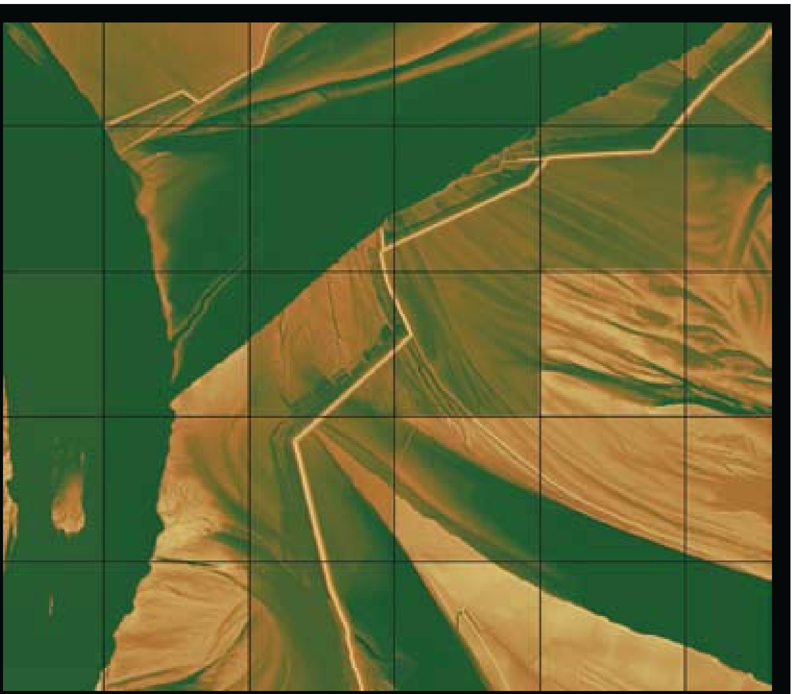
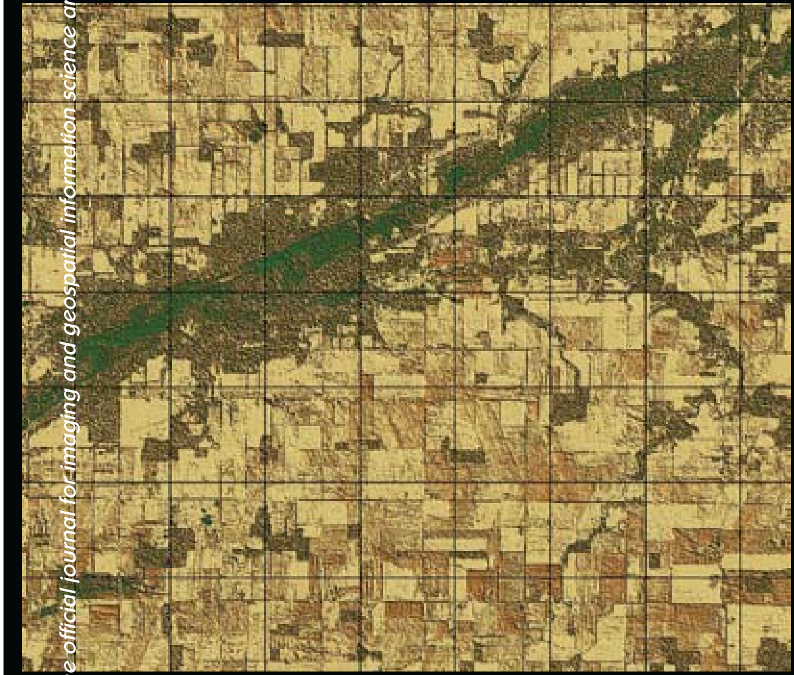
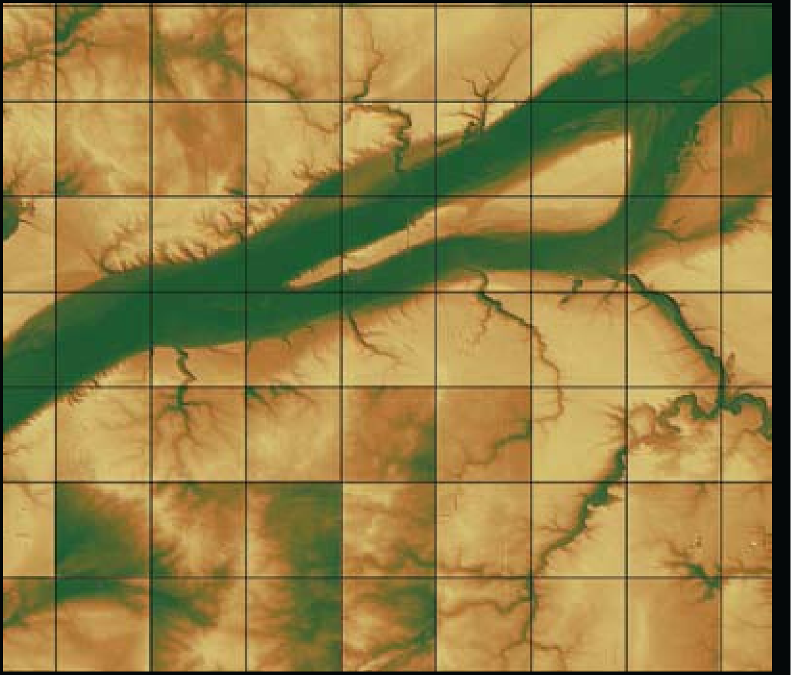
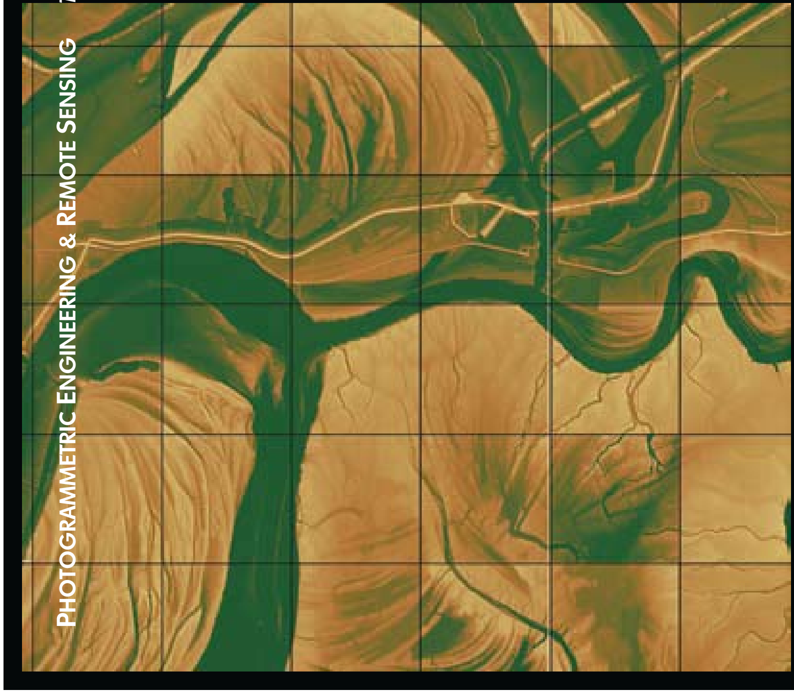
PE&RS

November 2011
Volume 77, Number 11









The official journal for imaging and geospatial information science and technology

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING

This month's cover shows elevation data from large-area lidar data collections as well as derived information products and orthophotos.

Three pairs of images show lidar elevation data with orthophotos (top and bottom pairs) or with a lidar intensity image (middle pair) for the same area. The upper pair (collected by Photo Science for the Vicksburg District U.S. Army Corps of Engineers) shows orthophotos and elevation data for a segment of the Mississippi River along "The Delta" emphasizing the mainline levee system that provides flood protection for areas along the Mississippi River. The middle images (collected in Michigan by Woolpert for USGS) show lidar intensity and elevation data depicting the complex data returns in flood plains and riparian zones and the refined elevation data products. The lower pair shows elevation data and orthophotos for the lower part of "The Delta" where levees and flood control structures protect the Yazoo Basin from backwater flooding that occurs when the Mississippi River is at high stages. These views of lidar and derived data products were generated in Topo Analyst, a new product from Spatial Information Solutions (SIS) of Starkville, MS, a spin out from Mississippi State University.

For additional information, contact Spatial Information Solutions (cgohara@spatialis.com) or visit http://www.spatialis.com/.

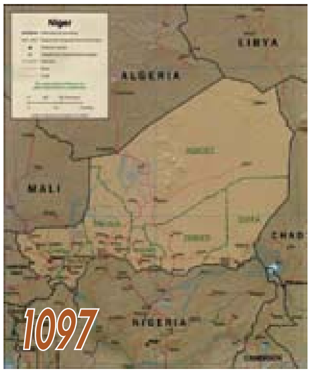
Highlight Article

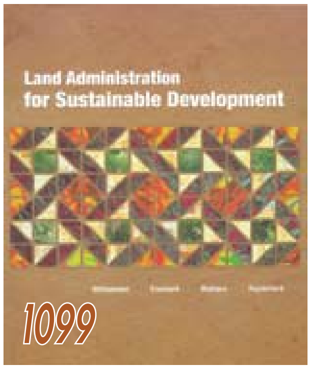
Feature Article

Columns & Updates

Announcements

Departments





PE&RS
November 2011 Volume 77, Number 11

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING
The official journal for imaging and geospatial information science and technology

Immediate electronic access to all peer-reviewed articles in this issue is available to ASPRS members at www.asprs.org. Just log in to the ASPRS web site with your membership ID and password and download the articles you need.

Cross-Walking “Lidar Guidelines and Base Specification” to Data Lifecycle Verification Approaches

by Charles O'Hara

Evolution in Aerial Lidar Data Accuracy Verification and Review

Rapid advances in aerial lidar technologies, growth in useful applications of the data from lidar acquisition projects, and increasing demand for improved elevation data and derived information products combine to drive the need for consistent and unified specifications as well as efficient, cost effective, and standardized lidar data accuracy and verification methods which are aligned with those specifications. According to version 13 of the U.S. Geological Survey National Geospatial Program “Lidar Guidelines and Base Specification”[1]

> “The U.S. Geological Survey National Geospatial Program (NGP) has cooperated in the collection of numerous lidar datasets across the nation for a wide array of applications. These collections have used a variety of specifications and required a diverse set of products, resulting in many incompatible datasets and making cross-project analysis extremely difficult. The need for a single base specification, defining minimum collection parameters and a consistent set of deliverables, is apparent.
>
> Beginning in late 2009, an increase in the rate of lidar data collection due to American Reinvestment and Recovery Act (ARRA) funding for The National Map makes it imperative that a single data specification be implemented to ensure consistency and improve data utility. Although the development of this specification was prompted by the ARRA stimulus funding, the specification is intended to remain durable beyond ARRA funded NGP projects.”

Lidar acquisition projects are evolving to include partnering of federal, state, county, and local participants. Partnering at multiple levels (federal, state, and local) on projects enables leveraging of funding resources while also requiring adjustments in the project acquisition data requirements and specification so that products deliver maximum benefits to the customers and the public and meet the combined requirements of the project partners. Differing participation can have significant impacts on the project specifications and requirements for their testing and verification. Major federal agencies such as the U.S. Geological Survey (USGS), the Federal Emergency Management Agency (FEMA), the U.S. Army Corps of Engineers (USACE), and the National Oceanic and Atmospheric Administration (NOAA) are providing leadership in developing specifications for lidar projects that are becoming increasingly a part of project requirements and specifications.

Vendor partnering for complex or large projects is also becoming commonplace. Projects that combine data from aerial rotary and fixed wing acquisitions of lidar and orthophotography, as well as terrestrial and mobile lidar, are delivering products that fuse data to include the best aspects of each sensor system involved. Such projects (Figure 1) will become more common as high resolution multi-scale, multi-source data are more effectively integrated with mapping and engineering software solutions. As the data become more complex, gathered from multiple sources, and collected to suit multi-scale applications, the methods for verifying the data must evolve to provide “best-practice” solutions that may be implemented across data sets, regardless of vendor specific systems or data types.

---

[1] USGS, 2010. “U.S. Geological Survey National Geospatial Program: Lidar Guidelines and Base Specification,” version 13 (http://lidar.cr.usgs.gov/USGS-NGP Lidar Guidelines and Base Specification, v13 ILMF.pdf).

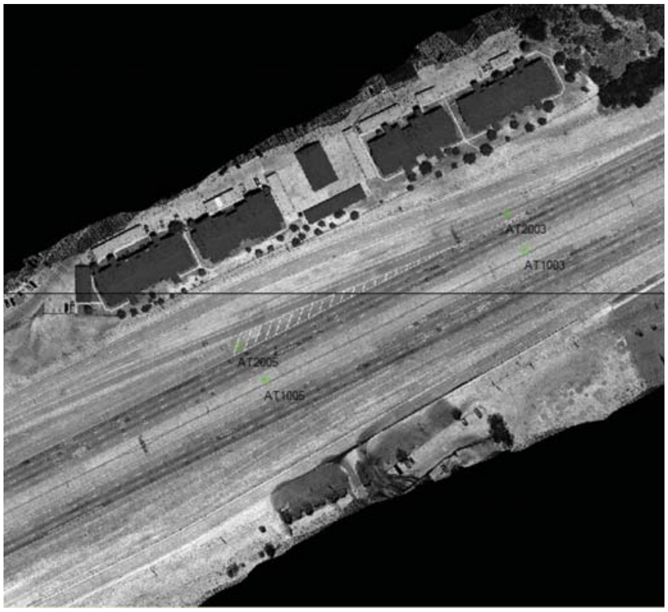


Figure 1. Surveyed checkpoint locations shown on intensity image derived from helicopter-based lidar acquired by Tuck Mapping as part of a multi-source lidar demonstration project conducted for the Texas Department of Transportation. (Data provided and used with the permission of Tuck Mapping)

Across the mapping industry, data vendors have made enormous investments in aircraft, lidar instruments, inertial measurement units and GPS navigation systems, as well as project planning and data production processing capabilities. Accompanying these investments in tangible assets, data producers have made further investments in both their production processes and personnel to develop and adhere to highly technical workflows, data processing, and internal quality control and assurance for delivering data products which will be acceptable to their customers. Despite the emergence of increasingly thorough guidelines and specifications for data and derived products, as well as vendors' production criteria, there exists a lack of commonly accepted best practices and standardized methods for verifying compliance to specifications.

Lidar – The Project Lifecycle Disconnect

A fundamental disconnect exists between lidar project data production and product delivery approaches, which inhibits the development of best practices that may be implemented across the project data production lifecycle. Lidar data and derived products are typically delivered at the end of the project; review, verification, and acceptance by customers typically await the delivery of final products. End-of-project delivery and verification approaches do not provide vendors with incentives for verifying, documenting, or delivering incremental products. At the same time, wide variation in the current ”state-of-the-practice” for end-of-project delivery and verification highlights the fact that there are no mutually agreed-upon best practices or commodity software tools for verifying and accepting final products let alone practices and tools which could be accepted as sufficient, best practices for authoritatively testing incremental products against specifications and requirements.

New guidelines and specifications include unprecedented attention to the accuracy, quality, correctness, compliance, and completeness of lidar point cloud data which include flightline swath data, tiled data products, and final classified tiled LAS data. The new specifications for lidar data and derived information products require testing of various data products developed during phases of the data project. As part of deliverable products verification, these tests are performed after-the-fact. Unless there are specific interim deliverables defined as project milestones, there are no mechanisms for engaging the producer and customer in ”real time” verification of incremental products as data flow through the project. These emerging guidelines imply that lidar data must be tested and verified at key project phases by vendors to ensure that data meet eventual delivery guidelines and specifications. Specifications present a framework for verifying products developed across the lidar project lifecycle process or face the prospect that downstream products will ultimately fail due to flaws that were not addressed in earlier acquisition, calibration, and processing phases.

This article suggests that there is a pressing need for an effective cross-walking that maps lidar data and derived information product guidelines and specifications to project data lifecycle phases. As an initial step in that direction, this article proposes a simple cross-walking that identifies key phases of a typical lidar project life cycle and corresponding verification steps that may be conducted as gateways for assuring that products will meet specifications. Additionally, a collection of recent lidar projects is highlighted in summary with a brief description of the projects, technical challenges, and aspects of accuracy or quality characteristics. Some key tests and methods are presented that offer potential to be considered as part of a collection of industry best practices for key phases of lidar data verification.

Cross-Walking Specifications to Lidar Data Project Lifecycle Phases

Verifying lidar products as an industry standard practice currently emphasizes testing finished end products. This approach can lack rigor if testing focuses only on final delivered products which include uncertainties more ideally addressed in earlier phases of the data product's lifecycle. An approach to lidar data verification is presented which cross-walks guidelines and specifications to key phases of a typical lidar project. Four basic phases are proposed for a lidar project in which testing and verification tasks are suggested as starting points to address ”gateway”' questions that may be answered prior to subsequent tasks in the lidar data project lifecycle.

Phase 1 – Flight Planning, Operations, and Lidar Data Acquisition

Gateway Question: Are flight operations complete; have data been collected fully covering the project area without voids within flightlines, gaps between flightline data swaths, and at the pulse spacing and density specified?

Preliminary planning and data acquisition produce data which are converted to LAS files. The following tasks are presented as a possible subset for which best practice methods may be developed and addressed during this phase to comply with guidelines and specifications:

- ***Collection Area:*** Determine whether the data sets from initial flightlines completely and adequately cover the defined project area as well as any designated buffer area.
- ***Overlap:*** Generate swath boundary file polygons. Intersect adjacent boundary files and quantify percentage of swath width in the overlap to determine whether adjacent flightline swath data sufficiently overlap (Figure 2).
- ***Spacing and Density:*** Quantify point cloud spacing and density and determine if data are spaced at sufficient density to meet project requirements (Figure 2).
- ***Gaps and Voids:*** Identify gaps in data, voids that must be filled, or areas where additional flightlines are required. Employ vector boundaries containing holes as well as pulse

continued on page 1076

continued from page 1075

count rasters to quantify the spatially varying distribution of pulses.

- ***LAS Compliance:*** Test LAS files for size, multiple returns, scan angle, intensity, compliance to standard, contents, and for completeness of required information for header, variable length records, and data values.

Benefits: Completing these tasks in a timely and efficient manner is vital so that lidar equipment may be effectively used and so that any gaps, voids, or areas missed in initial flightlines may be acquired in conditions as similar as possible to those present during initial collections.

Phase 2 – Flightline Swath Data Calibration and Early Processing

Gateway Question: Are acquired data calibrated for instrument and systematic errors and have flightline swath LAS residual errors been minimized such that relative error within and between flightlines meets accuracy requirements?

Lidar flightline data has been acquired, converted to LAS format, and calibrated to remove system errors and systematic bias between flightlines. Relative accuracy of data must be tested by systematically investigating LAS elevations in the areas of overlap between adjacent flightlines. According to the USGS version 13 specification, "Accuracy for the lidar point cloud data is to be reported independently from accuracies of derivative products. Point cloud data accuracy is to be tested against a TIN constructed from bare earth lidar points."[1] Therefore, verification of relative in-swath and between-swath accuracy should comprise the following tasks:

- **Bare Earth:** Preliminary classification of flightline swath data (or temporarily tiled in a reversible manner to enable compute effective preliminary classification) for bare earth lidar pulse returns.
- **Between Swath Relative Accuracy:** Quantify areas of overlap between adjacent swaths, segment overlap into desired subunits, and identify common points per subunit for extracting Z-values from TINs constructed from bare earth returns in the immediate vicinity of the points (Figure 3).
- **Within Swath Relative Accuracy:** Identify areas within swath of single (first and only) returns and uniform surface characteristics for quantifying relative accuracy within that swath.

Relative accuracy testing is conducted using raster and vector methods, but methods must emphasize using point cloud data, bare earth classed points, and Z-values extracted from TINs. The use of boundaries, overlap extraction, using point cloud data and TINs to extract set Z-values from each swath corresponding to a common location is superior to raster-to-raster comparisons which compute Z-difference. These Z-difference rasters compute a difference surface from two preliminary surface models that are of limited confidence. Lacking an agreed upon best practice in this area, the use of the Z-difference surface for estimating RMSE Z between the two adjacent swaths is currently an accepted aspect of the "state-of-the-practice". This practice may remain a "supplemental" relative accuracy test, but should be replaced by fully automated vector methods that might be similar to the overlap, segmentation, and centroid generation approach shown in Figure 3.

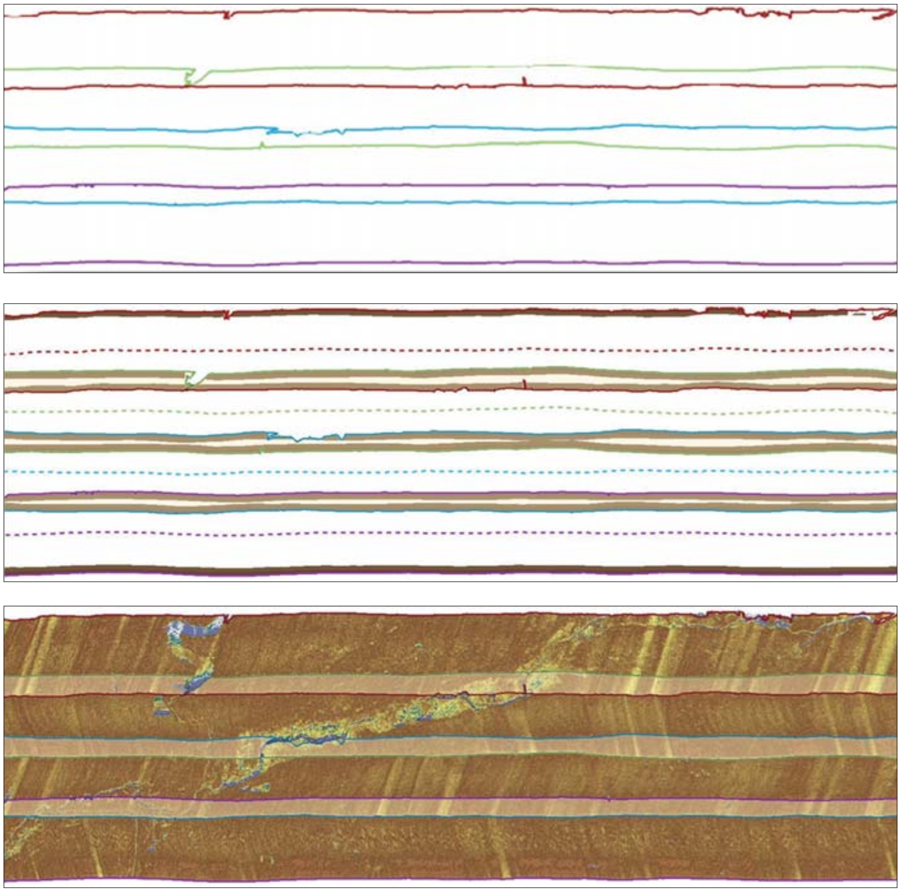


Figure 2. Boundary polygons created from lidar swaths (strips) shown in upper image quantify completeness of coverage, intersected overlap areas shown in the middle image with estimated centerline and outer 10% of the flightline swath shown for quantifying adequate overlap, and pulse count raster shown in the lower image created from first returns shows pulse distribution and voids or gaps in data.

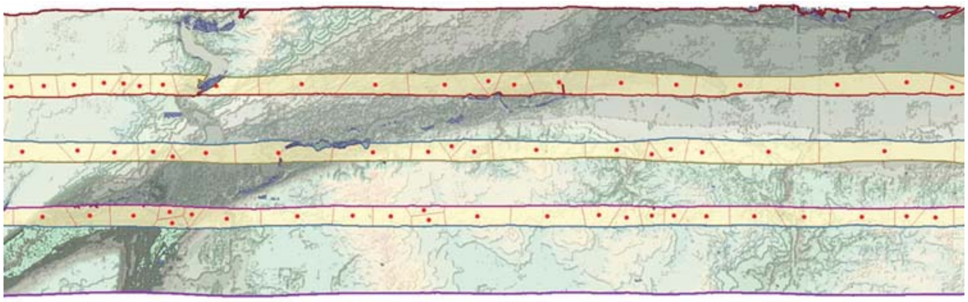


Figure 3. Lidar overlap areas are evaluated for quantifying relative error between swaths. Current methods largely use elevation surfaces created per swath and compute a difference surface. The graphic illustrates the segmentation of each overlap polygon and the generation of points (approximate segment centroids) which may be used to extract Z-values from TINs created from bare earth points from the individual swaths. The Z-values may be used for offset analysis and calculation of RMSE Z.

Benefits: Verifying that lidar flightline swath data are calibrated and that residual relative error within and between adjacent swaths falls within the tolerances of guidelines and specifications helps facilitate further classification and processing of tiled LAS data.

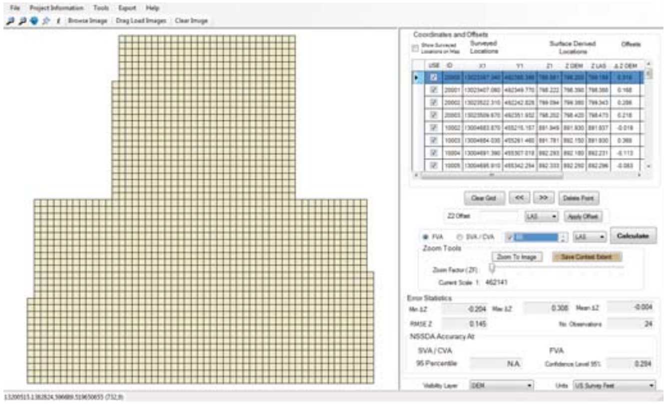
Phase 3 – Tiled Lidar Data Accuracy and Classification

Gateway Question: Have LAS data been properly tiled and classified and do point cloud data for LAS in open spaces and other land cover types meet absolute accuracy specifications?

The following tasks are presented as a possible subset for which best practice methods may be developed and addressed during this phase to comply with guidelines and specifications:

- **Tiling Scheme:** A single, non-overlapping tiling scheme will be developed and used for all tiled deliverables and must be an even integer multiple of the cell size of any raster deliverable (Figure 4).
- **Tiled LAS Deliverables Compliance:** LAS data should be tested against the requirement that all tiled deliverables comply with the tiling scheme without added overlap and seamlessly edge-match without gaps in the horizontal or vertical.
- **Absolute Accuracy Testing:** Tiled LAS files are tested using surveyed checkpoints in required landcover types to determine fundamental vertical accuracy (FVA) of LAS data in open space landcover areas. Supplemental vertical accuracies (SVAs) are compiled for other landcover types, and consolidated vertical accuracy (CVA) is quantified for all of the checkpoints (Figure 4). Best practice methods may also be defined for horizontal accuracy testing which may be derived from comparison between lidar intensity images and ortho photos. Per ASPRS guidelines, FVA is computed as 1.96 * RMSE Z = FVA at the 95% confidence level. SVA and CVA quantities are computed as the 95th percentile of the error quantities.
- **Classification Accuracy and Consistency:** Large open space areas are identified and point cloud data are tested for classification accuracy and consistency within large areas and among points that should be consistently classified.

Benefits: Verifying the absolute accuracy of tiled LAS data, as well as the accuracy and consistency of the point cloud classification, enables further steps to create and verify derived products such as DEMs and hydro-flattened surfaces. Failure to assure the accuracy of LAS data and classification at this phase may cause downstream problems. This is particularly true if LAS absolute accuracy is not rigorously verified and if DEM data produced at later stages are tested and found to not meet specifications.



Current Scale 1: 462141

Min ΔZ -0.204 Max ΔZ 0.308 Mean ΔZ -0.004

RMSE Z 0.145 No. Observations 24

SVA / CVA 95 Percentile N/A

FVA Confidence Level 95% 0.284

Visibility Layer DEM

Units US Survey Feet

Figure 4. Tile index for a multi-county project which includes 2,151 tiles to cover the project area. Each tile is the same size (5,000 ft × 5,000 ft), contains attributes for data products, and provides an exact boundary for tiled and classified LAS data as well as bare earth DEMs, intensity images and other derived information products. LAS data have been evaluated using SIS Topo Analyst and verified to have a fundamental vertical accuracy of .284' or approximately 8.65 cm. (Data provided courtesy of USGS Center for Lidar Information, Coordination, and Knowledge)

Phase 4 – Final Products and Delivery

Gateway Question: Do LAS point cloud, DEM data, and all derived information products meet requirements for accuracy, hydro-flattening, documentation and are all products accompanied by required metadata and project reports?

The following tasks are presented as a possible subset for which best practice methods may be developed and addressed during this phase to comply with guidelines and specifications:

- **DEM Testing:** DEM products are tested for tile agreement, lack of edge effects for tile boundaries, complete coverage, as well as absolute accuracy testing.
- **Hydro-Flattening:** DEM data are also tested for Hydro-Flattening for waterbodies.
- **Breakline Verification:** Breaklines for waterbodies are verified and elevations of lakes, 2-D streams, and 1-D streams are tested.
- **Enhanced Visual Checks:** Visualization methods are required for review of lidar point cloud data, colored by elevation or classification as well as to consider them in the context of breaklines data (Figure 5). Classification of point cloud data in the vicinity of breaklines may be evaluated as well as the relative vertical position of breaklines with respect to point cloud elevations to ensure that lake breaklines do not "float above" the nearby land and to generally verify the vertical relationships between breaklines, point cloud data, and derived surfaces.

continued on page 1078

continued from page 1077

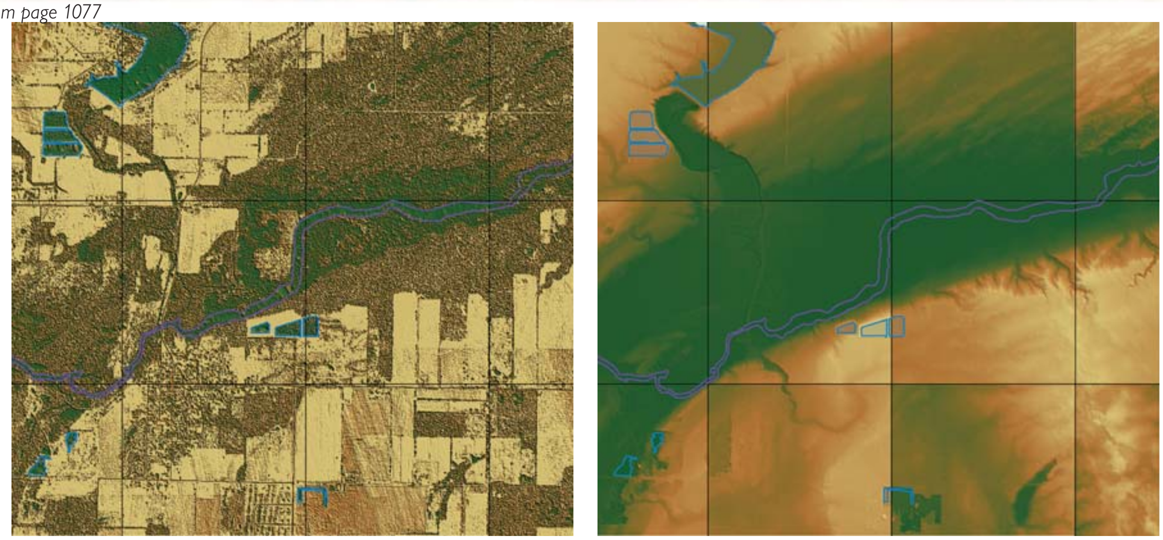


Figure 5. Breaklines for streams and waterbodies overlying intensity data on the left and hydro-flattened DEM data on the right. Note the TIN artifacts and data irregularities in the left-side image for areas within the breaklines and the smooth surface within the same areas in the right-side image. These forms of tests may be a typical part of the advanced visual checks performed on final data sets.

- **Products, Documents, Reports, and Metadata:** There is a complex set of deliverable components in a lidar data project. Checksheets and methods are needed to compile a complete list of deliverable products and to verify them against requirements and specifications. Best practice methods are needed that might include a project data directory structure for data and derived information products, as well as a possible project documents, reports, and metadata repository to assure proper organization and delivery of final products.

Benefits: Standardizing the organization and delivery of components of project products will help assure consistent products that may be effectively tested against guidelines and specifications. Implementing streamlined capabilities that enable real-time testing of incremental products during the lidar project life cycle will ensure that post production tests of end products will meet product specifications and requirements and will help improve the accuracy and quality of lidar data and derived information products.

Focus Projects

Spatial Information Solutions (SIS) has compiled a collection of recently acquired sample lidar data sets from federal agencies as well as industry lidar data producers. In some cases, data were provided along with surveyed checkpoints which were used in the SIS software product, Topo Analyst, to verify accuracy and perform other quantitative and qualitative checks on the data sets. A brief description of the data sets, a view of the data and fundamental accuracy characteristics follow for a set of highlighed focus projects.

Channel Islands, California (USGS)

The USGS funded a lidar collection for the Channel Islands off the coast of California. The project presented challenges in terms of difficulty of access, lack of ability to establish ground survey base stations, and the rugged terrain and steep cliffs on the sides of the islands. These factors combined to add difficulties to aspects of data acquisition and processing. The lidar data were collected by helicopter and survey checkpoint locations were accessed by helicopter. As shown in figure 6, the fundamental vertical accuracy on this project was determined to be 5.8 cm.

Multi-County Project, Michigan (USGS)

A USGS-funded acquisition project collected lidar data for multiple counties in Michigan (Figure 7). This data collection illustrates what is becoming more of the norm for lidar acquisition projects in which data are collected for multiple counties and deliverables conform to specifications of the USGS and FEMA for elevation updates, map data enhancements, and flood plain mapping updates. Lidar was collected by a fixed-wing aerial platform. Data management and organization of data and derived products becomes more of a challenge with large projects, in this case comprising over 2,500 tiles of classified LAS data, DEM data, and other derived data products. Accuracy of the lidar point cloud data was determined to be .284 US Survey Feet (8.66 cm) at the 95% confidence level (FVA) and for the DEM to be .342 US Survey Feet (10.42 cm) at the 95% confidence level (FVA).

Texas DOT Demonstration Project (Industry Collaboration)

In a project aimed at highway transportation data acquisition for design and other applications, high-accuracy helicopter and terrestrial mobile lidar datasets were acquired for a segment of I-30 in Texas. Data were acquired to improve understanding of water accumulation of the interstate segment studied. Tuck Mapping flew helicopter lidar and other project partners collected terrestrial mobile lidar datasets. The lidar collected by Tuck Mapping (Figure 8) was found to have an RMSE Z of 0.039 US Survey Feet (1.19 cm) and a FVA 95% confidence level accuracy of 0.076 US Survey Feet (2.32 cm).

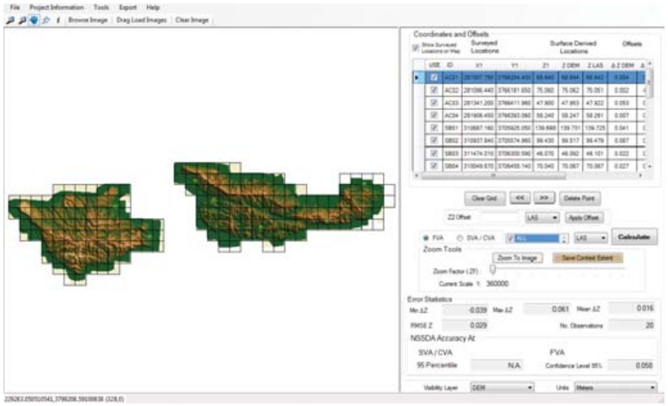


Figure 6. View of Channel Islands elevation data and absolute accuracy verification testing of LAS data by Fundamental Vertical Accuracy for quantifying accuracy to the 95 percent confidence level.

Mississippi River Alluvial Floodplain, Mississippi – The Delta (USACE)

The U.S. Army Corps of Engineers has been leading the way in developing lidar data for areas of need for flood protection, levee management, and flood control structures, as well as coastal areas within their jurisdiction. The Corps also leads the way in developing lidar technologies for advanced data acquisition. The Vicksburg District Corps acquired lidar data and updated ortho photos for the area commonly called "The Delta". This area is the Mississippi River Alluvial Floodplain in the state of Mississippi and is an area of both historical and recent severe flooding. The area covers many counties in Mississippi and the lidar dataset required over 11,000 tiles (at 5000 × 5000 feet per tile) to cover the project area. The broad geographic extent of the Delta (Figure 9), the mostly flat-lying terrain, and the complex nature of the stream networks, waterbodies, oxbow lakes, levees, and flood control structures created challenges in data collection and processing. Projects such as this lidar project for "The Delta" will provide data and derived products of unprecedented accuracy and quality that will strengthen capabilities to plan infrastructure and to help protect our nation's resources and population from the potentially disastrous impacts of flooding and other natural disasters.

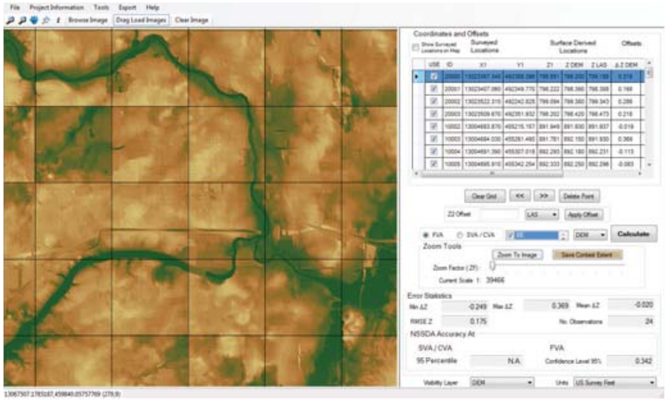


Figure 7. View of Michigan project elevation data and absolute accuracy verification testing of LAS as well as DEM data by Fundamental Vertical Accuracy for quantifying accuracy to the 95 percent confidence level.

Figure 8. View of Tuck Mapping DOT demonstration project elevation data and absolute accuracy verification testing of LAS data by Fundamental Vertical Accuracy for quantifying accuracy to the 95 percent confidence level.

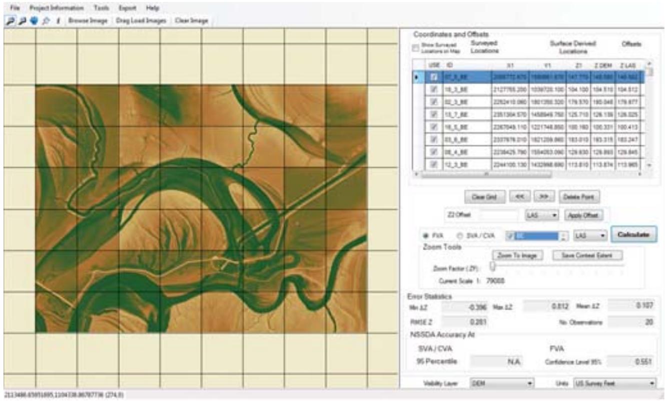


Figure 9. View of elevation data and absolute accuracy verification testing of LAS data by Fundamental Vertical Accuracy for quantifying accuracy to the 95 percent confidence level.

Conclusions

New lidar guidelines and specifications imply that incremental products must be verified to ensure the accuracy, quality, compliance, and completeness of final products. The implication of incremental product verification makes apparent the need for a set of best practices, not just for verifying final products, but for verifying products developed across the lidar project life cycle. Incremental product verification and best practices introduce an opportunity to transform industry practices for delivery and acceptance of products. Key to the transformation of industry practices would be the definition of specific interim deliverables defined as project milestones as well as best practices and mechanisms for engaging the producer and customer in "real time" verification of incremental products as data flow through the project.

Acceptance of incremental products opens the door to partial payment of project fees for partial completion of project deliverables. This transformation would improve project and product visibility and management capabilities for customers while improving data producers' capabilities to improve cash flow during lengthy and cost-intensive lidar projects at the expense of verifying and delivering incremental products. The sum of these factors should have the cumulative effects of improving projects and products, providing a consistent set of best practices, and advancing industry and mapping community practices; ultimately leading to improvement in lidar data and derived mapping product quality, accuracy, and cost-benefits.

Author

Charles O'Hara, *Spatial Information Solution, Inc.*

Spatial Information Solution, Inc. (SIS) was formed as a spin-out from Mississippi State University with a focus on the commercialization of software technologies for map data accuracy and quality verification. Accuracy Analyst and Topo Analyst are an automated QA software solution for orthophoto verification and a software solution for automated accuracy verification and quality assurance of elevation data respectively.

cgohara@spatialis.com
http://www.spatialis.com/