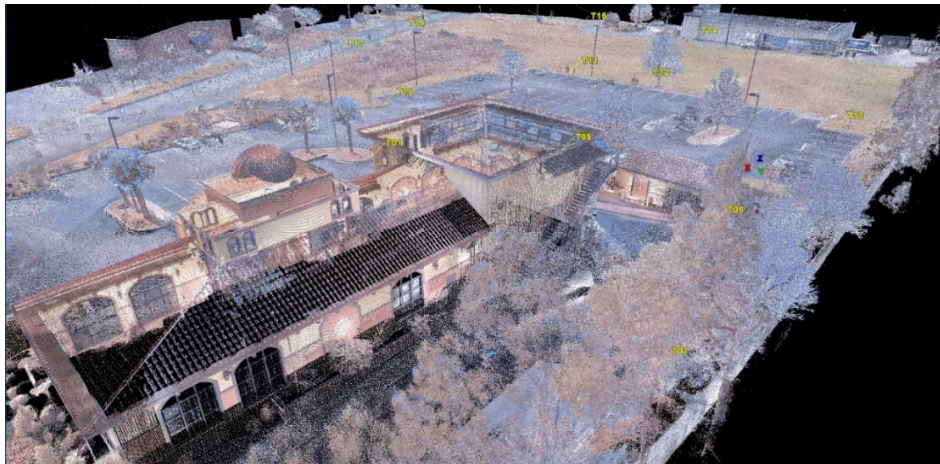


Figure 3. Façade of the main building.



Figure 4. Back side of the main building.

6.  *Noise Cleaning*: This procedure consists of removing pedestrian and vehicular noise from the final model. Figures 3 and 4 show screen captures of the cleaned georeferenced final point-cloud model near the main restaurant building. Even though this task is presented last in this list, it is advisable to complete it earlier before scans are stitched (registered) into the final model. Once each scan is

cleaned, they should be properly stored. This will allow to modify the final model (i.e., add one extra scan or georeference it into another reference system) without the need to clean the scans again.

# Results

In this distance discrepancy study, the number of considered distances is relatively large, more than 1000. They range from 4 to almost 700 feet. Table 1 shows the resulting Standard Deviations (SD) and Root Mean Square Values (RMSV) of the discrepancies in distances between the G and the NG models. The sample and population SDs were calculated using Microsoft Excel's STD S and STD P functions, respectively. Their magnitudes are very similar. This is due to the large number of sample points involved in this study. Additionally, the RMSVs are also close to the magnitudes of the SDs. This is since the mean values of the discrepancies are all close to zero. RMSVs approach the respective SD values when means approach zero. In particular, the global mean value of all discrepancies in distances, with outliers is -0.004 ft (-1.2 mm), and without outliers is -0.003 ft (-0.9 mm).

Table 1

*Comparison of local and global Standard Deviations and Root Mean Square Values of distance discrepancies between georeferenced and non-georeferenced models, with and without outliers. All quantities are in **US feet**.*

| Local Center Points | WITH OUTLIERS | | | WITHOUT OUTLIERS | | |
|---|---|---|---|---|---|---|
| | SD Pop. | SD Sample | RMSV | SD Pop. | SD Sample | RMSV |
| 1 (#11) | 0.110 | 0.111 | 0.110 | 0.034 | 0.034 | 0.034 |
| 2 (#25) | 0.398 | 0.400 | 0.400 | 0.042 | 0.042 | 0.043 |
| 3 (#43) | 0.240 | 0.241 | 0.241 | 0.034 | 0.035 | 0.035 |
| 4 (#58) | 0.363 | 0.364 | 0.363 | 0.042 | 0.042 | 0.044 |
| 5 (#86) | 0.269 | 0.270 | 0.269 | 0.034 | 0.035 | 0.036 |
| 6 (#127) | 0.761 | 0.764 | 0.763 | 0.034 | 0.034 | 0.034 |
| 7 (#129) | 0.204 | 0.205 | 0.204 | 0.033 | 0.033 | 0.040 |
| 8 (#131) | 0.133 | 0.134 | 0.133 | 0.034 | 0.034 | 0.034 |
| **Average of all Local Sets** | **0.301** | **0.311** | **0.310** | **0.036** | **0.036** | **0.037** |
| **Global Values** | **0.367** | **0.367** | **0.367** | **0.037** | **0.037** | **0.037** |

Table 1 presents SDs and RMSVs for distances calculated with and without outlying points. As expected, when distances are calculated without outlying points, the SDs and RMSVs of their discrepancies are much less than those with outliers. Each row of table 1 shows the statistics of the set of distances associated to one of the eight center points. That is, each row shows the local statistics of each CP. The second to last row of table 1 presents the averaged values of the 8 sets of CPs. The last row of this table shows the statistics corresponding to all discrepancies in distances involved in this study. It is noted a relatively small difference in the values presented in the last two rows, especially in those without outliers. In general, these numbers indicate that the resulting accurate models, including the accurately georeferenced one, produce only small discrepancies in distances measured from both. The overall value of the SD of those discrepancies is 0.037 ft (1.1 cm) when distances range from 4 to about 700 ft (1.2 - 213.4 m).

Figure 5 shows the discrepancies in all considered 1016 pair of distances calculated from within the G and NG models. Additionally, in this figure it is observed that most of the discrepancies is concentrated in the interval [-0.10 ft, +0.10 ft] which is expected when the SD = 0.037 ft and 3*SD = 0.11 ft.
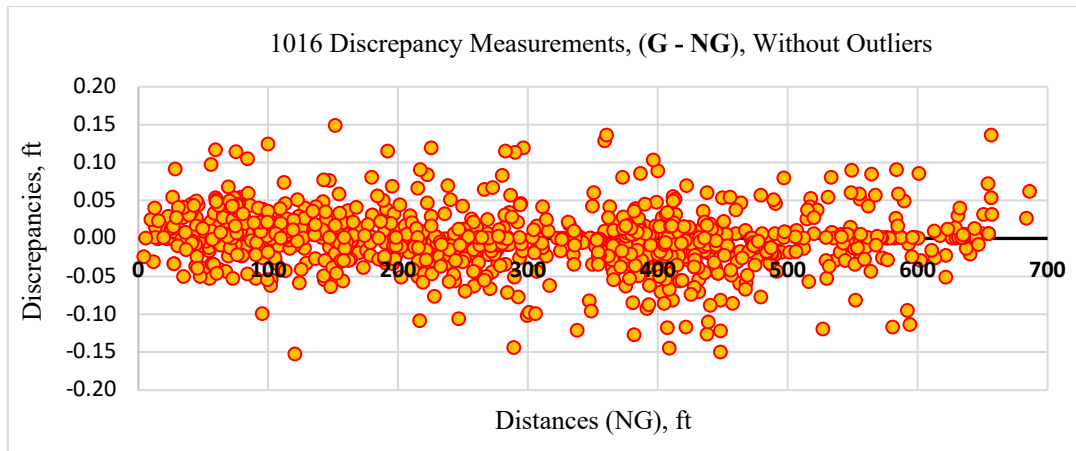


Figure 5. Distance Discrepancies between Georeferenced and Non-Georeferenced Models

Figure 6 shows the discrepancies in distances as a percent of the measured distance. This percentage is calculated as (G – NG) * 100 / NG, where G is the distance from the G model and NG is the distance from the NG model. Since the discrepancies are uniformly distributed along the lengths of the measured distances, it is observed that the percent discrepancy is much larger in the shorter distances.
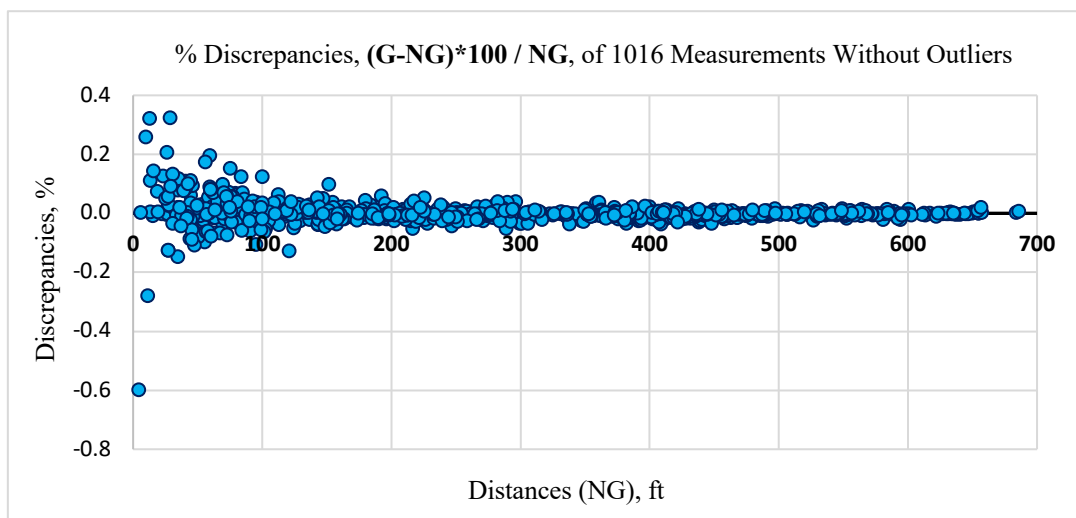


Figure 6. % Distance Discrepancies between Georeferenced and Non-Georeferenced Models

## Conclusions and Closing Remarks

The completed comparative statistical analysis in measured distances leads to the following concluding remarks:

- The resulting non-georeferenced model was accurate. It showed a low overall error, 0.033 ft (1.0 cm), as reported by the Leica's Cyclone software after registering the 64 individual scans.

- The coordinates of the benchmarks used for georeferencing purposes were accurate as well. This is concluded from the fact that the angular errors of closure were zero (0) in both traverses and the attained longitudinal precisions were 1:21,465 and 1:29,477 for the North and South traverses, respectively.

- The resulting georeferenced model was also accurate. This is inferred by the maximum error of 0.031 ft (0.94 cm) attained during the georeferencing process at one of the employed 5 benchmarks (T01). This maximum georeferencing error is even a bit less than the one attained while stitching all 64 scans into the NG model 0.033 ft (1.0 cm). The mean absolute error of this georeferencing process was reported by the Leica's Cyclone software as 0.021 ft (0.64 cm).

- All involved points in this analysis were selected in a manner that they were easily identified within the non-georeferenced and georeferenced models. For example, they were the center of small stickers placed on the walls of the buildings or they were easily identifiable within the models, such as the center of manholes or the tip of traffic arrows marked on the pavement. However, while acquiring the coordinates of these points from the virtual point-cloud models, an error could occur. This potential error has not been precisely quantified in this study. Nevertheless, after collecting numerous points in this and several other previous studies, it is estimated that such error could reach about 1/8 inch in the position of each point within dense clouds with approximately 64 points (8*8) per square inch. This could lead to errors in the calculated distance discrepancies, amounting from -0.5 to +0.5 inch. Its estimated RMSV would be approximately 0.5 inch / 3 = 0.17 inch or 0.014 ft (0.43 cm). This estimated discrepancy is smaller than the one captured by this study, RMSV = 0.037 ft (1.1 cm), but still ~1/3 of its magnitude. A more accurate estimation of this error could serve as the main objective of a future study. This coordinate acquisition error will depend on the point density of the involved clouds.

- The 1016 considered distances (without outliers) ranged from 4 to almost 700 ft in length. The Standard Deviation and Root Mean Square values of all 1016 discrepancies in distances, extracted from the non-georeferenced and the georeferenced models, are all equal to 0.037 ft (1.1 cm). As expected, all mean discrepancy values were very close to zero. Consequently, the corresponding RMSVs were almost identical to their associated SD values.

- The RMSV and SD parameters assist in the interpretation of the attained results via the six-sigma concept. Specifically, the six-sigma interval for the reported discrepancies is limited by the following magnitude: $\pm 3\sigma = \pm 3*SD = \pm 0.111$ ft ($\pm 3.4$ cm). According to the well-known 68-95-99.7 Gaussian rule, the six-sigma interval [-0.111 ft, +0.111 ft] contains almost the whole population (99.7%) of distance discrepancies between both models. In other words, for this range of distances, from 4 to 700 ft, almost none of them (actually, only 0.3% of them), when measured within the georeferenced model, produced discrepancy magnitudes larger than 0.111 ft (3.4 cm) with respect to the non-georeferenced ones. This conclusion is valid for the case presented in this

study and could be extrapolated to other cases employing similarly accurate georeferenced and non-georeferenced models.

## Acknowledgments

## References

Fan, L., Smethurst, J.A., Atkinson, P.M., and Powrie, W. (2015). Error in target-based georeferencing and registration in terrestrial laser scanning. Computers & Geosciences, Vol. 83, Oct. 2015, p. 54-64.

Maldonado, G.O., Maghiar, M., Jackson, N.M., Garrett, D.M. and Givens, K.E. (2015). Comparison of Building Measurements Acquired via Laser-Based Scanner and Modern Total Station. 51st ASC Annual International Conference Proceedings.

Maldonado, G.O., Maghiar, M., Tharp, B.W. and Patel, D.B. (2020). Point-Cloud Models of Historical Barns – Spatial Discrepancies of Laser Scanning versus Robotic Total Station. 56th ASC Annual International Conference Proceedings, Vol. 1, p. 292-300.

Maldonado, G.O., Maghiar, M., Van Stan, J.T., Peart, M.D. and Mesta, D.C. (2018). Discrepancies in Measurements of a Complex Street Intersection – Laser Scanner vs Accurate Total Station. 54th ASC Annual International Conference Proceedings, p. 371-378.

Newsome, S.R. Jr (2016), Discrepancy Analysis Between Close-Range Photogrammetry and Terrestrial LiDAR". Digital Commons, Georgia Southern University, Electronic Theses and Dissertations. Document 1423.

Salach A., Bakula, K., Pilarska, M., Ostrowski, W., Gorski, K. and Kurczynski, Z. (2018). Accuracy of Point Clouds from LIDAR and Dense Image Matching Acquired Using UAV Platform for DTM Creation. International Journal of Geo-Information, 7, 342.

Sheskin, D. J. Handbook of Parametric and Nonparametric Statistical Procedures, Fifth Edition (2011). CRC Press I, LLC. ISBN 10: 1439858012, ISBN 13: 9781439858011.

Slattery, K.T., Slattery, D.K., and Gordon, C. (2008). Evaluation of 3-D Laser Scanning for Highway Construction Applications. 44th ASC Annual International Conference Proceedings.