Mapping Matters

By Qassim A. Abdullah, Ph.D., PLS, CP**

Your Questions Answered

The layman's perspective on technical theory and practical applications of mapping and GIS

Question: What is the optimal location for a lidar checkpoint, and why don't we check the accuracy only in open terrain areas, which would increase the confidence in the fundamental vertical accuracy and save time and money? In addition, should every lidar project contain low-confidence area polygons?

Bob Ryan, CP, PLS
Geospatial Technology Team Leader, URS Corporation, Germantown, Maryland, USA

Dr. Abdullah: Before I start my answer, I would like to clarify the question by differentiating between the requirements for the vertical accuracy versus horizontal accuracy, as they differ from each other. For the vertical accuracy, the location of an individual checkpoint should have the following characteristics:

- Located in open space and away from any non-terrain features such as trees, shrubs, ditches, retaining walls, etc.
- The size of the open space is recommended to have a diameter of three to five times the nominal post spacing of the point cloud. This way the automated accuracy method will reliably interpolate the elevation for the terrain elevation model.
- The surface for such open space should be:
  - Smooth and free from high-frequency elevation changes
  - Flat or on evenly slopped terrain
  - Appropriately reflective for the laser energy of the lidar system.
- There may be other conditions adopted by different users.

"Averaging the accuracy of open terrain and different vegetated areas does not make much sense, as it does not represent a useful accuracy figure."

As for the horizontal accuracy checkpoints, they should have the following characteristics:

- Highly visible and reliably measurable on the intensity image
- Preferably, part of a group of points representing a linear feature, such as wall, pavement edge, railroad track, etc.

It has been very difficult to verify the horizontal accuracy of a lidar dataset. For that reason, most horizontal accuracy statements are labeled according to the manufacturers' recommendation for their system; however, an existing accurate planimetric map can play a great role in verifying the horizontal accuracy of a lidar dataset. Other methods could be achieved by flying bi-directional lines over a sloped feature, such as highways and ramps. I would have to admit that few users perform horizontal accuracy checks, due to the difficulties in performing such checks.

"There are many variables that affect the estimated accuracy in a block of data. One of these variables is the composition of the land at the project location—the percentage of the ground covered with vegetation versus open land or built-up areas."

Now I would like to discuss the locations of the checkpoints as a set. There is no straight-forward approach to this problem. There are many variables that affect the estimated accuracy in a block of data. One of these variables is the composition of the land at the project location—the percentage of the ground covered with vegetation versus open land or built-up areas. The preference is to see a fair and unbiased distribution of checkpoints according to the importance or the weight of the terrain category. For a lidar dataset collected for hydro modeling in a mostly vegetated area, the majority of the checkpoints should be located in the vegetated area around the project. A similar yet different approach can be followed for a corridor mapping project.

The problem does not end with the location, as there is no clear requirement based on pure statistical or scientific approach for the selection and distribution of checkpoints. Many of the projects call for 20 or more checkpoints to be used to verify the data; however, most of the time these projects do not specify where to

continued on page 601

continued from page 599

locate these points. I think this is the reason behind the question you raise. Statistically, you would want the checkpoints to follow the characteristics of a valid statistical sample randomly distributed around the project. Again, it becomes difficult to always meet the requirements for the reasons mentioned in the previous discussion. Each user has his or her own priority for the accuracy of the part of the terrain he or she is most concerned about. A road engineer will care about accuracy verification in areas that do not concern a forester, for example. The other variable that may influence the distribution of checkpoints is the geometry of the project and the way the lidar data are calibrated and adjusted to the ground reference points. If the lidar data went through some least squares adjustment to fix elevation biases and discrepancies, then you would like to see each individually adjusted block tested separately. That is true for a large project where the data is divided into blocks; however, a small dataset may require only a single set of checkpoints to validate its accuracy and therefore only one statistical report.

"Many of the projects call for 20 or more checkpoints to be used to verify the data; however, most of the time, these projects do not specify where to locate these points."

As for the second part of the question, checking the accuracy only in open terrain, although it makes a lot of sense, it leaves users who are interested in modeling the ground in vegetated areas in the dark. Many hydrological modeling projects involve an entire watershed that, in most cases, is made up of a vast majority of trees and other vegetation such as shrubs and grass. Therefore, such a user needs some sense of accuracy for the data he or she receives for vegetated areas. I personally do not agree with the current practices of dividing the lidar accuracy data into three categories: fundamental, supplemental and consolidated. I do not see how the consolidated accuracy helps any decision making on a lidar dataset. Averaging the accuracy of open terrain and different vegetated areas does not make much sense, as it does not represent a useful accuracy figure. The accuracy of an open terrain should never be mixed or confused with the accuracy of vegetated areas. For that reason, preliminary ideas for the new ASPRS map accuracy standard suggest dividing the terrain into two categories, vegetated and non-vegetated, and it never suggests a value to mix the two categories in one accuracy number. Here I am referring to the ongoing work of a working group (which I am part of) that was tasked by the ASPRS Primary Data Acquisition Division (PDAD) to draft a new national mapping accuracy standard for geospatial data. Categorizing the types of terrain as vegetated and non-vegetated makes more sense as the technology of lidar systems is advancing at an unprecedented rate, resulting in a dense point cloud that strengthens our ability to map the forest floor more accurately.

The previous discussion flows smoothly into the third part of the question, and that is whether we need to mark the obscured areas as a "low confidence" polygon as is practiced in image-based photogrammetric mapping. Although marking the obscured area as "low confidence" is essential, the situation in the lidar point cloud is a little different. In photogrammetric mapping, it is impossible for the compilers to see the ground under dense foliage, and therefore those areas have to be labeled as "low-confidence areas" or plotted as dashed contours. In such instances, compilers cannot measure the ground elevation, but they do their best in estimating the ground elevation based on the elevation of the ground surrounding the obscured areas that are visible to them. This is not the case with the lidar dataset, as there is always a point cloud that reaches the ground, and therefore an idea can be formed on the accuracy of the lidar-based digital elevation/terrain in vegetated areas.

"For a lidar dataset collected for hydro modeling in a mostly vegetated area, the majority of the checkpoints should be located in the vegetated area around the project. A similar yet different approach can be followed for a corridor mapping project."

An argument can be made on how accurate that elevation model is in vegetated areas. A low-confidence area in photogrammetry means no accuracy measure can be designated for the vegetated areas; while with lidar, low-confidence areas can be tagged with expected accuracy depending on the density of the points that reached the ground and the nature of the terrain. Quantifying the accuracy in vegetated areas is certainly a difficult task, but general rules can be developed and agreed upon within the mapping community.

***Dr. Abdullah is Senior Geospatial Scientist at Woolpert, Inc. He is the 2010 recipient of the ASPRS Photogrammetric (Fairchild) Award.*