U.S. Geological Survey (USGS) - National Geospatial Program (NGP)

and the

American Society for Photogrammetry and Remote Sensing (ASPRS)

Summary of Research and Development Efforts Necessary for Assuring Geometric Quality of Lidar Data

U.S. Department of the Interior
U.S. Geological Survey
American Society for Photogrammetry and
Remote Sensing

8/22/2015

Approved by:

____

, Date

ASPRS,

Executive Director

Introduction

Lidar has become the technology of choice for collecting three dimensional data of the earth's surface. Standards that ensure geometric quality and enable metric measurement from lidar data must be applied consistently across the industry. This also requires that quality assurance processes are applied consistently across the industry in a transparent manner. The goal of this document is to highlight further research and developmental efforts that are necessary for comprehensively quantifying the geometric quality of lidar data. These required research efforts are listed at the end of each section, and are also summarized at the end of this document.

A data set can be said to have good geometric quality if geometric measurements of identical features, regardless of their position or orientation, yield identical results. Good geometric quality indicates that the data are produced using sensor models that are working as they are mathematically designed, calibration has been adequately performed, and acquisition processes are not introducing any unforeseen distortion in the data. Geometric quality (measured in terms of relative accuracy of overlapping swaths) coupled with absolute accuracy assessments made using a limited number of Ground Control Points (GCPs) is required to assess the geometric accuracy of the entire point cloud.

Towards these objectives, the ASPRS and the U.S. Geological Survey formed a Calibration/Validation Working Group (WG from here), operating under the aegis of the ASPRS Lidar Division and the Airborne Lidar Committee, to recommend processes that must be followed to assure the geometric quality of lidar data. While a number of lidar quality metrics have been introduced within the broad community of lidar researchers and practitioners, the WG has limited its recommendations to the QA/QC of geometric quality of lidar data.

The WG recommends that QC processes associated with testing the quality of calibration/boresighting of lidar instruments be developed and standardized for the industry. The geometric quality of data, which includes the quality of calibration, most easily manifests itself in the overlapping regions of the various swaths of data. Therefore, procedures to test the relative accuracy of the data, i.e. testing the accuracy within the swaths of the data are recommended to be implemented in a transparent and standardized manner. The WG has made progress in developing metrics, which are detailed in the draft "ASPRS Guidelines of Measurement of Inter Swath Geometric Quality of Lidar data".

The WG also recommends that affordable absolute accuracy measurement techniques for assessing the three dimensional accuracy of point clouds be researched and developed. On most occasions, only the vertical accuracy is tested, with the horizontal accuracy assessments ignored due to lack of easy means of measurements and standardized metrics. This level of testing is adequate to assess low resolution point cloud data meant for generating Digital Elevation Models. However, for 3D analytical applications, the quality of all three dimensions of the point cloud is essential. It is important, therefore, to develop cost effective means of assessing the three dimensional accuracy of lidar point clouds.

The WG also recommends that the data acquisition procedures follow sensor manufacturer recommended procedures for instrument calibration and boresight angle determination while performing data acquisition. This includes data driven determination of optimal boresight parameters using the lidar instrument's sensor model. Since most of the lidar instrument sensor models are proprietary, the WG cannot recommend any specific methodology.

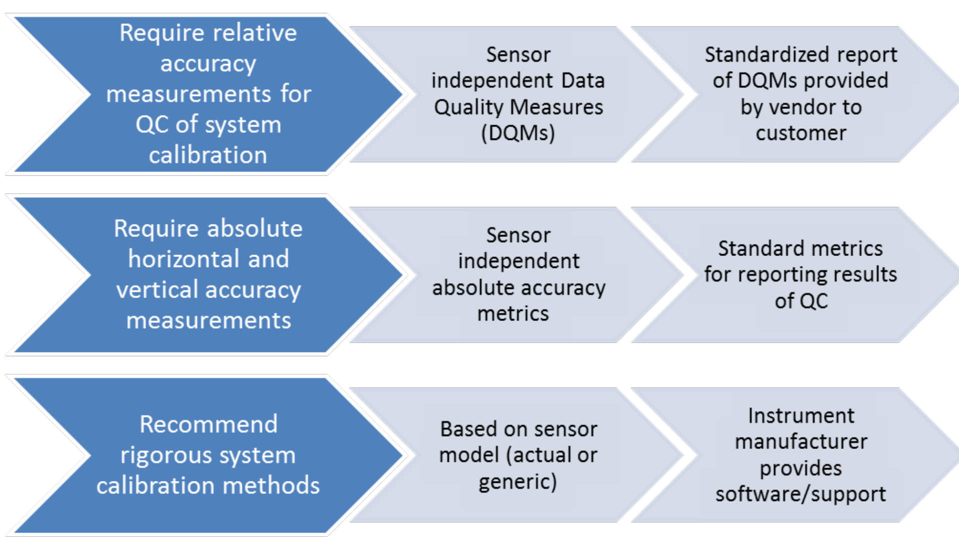

Fig. 1 Framework of recommended QA/QC guidelines

In most cases, lidar data over a project area collected in swaths. The geometric quality can be assessed using measurements listed below:

1. Intra-Swath point cloud consistency measurements.
2. Inter-Swath point cloud consistency measurements.
3. Absolute accuracy measurements using Ground Control Points, Surfaces, 3D targets, or other means.
   - Effects of errors in determination of lidar trajectory
4. Inter-Project consistency of features measurements. In case of national projects like 3DEP, consistency of features collected between adjacent projects, such as streams, roads, etc. must also be quantified.

The USGS-ASPRS led WG has had great success in defining and developing Inter-Swath Data Quality Measures (DQMs) for measuring the consistency of features found in the overlapping regions of different swaths of lidar point cloud. The effort has been supported by The U.S.USGSS's National Geospatial Program (NGP) and the Land Remote Sensing (LRS) program. The process is presented in a detailed manner in the draft "ASPRS Guidelines on Inter-Swath Geometric Accuracy and Quality of Lidar Data".

Lidar datasets processed without standardized and theoretically correct QA/QC processes cannot be easily integrated into scientific investigations, and such QA/QC assessment techniques may also be required for other non-scientific (including legal) purposes. In this document, the research and developmental efforts that are necessary to provide solutions to the remaining problems are briefly described. It is hoped that the template provided by the success of the USGS-ASPRS WG can be replicated to solve these problems.

QC of Lidar System Calibration and Data Acquisition using Relative Accuracy Measurements

Introduction

Calibration is often called for with a new instrument. However, instrument error can occur due to a variety of factors such as environment, electrical supply, addition of components, process changes, etc. Therefore, calibration should be performed, at a minimum, as directed by the instrument manufacturer either periodically, or based on usage. Calibration should also be performed when an instrument has had a shock or vibration (which potentially may put it out of calibration), any time it is reinstalled on the host platform (e.g., after aircraft maintenance), or whenever observations appear questionable. In some cases, calibration of a partial set of sensor system parameters can be performed in situ (field tests or self-calibration), and it is a standard practice among lidar data providers to perform such a calibration of boresight parameters.

The users of lidar point cloud data require quantifiable means to be assured of the calibration of the lidar data acquisition system. This quality of calibration of lidar systems can be quantified without requiring externally referenced data. These means are Intra Swath analysis of data and Inter Swath analysis of data.

Intra Swath Analysis

The Intra Swath analysis of data quantifies the level of noise present in the lidar sensor. For example, consider a situation where a flat/horizontal building roof can be detected in portions of the point cloud. In the current specifications (Heidemann 2014), the intra swath quality of data is in general defined as the repeatability of measurements over a smooth surface, and in particular, repeatability of elevation values over a flat surface.

Inter Swath Analysis

Lidar data for large projects are collected by flying swaths, with overlap between portions of the swaths to prevent data gaps or to ensure the required level of point density. The quality of calibration of a lidar instrument and data acquisition most readily manifests itself in these overlapping areas. The USGS-ASPRS WG has defined several measures to quantify the Inter Swath accuracy or relative accuracy of lidar data. These measures are termed Data Quality Measures (DQMs).

Table 1 Data Quality Measures (DQMs) or inter-swath goodness of fit measures

Nature of surface	Examples	Data Quality Measures (DQMs)/Goodness of fit measures	Units
Natural surfaces	Ground surface, i.e. not trees, chimneys, electric lines etc.	Point to natural surface (tangential plane to surface) distance	Meters
		Point to surface vertical distance	Meters
Man-made surfaces	Roof planes	Perpendicular distance from the centroid of one plane to the conjugate plane	Meters
	Roof edges	Perpendicular distance of the centroid of one line segment to the conjugate line segment	Meters

The WG have also defined the process of measuring these DQMs and defined how these could be summarized to robustly quantify the lidar point cloud's geometric quality. More details can be found in the draft "ASPRS Guidelines of Measurement of Inter Swath Geometric Quality of Lidar data". The WG recommends that the Inter Swath quality may be reported in three parts

a) The sampled locations where DQMs are measured be categorized as functions of slope of terrain: Flat terrain and slopes greater than 20 degrees. For higher slopes, the DQM errors are further categorized as belonging to slopes along flight line and across flight line

b) For flat areas:

   a. The distance of each sampled location from the mean centerline of overlap is measured (Dco)

   b. The discrepancy angle (dS**i**) at each sampled location, defined as the arctangent of DQM error divided by Dco, is measured

c. The summary statistics for the discrepancy angles is recorded. For high quality data, the discrepancy angles must be close to zero, and the standard deviation must also be low.

The Discrepancy Angle (DA) is a measure of systematic errors present in the lidar point cloud, and is a direct indicator of the geometric quality of the data.

Recommendations for Further Research

To quantify the noise level in the data, the WG recommends research and development efforts towards answering the following questions:

- How should the "noise" be mathematically defined?
- How should measurements be made to verify noise?
- How should these measurements be summarized (in terms of mean, median, standard deviation, and spatial distribution, etc.)?

On Inter Swath analysis, the WG further recommends that government agencies, academia and the industry work together towards determining acceptable solutions to the following questions:

- Are current thresholds for relative accuracy valid?
- Are the thresholds measurable?
- Should the thresholds be different for different terrain?
- What should be the thresholds for the discrepancy angles?

It is also recommended that a comprehensive error library consisting of systematic errors found in lidar data and their source be maintained by the ASPRS.

QC of Lidar System Calibration and Data Acquisition using Absolute Accuracy Measurements

Introduction

The absolute accuracy of lidar data ideally must consider the accuracy of all three dimensions of the data. Due to the legacy of lidar data being used primarily for Digital Elevation Models, only the vertical accuracy of the data are generally reported. Another reason could be that in the initial years of the development of the technology, the point spacing of the data was low, and hence it was difficult to identify planimetric features from the data with certainty.

However, the applications of lidar data have grown tremendously since the early days. It is not limited to being an input to DEMs, and the data are used in solving many problems which require their accuracy be known in all three dimensions. Also, with the increasing use of return intensity data, development of 3D feature extraction algorithms, increasing point density, as well as increased computational resources, it is possible to measure the accuracy of all three dimensions of the data. Currently, the vertical accuracy of the data is measured on mostly flat, clear and open areas. However, these measurements do not provide any quantifiable estimate of the horizontal accuracy of the data set. However, there are very few studies that have attempted to provide standardized processes to measure the absolute accuracy. Some of these are mentioned below. Almost all of them use 3D targets specifically designed to take advantage of the 3D nature of the data, or the intensity of the return.

However, deploying targets can be expensive. Few, if any, of the studies have focused on the cost of deploying such targets, which are assumed to be high based on current costs of obtaining ground control points.

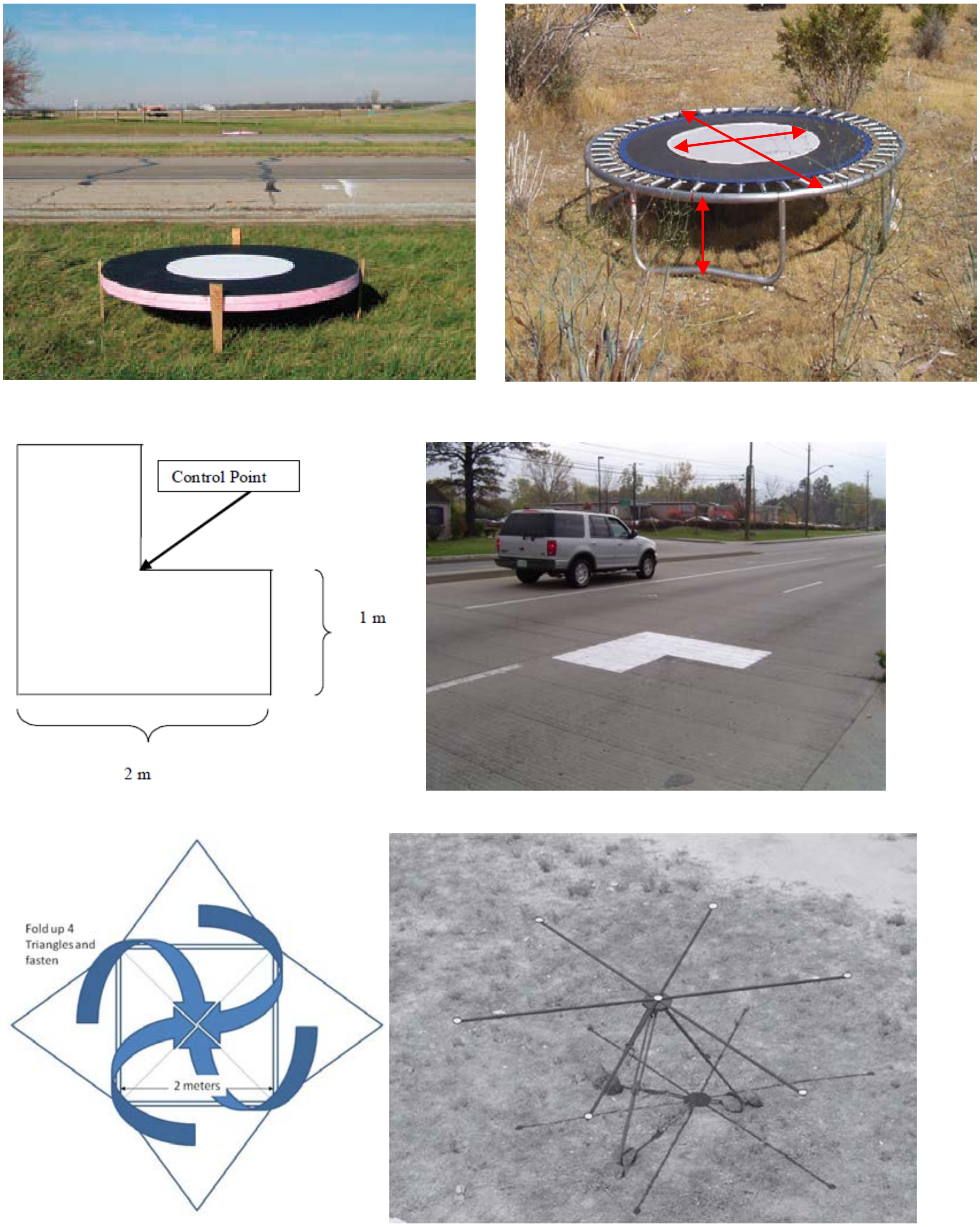

Figure 2 Targets used in planimetric and vertical elevation assessment of lidar data. First row shows targets developed by the Ohio State University for the Ohio Department of Transportation; Second row depicts targets developed for Indiana Department of Transportation and Third row depicts a 3D pyramid shaped target designed by Jason Stoker of USGS and retro reflector based targets designed by the NGA

Recommendations for Further research

The WG recommends that the following studies be undertaken:

- Cost of deploying targets as a percentage of total project cost, project size, terrain conditions, etc.
- Research on estimation of absolute three-dimensional accuracy of lidar data using latest technologies, including mobile lidar, static lidar, and UAV data.

QC of Lidar data for National Lidar Projects such as 3DEP

Introduction

National projects including the 3D Elevation Program (3DEP) are envisaged to support national and pan-regional scientific and engineering goals. Many of the lidar acquisition projects have scientific goals that are regional or sub-regional in scope. Their geometrical quality assessments are also limited in their spatial extent. Lidar data for scientific projects such as river systems analysis, carbon sink assessments, interstate road networks, watersheds, etc.; require data assessments that are similar in scope, i.e. across projects. example,initially . In practice, individual projects are executed by USGS with other federal and local partners in a more ad-hoc (spatially) manner.

Many features such as streams, roads, forests etc. will span multiple collectionsor projects. It is important to the scientific community that these features are represented in a consistent manner by the lidar data, even though the data may be acquired by different vendors, using different sensors, and at different times. The consistency between adjacent datasets must be quantified and documented in the metadata.

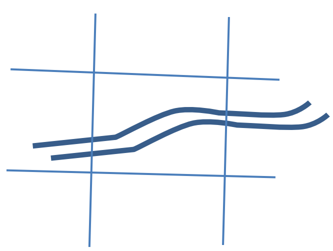

Figure 3 Illustration of pan regional features of scientific importance that require quantification of inter project geometric alignments

Recommendations for Further Research

The WG recommends that government, academia and the industry stakeholders research and develop methods for inter project lidar data consistency. The specific questions that are required to be answered include

- How should inter project lidar data consistency be mathematically defined?
- How should the data consistency metrics be measured?

- How should data acquisition be specified to ensure consistency of features across spatial extents of different projects?
- Are there new technologies that can assist in enhancing inter project lidar data quality?

QA of Lidar data using sensor Model based Calibration Methods

Introduction

The most commonly used calibration is an *ad hoc* methodology where users observe discrepancies in adjacent swaths of data and sequentially tweak boresight parameters until an acceptable product is obtained. The traditional procedure is a manual adjustment process based on the operator's expertise. However, calibration parameters are often correlated, and since the process does not take these parameters into account, it is easy to misinterpret error source. These procedures also implicitly assume that the other calibration parameters hold their previously calibrated values, which may not be the case. These results depend on individual skill, experience of the operator, and are exhausting, time-consuming and expensive. Therefore, depending on the operators, there may be multiple calibration methods, and these methods may produce different results for the same inputs. These *ad hoc* methodologies cannot be termed true calibration because:

- They do not allow estimation of error in the generated data,
- Data sets generated by different users cannot be easily compared to each other.
- The processes cannot be traced to an accepted standard, i.e. the calibration is not conducted using a measure of higher accuracy.

It is recommended that these methodologies not be used unless it is absolutely necessary i.e. only be used when the use of rigorous and quasi-rigorous means are not possible.

Rigorous calibration methods

Rigorous calibration methods are based on determining parameters describing the sensor model completely. The parameters include the range and the trajectory files (including GPS/GNSS and IMU data), as well as the boresight parameters. Since many parameters associated with a complete sensor model are proprietary, software to perform rigorous calibration should be provided by the instrument manufacturer. Alternately, sensor manufacturers should be encouraged to make their sensor models open. Conceptually, the rigorous calibration process relies on minimizing discrepancies between tie features using redundant observations (data). These tie features should be features that can be reliably extracted from point cloud data. The most basic geometric feature that can be reliably extracted from a point cloud is a plane. Breaklines can be treated as the intersection of two planes, and points can be treated as intersection of at least three planes. Rigorous calibration processes have focused on extracting (sloping) planar features (e.g. roof planes, etc.) from the overlapping portions of point cloud, including those obtained from flying opposite directions, cross-strips and from different altitudes. Each point lying on the planar feature forms an observation equation (using the complete sensor model), and a constraint equation is used for points that are known to lie on the same plane. The parameters of the sensor model are estimated by minimizing the errors (e.g. least squares). Other methods of calibration that use extracted breaklines and points have also been successfully implemented (Habib et al., 2010)

Quasi-rigorous calibration methods

Quasi-rigorous calibration methods are very similar to the rigorous calibration methods. However, these methods do not have access to the actual sensor model, or all the parameters involved in the sensor model. These methods therefore use a generic sensor model substituting proprietary or unknown sensor parameters with generic sensor parameters. Examples of a generic sensor model include the Universal Lidar Error Model (ULEM), now renamed as the Generic Point Cloud Model (NGA, 2012). Some methods also use a limited set of parameters (e.g. using trajectory information, but not the range). However, the process of calibration is the same as the rigorous methods, i.e. these methods also use extracted planar features or combine extracted planar features to make measurements on derived features such as breaklines and points (Habib et al., 2010). The results of using such methods for calibration have been shown to approach a high level of accuracy (Habib et al., 2010).

Recommendations for Further Research

The Sensor Model based calibration methods are essential for the success of error propagation models that estimate the error present in the lidar data point cloud. The WG recommends that

- The possibility of expressing error in terms of generic sensor models be investigated further.
- Isolating the contribution of generic subsystems towards the total error budget be studied further. In particular, the errors introduced in the point cloud data due to errors in sensor trajectory must be further investigated

Summary

To summarize, the USGS-ASPRS WG has investigated various factors associated with the geometric quality of lidar data. The WG has noted that while the quality of lidar data have improved tremendously in the past few years, the QA/QC of these data are not standardized including the semantics, processes for measurement and reporting and meta data. Therefore, to ensure the geometric quality of that lidar data required for scientific (statistical error propagation and estimation) and non-scientific purposes (including legal), the WG has recommended several topics for research and development.

To quantify the noise level in the data, the WG recommends research and development efforts towards answering the following questions:

- How should the "noise" be mathematically defined?
- How should measurements be made to verify noise?
- How should these measurements be summarized (in terms of mean, median, standard deviation, and spatial distribution, etc.)?

On Inter Swath analysis, the WG further recommends that government agencies, academia and the industry work together towards determining acceptable solutions to the following questions

- Are current thresholds for relative accuracy valid?
- Are the thresholds measurable?
- Should the thresholds be different for different terrain?
- What should be the thresholds for the discrepancy angles?

The WG also recommends a comprehensive error library consisting of systematic errors and their sources be documented and maintained, for current as well as future lidar technologies.

Towards measuring the absolute accuracy of lidar data, the WG recommends that the following studies be undertaken:

- Cost of deploying 3D accuracy assessment targets as a percentage of total project cost, project size, terrain conditions, etc.

- Research on estimation of absolute three-dimensional accuracy of lidar data using latest technologies, including mobile lidar, static lidar, and UAV data.

Towards assessment of national lidar geometric quality and consistency, the WG recommends that research efforts are directed to answer the following topics:

- How should inter project lidar data consistency be mathematically defined?
- How should the data consistency metrics be measured?
- How should must data acquisition be specified to ensure consistency of features across spatial extents of different projects?
- Are there new technologies that can assist in enhancing inter project lidar data quality?