ASPRS Guidelines
Vertical Accuracy Reporting for Lidar Data

Version 1.0
Drafted May 15, 2004
Released May 24, 2004
Ownership ASPRS Lidar Committee (PAD)
Editor Martin Flood

Scope

This document identifies the vertical accuracy reporting requirements that are recommended by the American Society for Photogrammetry and Remote Sensing (ASPRS) when analyzing elevation data generated using airborne light detection and ranging or laser radar (lidar) technology. ASPRS recommends all mapping professionals adhere to and follow these guidelines when generating mapping products derived from lidar data.

Reference Standards

These ASPRS guidelines are harmonized with the relevant sections of the Guidelines for Digital Elevation Data (Version 1.0) released by the National Digital Elevation Program (NDEP). The sections on vertical accuracy testing and reporting from the NDEP guidelines have been submitted to the Federal Geographic Data Committee (FGDC) for inclusion as approved revisions to the National Standard for Spatial Data Accuracy (NSSDA). The NDEP guidelines are available online at www.ndep.gov. For reference, the corresponding section references from the NDEP guidelines are cross-referenced and tabulated against section numbers in this document in Appendix A. If cases occur where these ASPRS guidelines are found to be in conflict with the NSSDA, the NSSDA is the controlling document and takes precedent.

Ethical Conduct and Public Health & Safety

ASPRS reminds mapping practitioners who are engaged in the use, development, and improvement of the mapping sciences and related disciplines such as lidar, that they should abide by the principles outlined in the ASPRS Code of Ethics, especially as it relates to the appropriate and honest application of photogrammetry, remote sensing, geographic information systems, and related spatial technologies. ASPRS recommends that mapping professionals always review lidar data vertical accuracy reporting requirements in terms of the potential harm that could be done to the public health and safety in the event that the elevation data fail to satisfy the specified vertical accuracy.

Guidelines

1. Accuracy Requirements

1.1. Vertical Accuracy

Vertical accuracy is the principal criterion in specifying the quality of elevation data, and vertical accuracy requirements depend upon the intended user applications.[1] There are five principal applications where high vertical accuracy is normally required of digital elevation datasets:

(1) For marine navigation and safety.
(2) For storm water and floodplain management in flat terrain.
(3) For management of wetlands and other ecologically sensitive flat areas.
(4) For infrastructure management of dense urban areas where planimetric maps are typically required at scales of 1 inch = 100 feet and larger scales.
(5) For special engineering applications where elevation data of the highest accuracy are required.

Whereas there is a tendency to specify the highest accuracy achievable for many other applications, users of elevation data must recognize that lesser standards may suffice, especially when faced with the increased costs for higher accuracy elevation data.

When contracting for lidar-derived elevation data, it is important to specify the vertical accuracy expected for all final products being delivered. For example, when contours or gridded digital elevation models (DEMs) are specified as deliverables from lidar-generated mass points, a TIN may first be produced from which a DEM or contours are derived. If done properly, error introduced during the TIN to contour/DEM process should be minimal; however, some degree of error will be introduced. Accuracy should not be specified and tested for the TIN with the expectation that derivatives will meet the same accuracy. Derivatives may exhibit greater error, especially when generalization or surface smoothing has been applied to the final product. Specifying accuracy of the final product(s) requires the data producer to ensure that error is kept within necessary limits during all production steps.

It should be noted that many states have regulations that require elevation data to be produced by licensed individuals to protect the public from any harm that an incompetent data producer may cause. Traditionally, such licensing is generally linked to experience in proving that products are delivered in accordance with the National Map Accuracy Standards (NMAS), or equivalent. Information about the National Map Accuracy Standard (NMAS) and the National Standard for Spatial Data Accuracy (NSSDA) is available from a variety of sources, including "Digital Elevation Model Technologies and Applications: The DEM Users Manual" and Part 5 of the referenced NDEP guidelines. An understanding of the basic principles of these Standards will be helpful for understanding the following guidelines for determining vertical accuracy requirements for lidar-derived elevation data.

With the NSSDA, the vertical accuracy of a data set ($\text{Accuracy}_{(z)}$) is defined by the root mean square error ($\text{RMSE}_{(z)}$) of the elevation data in terms of feet or meters at ground scale, rather than in terms of the published map's contour interval. Because the NSSDA does not address

---

[1] See Chapter 11, "Digital Elevation Model Technologies and Applications: The DEM Users Manual," ASPRS, 2001 for a more detailed discussion.

the suitability of data for any particular product, map scale, contour interval, or other application, no error thresholds are established by the standard. However, it is often helpful to use familiar NMAS thresholds for determining reasonable NSSDA accuracy requirements for various types of terrain and relief. This relationship can be shown to be:

[1] $\text{NMAS CI} = 3.2898 * \text{RMSE}_{(z)}$
[2] $\text{NMAS CI} = \text{Accuracy}_{(z)}/0.5958$

where

[3] $\text{Accuracy}_{(z)} = 1.9600 * \text{RMSE}_{(z)}$ (Normally Distributed Error)

Note that for error that is not normally distributed, ASPRS recommends $\text{Accuracy}_{(z)}$ be determined by 95th percentile testing, not by the use of Equation [3]. A normal distribution can be tested for by calculating the skewness of the dataset. If the skew exceeds ±0.5 this is a strong indicator of asymmetry in the data and further investigation should be completed to determine the cause. Based on this relationship, the $\text{Accuracy}_{(z)}$ values shown in Table 1 below are NSSDA equivalents to the NMAS error thresholds for common contour intervals and should be taken as the recommended ASPRS vertical accuracy requirements for lidar data to support mapping products that meet the corresponding NMAS standard.

NMAS Equivalent Contour Interval	NSSDA $\text{RMSE}_{(z)}$	NSSDA $\text{Accuracy}_{(z)}$	Required Accuracy for Reference Data for "Tested to Meet"
0.5	0.15 ft or 4.60 cm	0.30 ft or 9.10 cm	0.10 ft
1	0.30 ft or 9.25 cm	0.60 ft or 18.2 cm	0.20 ft
2	0.61 ft or 18.5 cm	1.19 ft or 36.3 cm	0.40 ft
4	1.22 ft or 37.0 cm	2.38 ft or 72.6 cm	0.79 ft
5	1.52 ft or 46.3 cm	2.98 ft or 90.8 cm	0.99 ft
10	3.04 ft or 92.7 cm	5.96 ft or 181.6 cm	1.98 ft

Table 1 Comparison of NMAS/NSSDA Vertical Accuracy

In contracting for lidar data production, the required vertical accuracy should be specified in terms of $\text{Accuracy}_{(z)}$, rather than NMAS CI, the correct value for which may be calculated from Equation [2] above for any given NMAS CI, extracted from the third column of Table 1 or uniquely derived for a particular application. Consistent use of the User Requirements Menu when contracting/specifying lidar-derived elevation data is highly recommended by ASPRS. Details of the User Requirements Menu can be found in the NDEP Guidelines or the DEM Manual referenced earlier.

However, it should be noted that stating a single vertical accuracy requirement without providing additional clarification and details of the intended purpose of the lidar-derived elevation dataset may not be sufficient information to allow for proper planning and implementation of the field data collection by the data provider. Testing of lidar-derived elevation data over various ground cover categories has revealed that the magnitude and distribution of errors often vary between different land cover types. To account for this, ASPRS recommends the following:

1. For ASPRS purposes, the lidar dataset's required "fundamental" vertical accuracy, which is the vertical accuracy in open terrain tested to 95% confidence (normally distributed error), shall be specified, tested and reported. If no distinction is made when a document references "vertical accuracy", it shall be assumed to be "fundamental" (best case) vertical accuracy.
2. If information is required on the vertical accuracy achieved within other ground cover categories outside open terrain, either to meet the same specification as the fundamental vertical accuracy or a more relaxed specification, then "supplemental" vertical accuracies, that is vertical accuracy tested using the 95th percentile method (not necessarily normally distributed) shall be specified, tested and reported for each land cover class of interest.
3. If contour maps or similar derivative products are to be generated across an entire project area, the project-wide vertical accuracy requirement shall be the same as calculated by Equation [1] or listed in Table 1 across all land cover classes. For ASPRS purposes this means that vertical accuracy in such cases shall be specified, tested and reported for each land cover class, reporting a fundamental vertical accuracy in open terrain and a supplemental vertical accuracy in each unique land cover class, each of which must independently meet the requirements for the desired contour interval.
4. Contour maps or similar derivative products that cover several different land cover classes in a project shall only be reported as "Tested" or "Compiled to Meet" (see Section 3.2) a given accuracy in accordance with the worst vertical accuracy, fundamental or supplemental, of any of the land cover classes to be included in the mapping product.
5. In some circumstances, it may be preferable to specify a different vertical accuracy in different land cover classes, specifying a relaxed vertical accuracy in forested areas, for example, than in tall grass. Such situations shall be explicitly stated in the project specifications.
6. It is commonly accepted that vertical accuracy testing in very irregular or steep sloping terrain is inappropriate due to the high probability that the error in the testing process is a significant contributor to the final error statistic and thus biases the results. For example a small but acceptable horizontal shift in the data may reflect in an unacceptable vertical error measurement. Because of this concern, ASPRS recommends that vertical accuracy testing always be done in areas where the terrain is as level and consistent as possible. In mountainous areas, level areas may not be easy to access, but attempts should be made to keep test points in reasonably low slope and smooth terrain as possible.

Note that for the specific case of contour mapping, ASPRS does not support extrapolating a fundamental vertical accuracy across different land cover classes with the assumption the vertical accuracy will meet the stated mapping standard. For example, if a dataset is reported with a fundamental vertical accuracy that just meets the vertical accuracy requirement listed in Table 1 for the desired contour interval, it is probable that it will not meet that mapping standard outside of open terrain. Supplemental vertical accuracy reporting shall always be requested and provided for every land cover class for which it is intend to generate contour maps and care should be taken to verify that the required vertical accuracy for the given contour interval is met in each and every land cover class.

For legacy datasets for which only a "vertical accuracy" was reported with no indication if this is fundamental, supplemental or consolidated accuracy, ASPRS recommends assuming this is a fundamental (best-case) vertical accuracy and recommends caution when working with the dataset in different land cover classes. If possible, review the QA/QC data and re-test the data to measure supplemental vertical accuracies (95th percentile testing) in areas outside open terrain.

1.2. Horizontal Accuracy

Horizontal accuracy is another important characteristic of elevation data; however, it is largely controlled by the vertical accuracy requirement. If a very high vertical accuracy is required then it will be essential for the data producer to maintain a very high horizontal accuracy. This is because horizontal errors in elevation data normally, but not always, contribute significantly to the error detected in vertical accuracy tests.

As a general rule, horizontal error is more difficult than vertical error to assess in lidar datasets. This is because the land surface often lacks distinct (well defined) topographic features necessary for such tests or because the resolution of the elevation data is too coarse for precisely locating distinct surface features. For these reasons, ASPRS does not require horizontal accuracy testing of lidar-derived elevation products. Instead, ASPRS requires data producers to report the expected horizontal accuracy of elevation products as determined from system studies or other methods. See *ASPRS Guidelines: Horizontal Accuracy Reporting for Lidar Data* (to be published) and *ASPRS Guidelines: Sensor Calibration and Reporting* (to be published) as well as section 1.5.3.4 of the NDEP Guidelines for further information on testing and reporting of the horizontal accuracy of lidar data.

However, when considering vertical accuracy, it is important to specify some minimum expectation of horizontal accuracy for elevation data acquired using lidar, so ASPRS recommends Table 2 shall be used as a guideline. Note that a contractual horizontal accuracy specification for lidar data collection still requires the lidar data producer to ensure that an appropriate methodology and horizontal control structure is applied during the collection and processing of the elevation data and acceptable reporting procedures identified and agreed to by the contractor.

NMAS Map Scale	**NMAS** CMAS 90%	**NSSDA** RMSE(r)	**NSSDA** Accuracy(r) 95% confidence level
1" = 100' or 1:1,200	3.33 ft	2.20 ft or 67.0 cm	3.80 ft or 1.159 m
1" = 200' or 1:2,400	6.67 ft	4.39 ft or 1.339 m	7.60 ft or 2.318 m
1" = 400' or 1:4,800	13.33 ft	8.79 ft or 2.678 m	15.21 ft or 4.635 m
1" = 500' or 1:6,000	16.67 ft	10.98 ft or 3.348 m	19.01 ft or 5.794 m
1" = 1000' or 1:12,000	33.33 ft	21.97 ft or 6.695 m	38.02 ft or 11.588 m
1" = 2000' or 1:24,000 *	40.00 ft	26.36 ft or 8.035 m	45.62 ft or 13.906 m

Table 2 Comparison of NMAS/NSSDA Horizontal Accuracy

* The 1:24,000- and 1:25,000-scales of USGS 7.5-minute quadrangles are smaller than 1:20,000; therefore, the NMAS horizontal accuracy test for well-defined test points is based on 1/50 inch, rather than 1/30 inch for maps with scales larger than 1:20,000.

2. Accuracy Assessment and Reporting

2.1. General Guidance

The NSSDA specifies that vertical accuracy should be reported at the 95 percent confidence level for data tested by an independent source of higher accuracy as:

"Tested __ (meters, feet) vertical accuracy at 95 percent confidence level."

For ASPRS purposes, the independent source of higher accuracy should be at least three times more accurate than the dataset being tested, whenever possible. The NSSDA further states that an alternative "Compiled to Meet" statement shall be used when the guidelines for testing by an independent source of higher accuracy cannot be followed and an alternative means is used to evaluate accuracy. Accuracy should be reported at the 95th percent confidence level for data produced according to procedures that have been consistently demonstrated to achieve particular vertical accuracy values as:

"Compiled to meet __ (meters, feet) vertical accuracy at 95 percent confidence level."

For ASPRS purposes, the "Compiled to Meet" statement should be used by data producers when no independent test results are available or can be practically obtained. For example, vertical accuracy may be impossible to test against an independent source of higher accuracy in very remote or rugged terrain.

It is important to note that the present NSSDA test for vertical accuracy is valid only if errors for the dataset follow a normal or Gaussian distribution, i.e., one defined by a bell-shaped curve. NSSDA modifications for testing and reporting accuracy of non-normal error distributions are being recommended to the FGDC by the NDEP. Whereas vertical errors in open terrain typically have a normal distribution, vertical errors do not typically follow a normal distribution in other land cover categories, especially in dense vegetation where even active sensors such as lidar may be unable to detect the ground. For this reason, additional ASPRS guidelines are provided below for reporting the vertical accuracy of lidar-derived elevation data in land cover categories other than open terrain. For example, forested areas, scrub, wheat or corn fields, tall weeds, mangrove, sawgrass, or urban terrain.

2.2. Designing Accuracy Tests

The NSSDA specifies:

> *If data of varying accuracies can be identified separately in a dataset, compute and report separate accuracy values.*

Many factors will vary over time and space for any particular elevation production project. Major variations in certain factors may have significant influence on the resulting accuracy of the data. To derive an accuracy statistic that is meaningful and representative of the data, potential variables, such as those discussed below, should be considered during the design of the accuracy tests.

2.2.1. Continuity of Data Collection and Processing

Data producers have unique systems and procedures for collecting and processing lidar data. Any time multiple producers and collection systems are utilized to gather lidar data over the same project area, the data should be tested separately for each producer or collection system. System components (equipment, procedures, software, etc.) may also vary over the life of a project. When there is reason to suspect that such changes may have a significant effect on accuracy, these variations should be tested separately.

2.2.2. Topographic Variation

Varying types of topography (such as mountainous, rolling, or flat terrain) within a project may affect the accuracy at which the elevation surface can be modeled. Also, for many applications, the accuracy requirement in high-relief terrain may be less than that for flat terrain. In such situations, it may be preferable to specify different accuracy requirements for the various terrain types and to design separate tests for each.

2.2.3. Ground Cover Variation

Studies have shown lidar errors to be significantly affected by various ground cover types. Because vegetation can limit ground detection, tall dense forests and even tall grass tend to cause greater elevation errors than unobstructed (short grass or barren) terrain. Errors measured in areas of different ground cover also tend to be distributed differently from errors in unobstructed terrain. For these reasons, ASPRS requires open terrain to be tested separately from other ground cover types. Testing over any other ground cover category is required only if that category constitutes a significant portion of the project area deemed critical to the customer.

2.3. Selecting and Collecting Checkpoints

ASPRS recommends that all checkpoint survey work to be used in verifying the vertical accuracy meets contractual specifications (as opposed to checkpoint data used by the lidar data provider for their own internal accuracy assessment tests) be undertaken by, or be under the supervision of, an independent survey firm licensed in the particular state where the project area is located. Independent in this case is taken to mean having no contractual or financial connections to the lidar data provider; in particular the survey firm should not have a subcontractor relationship to the data provider.

Checkpoints should be well distributed throughout the dataset. ASPRS recommends the following NSSDA guidance be followed when choosing checkpoint locations:

> *Checkpoints may be distributed more densely in the vicinity of important features and more sparsely in areas that are of little or no interest. When the distribution of error is likely to be nonrandom, it may be desirable to locate checkpoints to correspond to the error distribution. For a dataset covering a rectangular area that is believed to have uniform positional accuracy, checkpoints may be distributed so that points are spaced at intervals of at least 10 percent of the diagonal distance across the dataset and at least 20 percent of the points are located in each quadrant of the dataset.*

2.3.1. Land Cover Categories

The NSSDA states:

> *A minimum of 20 checkpoints shall be tested, distributed to reflect the geographic area of interest and the distribution of error in the dataset. When 20 points are tested, the 95 percent confidence level allows one point to fail the threshold given in product specifications.*

However, ASPRS recommends collecting a minimum of 20 checkpoints (30 is preferred) in each of the major land cover categories representative of the area for which lidar data vertical accuracy is to be verified. This provides more robust characterization of the error distribution across the dataset and helps to identify potential systematic errors. Thus if five major land cover categories are determined to be applicable for a particular project, then a minimum of 100 total checkpoints are required. Note that in cases where more than 20 checkpoints are collected in a particular land cover class, vertical accuracy reporting is to be based on the 20 (or 30 preferred) worst or least accurate checkpoints in that land cover class, after eliminating checkpoints that have been identified as containing errors and blunders in the ground survey. It is not acceptable practice to collect an abundance of checkpoints and retain only the best for vertical accuracy reporting.

The most common land cover categories are as follows:

- ✓ Open terrain (sand, rock, dirt, plowed fields, lawns, golf courses).
- ✓ Tall weeds and crops.
- ✓ Brush lands and low trees.
- ✓ Forested areas fully covered by trees.
- ✓ Urban areas with dense man-made structures.

It is up to the lidar data producer and customer to determine the significant land cover categories to be tested. The selection and definition of land cover categories should be based on the unique mix and variations of land cover for the project site and the potential effect of each on the surface application. Care should be taken to ensure adequate planning and QA/QC control is in place for each land cover class when derivative products, such as contour maps, are to be generated across an area with several different land cover classes. For some applications, distinction between grass, brush, and forest may not be sufficient. For example, where very high vertical accuracy is a must, it may be important to understand how variations in grass height and density affect the final vertical accuracy. In such situations, it may be preferable to break “grasses” into two or more categories based on species or stand characteristics.

Whether land cover categories are user defined or chosen based on existing land cover categories such as the Anderson[2] or National Land Cover Dataset[3] land-use and land-cover classification systems, they need to be reported in the metadata. User defined categories should be simple, descriptive and representative of existing major land cover categories. For example, there is no Anderson Level II for lawns, but there is (11) residential, (16) mixed urban or built-up land, and (17) other urban or built-up land. There is a category for pasture, it is (21) cropland and pasture, but cropland normally has taller vegetation than "open terrain."

2.3.2. Checkpoints

The QC (quality control) checkpoints should be selected on flat terrain, or on uniformly sloping terrain for x-meters in all directions from each checkpoint, where "x" is the nominal spacing of the DEM or mass points evaluated. Whereas flat terrain is preferable, this is not always possible. Whenever possible, terrain slope should not be steeper than a 20 percent grade because horizontal errors will unduly influence the vertical RMSE calculations. For example, an allowable 1-meter horizontal error in a DEM could cause

---

[2] For a detailed description of this system, see USGS Professional Paper 964, *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. This system is commonly referred to as the "Anderson Classification".

[3] The NLCD 1992 Classification System is available online at http://www.epa.gov/mrlc/definitions.html.

an apparent unallowable vertical error of 20 cm in the DEM. Furthermore, checkpoints should never be selected near severe breaks in slope, such as bridge abutments or edges of roads, where subsequent interpolation might be performed with inappropriate TIN or DEM points on the wrong sides of the break lines.

Checkpoint surveys should be performed relative to National Spatial Reference System (NSRS) monuments of high vertical accuracy, preferably using the very same NSRS monuments used as GPS base stations for airborne GPS control of the mapping aircraft. This negates the potential that elevation differences might be attributed to inconsistent survey control.

To extend control from the selected NSRS monuments into the project area, it is recommended that *NOAA Technical Memorandum NOS NGS-58, "Guidelines for Establishing GPS-Derived Ellipsoid Heights (Standards: 2 cm and 5 cm)," November, 1997 (NOAA, 1997)* be used, using the National Geodetic Survey's latest geoid model to convert from ellipsoid heights to orthometric heights. GPS real-time-kinematic (RTK) procedures are acceptable as long as temporary benchmarks within the project area are surveyed twice with distinctly different satellite geometry to overcome the possibility of GPS multipath error. Subsequent to GPS surveys to extend control into the project area, conventional third-order surveys can be used to extend control to checkpoints that are typically located within forested areas or "urban canyons" where GPS signals would be blocked. QC surveys should be such that the checkpoint accuracy is at least three times more accurate than the dataset being evaluated. For example, if a DEM is supposed to have a vertical $\mathrm{RMSE}_{(z)}$ of 18.5-cm, equivalent to the accuracy required of 2' contours, then the checkpoints should be surveyed with procedures that would yield vertical $\mathrm{RMSE}_{(z)}$ of 6.0 cm or better.

In all methods of accuracy testing and reporting, there is a presumption that the checkpoint surveys are error free and that discrepancies are attributable to the lidar technology assumed to have lower accuracy. This is especially true when the checkpoint surveys are performed with technology and procedures that should yield accuracies at least three times greater than the expected accuracy of the remote sensing data being tested. However, checkpoint surveys are not always error free, and care must be taken to ensure that all survey errors and blunders are identified. When discrepancies do appear, resurveying questionable checkpoints themselves, or asking for the original checkpoint survey data to be reviewed are ways to challenge the accuracy, or inaccuracy, of the checkpoints. Because of potential challenges to the surveyed checkpoints, it is recommended that each checkpoint be marked with a recoverable item, such as a 60d nail and an adjoining flagged stake, to assist in recovery of the checkpoints for resurveys.

2.4. Deriving Dataset Elevations for Checkpoints

Once checkpoints are collected and checked for blunders, elevations corresponding to each checkpoint must be derived from each lidar dataset to be tested. Exact procedures for obtaining these elevations will vary depending on the elevation data model and on software tools available for the test.

Whereas checkpoints may be considered to be well-defined and recoverable, mass points, TIN/DEM points, and contours are not. Because digital elevation models derived from lidar data do not contain well-defined points, it is difficult to test exactly the same points measured as checkpoints in the lidar-derived DEM or TIN dataset. Therefore, it is usually necessary to interpolate an elevation from the surface model at the horizontal (xy) location of each checkpoint.

2.4.1. TIN Interpolation

When mass points are specified as a deliverable, a TIN derived from the mass points provides a surface from which elevations can be directly interpolated at the horizontal location of each checkpoint. A number of commercial software packages have commands[4] that perform this interpolation automatically for a list of checkpoints.

2.4.2. DEM Interpolation

When a gridded DEM is specified as a deliverable, it must be tested to ensure it meets required accuracies even when a TIN with a “Tested to Meet” accuracy statement is used as the DEM source. This is because generalization or smoothing processes employed during DEM interpolation may degrade the elevation surface. If a gridded DEM is to be tested, surface elevations at the checkpoint locations can be interpolated using a suitable interpolation scheme.[5]

2.4.3. Contour Interpolation

Contours may be directly collected from stereoscopic source by a compiler or may be generated from a lidar-derived TIN or DEM. The contours should be tested when specified as a deliverable whether they were directly compiled or derived from another data model, even if the source model meets required accuracies. This is because the accuracy of any derived product can be degraded by interpolation, generalization, or smoothing. Contour tests can be performed two ways. One method consists of plotting checkpoint locations in relationship with surrounding contours and mentally interpolating an elevation for that checkpoint from surrounding contours. Another method requires the contours to be converted to a TIN, from which elevations can be automatically interpreted with software. The TIN method is somewhat risky because TINing software cannot apply the rationale that may be required of the human during interpolation. Therefore, the TINing process may introduce additional error into the interpolated elevations. However, if the TIN test meets accuracy, one can be fairly confident that the contours meet accuracy. If the TIN accuracy fails, it may be necessary to perform the mental interpolation and retest.

2.4.4. Direct Measurement

It should be noted that a direct test of the accuracy of the elevation data may be performed by conducting field measurements after the locations of the unique lidar returns are known. By surveying checkpoints on the known xy coordinates of ground surface lidar point data, a direct comparison can be used as a means of verifying vertical accuracy, removing the introduction of errors caused by interpolating the elevation value from a TIN, a DEM or contour interpolation. Given that most elevation data will be used to generate additional mapping products through TINing, DEM generation, or contour maps, ASPRS does not recommend this procedure as a final QC procedure for reporting the vertical accuracy of datasets. When direct measurements are used to calculate the vertical accuracy of the dataset, this fact must be clearly reported along with a caution about the potential of introducing further interpolation error in any derived mapping products.

---

[4] For example, Arc/Info TINSPOT.

[5] For example a 4-neighbor bilinear interpolation such as that used in the ArcInfo Latticespot command.

2.5. Computing Errors

The "difference" or error for each checkpoint is computed by subtracting the surveyed elevation of the checkpoint from the lidar dataset elevation interpolated at the x/y coordinate of the checkpoint. Thus, if the difference or error is a positive number, the evaluated dataset elevation is higher than true ground in the vicinity of the checkpoint, and if the difference is a negative number, the evaluated dataset elevation is lower.

For Checkpoint$_{(i)}$, the Vertical Error$_{(i)}$ = $(Z_{data(i)} - Z_{check(i)})$

Where:

$Z_{data(i)}$ is the vertical coordinate of the i^{th} checkpoint in the dataset

$Z_{check(i)}$ is the vertical coordinate of the i^{th} checkpoint in the independent source of higher accuracy

i is an integer from 1 to n; n = the number of points being checked

2.6. Analyzing Errors

2.6.1. Blunders, Systematic Error, and Random Error

The "errors" measured in accuracy calculations, in theory, pertain only to random errors, produced by irregular causes whose effects upon individual observations are governed by no known law that connects them with circumstances and so cannot be corrected by use of standardized adjustments. Random errors typically follow a normal distribution. Systematic errors follow some fixed pattern and are introduced by data collection procedures and systems. Systematic errors may occur as vertical elevation shifts across a portion or all of a dataset. These can be identified through spatial analysis of error magnitude and direction or by analyzing the mean error for the dataset. Systematic errors may also be identified as large deviations from the true elevations caused by misinterpretations of terrain surfaces due to trees, buildings, and shadows, fictitious ridges, tops, benches, and striations. A systematic error is predictable in theory and is, therefore, not random. Where possible, systematic errors should be identified and eliminated from a set of observations prior to accuracy calculations.

A blunder is an error of major proportion, normally identified and removed during editing or QC processing. A potential blunder may be identified as any error greater than three times the standard deviation (3 sigma) of the error. Errors greater than 3 sigma should be analyzed to determine the source of the blunder and to ensure that the blunder is not indicative of some unacceptable source of systematic error. Checkpoints with large error should not simply be thrown out of the test sample without investigation; they may actually be representative of some error characteristic remaining in the elevation surface and should be addressed in the metadata.

It is generally accepted that errors in open terrain represent random errors in the lidar sensor system, whereas errors in vegetated areas may include systematic errors. For example, systematic inability to penetrate dense vegetation, and/or systematic deficiencies in procedures used to generate bare-earth elevation datasets. A single large error (outlier) in a forested area, for example, can totally skew RMSE calculations of a large population of checkpoints that otherwise satisfy the accuracy criteria.

3. Calculating and Reporting Vertical Accuracy – ASPRS Requirements

3.1. Fundamental Vertical Accuracy

The **fundamental vertical accuracy** of a dataset must be determined with checkpoints located only in open terrain, where there is a very high probability that the sensor will have detected the ground surface. The fundamental accuracy is the value by which vertical accuracy can be equitably assessed and compared among different datasets. *Fundamental accuracy is calculated at the 95-percent confidence level as a function of* $RMSE_{(z)}$.

3.2. Supplemental and Consolidated Vertical Accuracies

In addition to the fundamental accuracy, **supplemental or consolidated accuracy** values may be calculated for other ground cover categories or for combinations of ground cover categories respectively. Because elevation errors often vary with the height and density of ground cover, a normal distribution of error cannot be assumed and, therefore, $RMSE_{(z)}$ cannot be used to calculate the 95-percent accuracy value. *Consequently a nonparametric testing method (95th Percentile) is required for supplemental and consolidated accuracy tests.*

95th Percentile Error

For supplemental and consolidated accuracy tests, the 95th percentile method shall be employed to determine accuracy. The 95th percentile method may be used regardless of whether or not the errors follow a normal distribution and whether or not errors qualify as outliers. Computed by a simple spreadsheet command, a "percentile" is the interpolated absolute value in a dataset of errors dividing the distribution of the individual errors in the dataset into one hundred groups of equal frequency. The 95th percentile indicates that 95 percent of the errors in the dataset will have absolute values of equal or lesser value and 5 percent of the errors will be of larger value. With this method, $Accuracy_{(z)}$ is directly equated to the 95th percentile, where 95 percent of the errors have absolute values that are equal to or smaller than the specified amount.

Prior to calculating the data accuracy, these steps should be taken:

- ✓ Separate checkpoint datasets according to important variations in expected error such as by land cover class (see for example NDEP Guidelines section 1.5.2.2).
- ✓ Edit collected checkpoints to identify, remove or minimize errors and blunders (see for example NDEP Guidelines section 1.5.2.3).
- ✓ Interpolate the elevation surface for each checkpoint location (see for example NDEP Guidelines section 1.5.2.4)
- ✓ Identify and eliminate lidar sensor systematic errors and/or blunders in the lidar data processing (see for example NDEP Guidelines sections 1.5.2.5 and 1.5.2.6).

Once these steps are completed, the fundamental vertical accuracy must be calculated. If additional land cover categories are to be tested, supplemental and/or consolidated accuracies may also be computed.

Fundamental Vertical Accuracy Test

Using checkpoints in open terrain only:

1. Compute $RMSE_{(z)} = Sqrt[(\Sigma(\mathbf{Z}_{\mathbf{data(i)}} - \mathbf{Z}_{\mathbf{check(i)}})^2)/n]$

2. Compute Accuracy$_{(z)}$ = 1.9600 * RMSE$_{(z)}$ = Vertical Accuracy at 95 percent confidence level.
3. Report Accuracy$_{(z)}$ as:

   "Tested ______(meters, feet) fundamental vertical accuracy at 95 percent confidence level in open terrain using RMSE$_{(z)}$ x 1.9600."

Supplemental Vertical Accuracy Tests

The following accuracy tests are considered optional, except in cases where the mapping products being generated require specific vertical accuracies be met by each land cover class (e.g. contour mapping). When used, these tests must be accompanied by fundamental vertical accuracy tests. The only possible exception to this rule is the rare situation where accessible pockets of open terrain (road clearings, stream beds, meadows, or isolated areas of exposed earth) do not exist in sufficient quantity for collecting the minimum test points. Only in this instance may supplemental or consolidated accuracies be reported without an accompanying fundamental accuracy. However, this situation must be explained in the metadata. Most likely, when producing an elevation surface where little or no accessible open-terrain exists, the data producer will employ a collection system that has been previously tested to meet certain accuracies and a "Compiled to Meet" statement would be used in lieu of a "Tested to Meet" statement.

When testing ground cover categories or combinations of categories excluding open terrain:

1. Compute 95th percentile error (described above) for each category (or combination of categories).
2. Report:

   "Tested ______(meters, feet) supplemental vertical accuracy at 95th percentile in (specify land cover category or categories)"
3. In the metadata, document the errors larger than the 95th percentile. For a small number of errors above the 95th percentile, report x/y coordinates and z-error for each QC checkpoint error larger than the 95th percentile. For a large number of errors above the 95th percentile, report only the quantity and range of values.

Consolidated Vertical Accuracy Tests

When 40 or more checkpoints are consolidated for two or more of the major land cover categories, representing both the open terrain and other land cover categories (for example, forested), a *consolidated vertical accuracy* assessment may be reported as follows:

1. Compute 95th percentile error (described above) for open terrain and other categories combined.
2. Report

   "Tested ______(meters, feet) consolidated vertical accuracy at 95th percentile in: open terrain, (specify all other categories tested)"
3. In the metadata, document the errors larger than the 95th percentile. For a small number of errors above the 95th percentile, report x/y coordinates and z-error for each QC checkpoint error larger than the 95th percentile. For a large number of errors above the 95th percentile, report only the quantity and range of values.

Failed Accuracy Tests

If the fundamental vertical accuracy test fails to meet the prescribed accuracy, there is a serious problem with the control, collection system, or processing system or the achievable accuracy of the production system has been overstated. If a systematic problem can be identified, it should be corrected, if possible, and the data should be retested. If a systematic problem cannot be identified, it is probable that the entire dataset will need to be recollected, depending on the contractual agreement with the data provider.

If a dataset passes the fundamental vertical accuracy test, but fails to meet supplemental or consolidated vertical accuracy tests (e.g. meets prescribed accuracy in open terrain, but not in forested areas), there may be a problem with the control, collection system, or processing system as above. It is also possible that the data was collected with equipment, procedures or methods designed to just meet the fundamental vertical accuracy requirement in open terrain, but that fail to take in to account the degradation of accuracy in other land cover classes. However, a more probable explanation is that serious errors may have occurred in automated or manual filtering of the lidar data. If such errors can be identified, they should be corrected, if possible, and the data should be retested. If such errors cannot be identified, the areas impacted may need to be excluded from the dataset or data recollected for the affected areas, depending on the contractual agreement with the data provider.

3.3. Reporting Vertical Accuracy of Untested Data – ASPRS Requirements

Use the 'Compiled to Meet' statement below when the above guidelines for testing by an independent source of higher accuracy cannot be followed and an alternative means is used to evaluate accuracy. Report accuracy at the 95th percent confidence level for data produced according to procedures that have been demonstrated to produce data with particular vertical accuracy values as:

"Compiled to meet ___ (meters, feet) fundamental vertical accuracy at 95 percent confidence level in open terrain."

The following accuracy statements are optional. When used they must be accompanied by a fundamental vertical accuracy statement. For ground cover categories other than open terrain, report:

"Compiled to meet ___ (meters, feet) supplemental vertical accuracy at 95th percentile in (specify land cover category or categories)."

For all land cover categories combined, report:

"Compiled to meet ___ (meters, feet) consolidated vertical accuracy at 95th percentile in: open terrain, (list all other relevant categories)."

3.4. Testing and Reporting Horizontal Accuracy – ASPRS Requirements

ASPRS does not require independent testing of horizontal accuracy for lidar-derived elevation products. Instead, ASPRS requires data producers to report the expected horizontal accuracy of elevation products as determined from system studies or other methods. See *ASPRS Lidar Guidelines: Horizontal Accuracy Reporting for Lidar Data* (to be published) and *ASPRS Lidar Guidelines: Sensor Calibration and Reporting* (to be published) as well as section 1.5.3.4 of the NDEP Guidelines for further information on testing and reporting of the horizontal accuracy of lidar data.

3.5. Accuracy Assessment Summary

Providers of lidar elevation data use a variety of methods to control the accuracy of their products. Lidar data providers may collect hundreds of static or kinematic control points for internal quality control and to adjust their datasets to these control points. To the degree that such control points are used in a fashion similar to control for aerotriangulation in photogrammetry, by which the datasets are adjusted to better fit such control points, then lidar data providers may use the "Compiled to Meet" accuracy statements listed above for such datasets.

Note, however, that with mature technologies such as photogrammetry, users generally accept "Compiled to Meet" accuracy statements without independent accuracy testing due to the rigorous methodology and recognized best practices that support such mature technology. However, with developing technologies such as lidar, users often require independent accuracy tests even for “Compiled to Meet” accuracy statements; testing for which, as outlined above, is more complex, especially when errors include "outliers" or do not follow a normal distribution as required for the use of RMSE in accuracy assessments. Because of these complexities, ASPRS strongly recommends the "truth in advertising" approach, whereby lidar data providers report vertical accuracies in open terrain separately from other land cover categories, report vertical accuracies in other than open terrain using 95th percentile testing and document the size of the errors larger than the 95th percentile in the metadata, regardless of the intention of the contractor to independently test (or not) such statements.

3.6 Relative Vertical Accuracy

The accuracy measurement discussed in these guidelines refers to absolute vertical accuracy, which accounts for all effects of systematic and random errors. For some applications of lidar elevation data, the point-to-point (or relative) vertical accuracy is more important than the absolute vertical accuracy. Relative vertical accuracy is controlled by the random errors in a dataset. The relative vertical accuracy of a dataset is especially important for derivative products that make use of the local differences among adjacent elevation values, such as slope and aspect calculations. Because relative vertical accuracy may be difficult to measure unless a very dense set of reference points is available, this ASPRS guideline does not prescribe an approach for its measurement. If a specific level of relative vertical accuracy is a stringent requirement for a given project, then the plan for collection of reference points for validation should account for that. Namely, reference points should be collected at the top and bottom of uniform slopes. In this case, one method of measuring the relative vertical accuracy is to compare the difference between the elevations at the top and bottom of the slope as represented in the elevation model vs. the true surface (from the reference points). In many cases, the relative vertical accuracy will be much better than the absolute vertical accuracy, thus the importance of thoroughly measuring and reporting the absolute accuracy, as described in these guidelines, so the data users can have an idea of what relative accuracy to expect.

GLOSSARY:

Accuracy – the closeness of an estimated (for example, measured or computed) value to a standard or accepted [true] value of a particular quantity. Note: Because the true value is not known, but only estimated, the accuracy of the measured quantity is also unknown. Therefore, accuracy of coordinate information can only be estimated.

- **Absolute Vertical Accuracy** – a measure that relates the stated elevation to the true elevation with respect to an established vertical datum. The computed value for the absolute vertical accuracy (tested, or compiled to) should be included in the metadata file.

Artifacts – Buildings, trees, towers, telephone poles or other elevated features that should be removed when depicting a DEM of the bare-earth terrain. Artifacts are not just limited to real features that need to be removed. They also include unintentional byproducts of the production process, such as stripes in manually profiled DEMs. Any feature, whether man-made or system-made, that unintentionally exists in a digital elevation model.

ASPRS - American Society for Photogrammetry and Remote Sensing

Calibration – Procedures used to identify systematic errors in hardware, software, and procedures so that these errors can be corrected in preparing the data derived there from.

Checkpoint – One of the points in the sample used to estimate the positional accuracy of the dataset against an independent source of higher accuracy.

Confidence level – The probability that errors are within a range of given values.

Consolidated vertical accuracy – The result of a test of the accuracy of 40 or more check points (z-values) consolidated for two or more of the major land cover categories, representing both the open terrain and other land cover categories. Computed using a nonparametric testing method (95th Percentile), a consolidated vertical accuracy is always accompanied by a fundamental vertical accuracy. See fundamental and supplemental vertical accuracies.

Contour – A line connecting points of equal elevation.

Contour interval – The difference in z-values between contours.

DEM - Digital Elevation Model - has at least three different meanings:

- "DEM" is a generic term for digital topographic and/or bathymetric data in all its various forms. Unless specifically referenced as Digital Surface Models (DSMs), the generic DEM normally implies elevations of the terrain (bare earth z-values) void of vegetation and manmade features.

- As used by the U.S. Geological Survey (USGS), a DEM is the digital cartographic representation of the elevation of the land at regularly spaced intervals in x and y directions, using z-values referenced to a common vertical datum. There are many types of standard USGS DEMs.

- As used by other users in the U.S. and elsewhere, a DEM has bare earth z-values at regularly spaced intervals in x and y, but normally following alternative specifications, with narrower grid spacing and State Plane coordinates for example.

DTED - Digital Terrain Elevation Data – Standard elevation datasets of the National Geospatial-Intelligence Agency (NGA), similar to standard USGS DEMs described above.

DTM - Digital Terrain Model - has at least two different definitions:

- In some countries, DTMs are synonymous with DEMs, representing the bare earth terrain with uniformly spaced z-values.

- As used herein, DTMs may be identical to DEMs, but they may also incorporate the elevation of significant topographic features on the land and change points and breaklines that are irregularly spaced so as to better characterize the true shape of the bare earth terrain. The net result of DTMs is that the distinctive terrain features are more clearly defined, and contours generated from DTMs more closely approximate the real shape of the terrain. Such DTMs are normally more expensive and time consuming to produce than uniformly spaced DEMs because breaklines are ill suited for automation; but the DTM results are technically superior to standard DEMs for many applications.

DSM - Digital Surface Model - – Similar to DEMs or DTMs, except that they depict the elevations of the top surfaces of buildings, trees, towers, and other features elevated above the bare earth. DSMs are especially relevant for telecommunications management, forest management, air safety, 3-D modeling and simulation.

Elevation – The distance measured upward along a plumb line between a point and the geoid. The elevation of a point is normally the same as its orthometric height, defined as "H" in the equation: $H = h - N$. This is the "official" geodesy definition of elevation, but the term elevation is also used more generally for height above a specific vertical reference, not always the geoid.

FGDC - Federal Geographic Data Committee (FGDC) http://www.fgdc.gov/

Fundamental vertical accuracy – The fundamental vertical accuracy is the value by which vertical accuracy can be equitably assessed and compared among datasets. The fundamental vertical accuracy of a dataset must be determined with check points located only in open terrain where there is a very high probability that the sensor will have detected the ground surface. It is obtained utilizing standard tests for RMSE. See supplemental and consolidated vertical accuracies.

Grid – A geographic data model that represents information as an array of equally sized square cells. Each grid cell is referenced by its geographic or x/y orthogonal coordinates.

Horizontal accuracy – Positional accuracy of a dataset with respect to a horizontal datum.

Independent source of higher accuracy – Data acquired independently of procedures to generate the dataset that is used to test the positional accuracy of a dataset. The independent source of higher accuracy shall be of the highest accuracy feasible and practicable to evaluate the accuracy of the dataset.

Interpolation – The estimation of z-values at a point with xy coordinates, based on the known z-values of surrounding points.

Lidar – Light Detection and Ranging – An instrument that measures distance to an object by emitting timed pulses of light and measuring the time between emission and reception of reflected pulses. The measured time interval is converted to a distance.

Mass points – Irregularly spaced points, each with an xy location and a z-value, used to form a TIN. When generated manually, mass points are ideally chosen to depict the most significant variations in the slope or aspect of TIN triangles. However, when generated by automated methods, For example, by lidar, mass point spacing and pattern depend on characteristics of the technologies used to acquire the data. Mass points are most often used to make a TIN, but not always. They can be used as XYZ point data for interpolation of a grid without an intermediate TIN stage.

NDEP - National Digital Elevation Program

NSRS - National Spatial Reference System

NSSDA - National Standard for Spatial Data Accuracy

Post spacing – The z-values at regularly spaced intervals of a grid (the ground distance in x and y ("post spacing" = .x = .y)). The post spacing is usually specified in units of whole feet or meters. Actual grid spacing, datum, coordinate system, data format, and other characteristics may vary widely from grid to grid.

Profile – A vertical view of a surface derived by sampling surface values along a specified line. In USGS DEMs, profiles are the basic building blocks of an elevation grid and are defined as one-dimensional arrays, i.e., arrays of n columns by 1 row, where n is the length of the profile.

Relative Accuracy – A measure that accounts for random errors in a dataset. Relative accuracy may also be referred to as point-to-point accuracy. The general measure of relative accuracy is an evaluation of the random errors (systematic errors and blunders removed) in determining the positional orientation (For example, distance, azimuth, elevation) of one point or feature with respect to another.

Relative Vertical Accuracy – A measure of the point-to-point vertical accuracy within a specific dataset. To determine relative vertical accuracy, the vertical difference between two points is measured. That difference is then compared to the difference in elevation for the same two points on the reference. The difference between the two measures represents the relative accuracy. The reference must have at least three times the accuracy of the intended product accuracy, insuring that all systematic errors and blunders have been removed. Relative vertical accuracy, an important characteristic of elevation data used for calculating slope, should be documented in the DEM metadata file.

Resolution – In the context of gridded elevation data, resolution is related to the horizontal post spacing and the vertical precision. Other definitions include:

- The size of the smallest feature that can be represented in a surface or image.

- Sometimes used to state the number of points in x and y directions in a lattice, for example, 1201 x 1201 mesh points in a USGS one-degree DEM

Root mean square error – The square root of the mean of squared errors for a sample.

Slope – The measure of change in z-value over distance, expressed either in degrees or as a percent. For example, a rise of 4 meters over a distance of 100 meters describes a 2.3° or 4 percent slope.

Surface – a 3-D geographic feature represented by computer models built from uniformly- or nonuniformly-spaced points with x/y coordinates and z-values.

Supplemental vertical accuracy – The result of a test of the accuracy of z-values over areas with ground cover categories or combinations of categories other than open terrain. Obtained utilizing the 95th percentile method, a supplemental vertical accuracy is always accompanied by a fundamental vertical accuracy. See fundamental and consolidated vertical accuracies.

Triangulated Irregular Networks (TINs) – A set of adjacent, nonoverlapping triangles computed from irregularly spaced points with xy coordinates and z-values. The TIN data structure is based on irregularly spaced point, line, and polygon data interpreted as mass points and breaklines. The TIN model stores the topological relationship between triangles and their adjacent neighbors. The TIN data structure allows for the efficient generation of surface models for the analysis and display of terrain and other types of surfaces. TINs usually require fewer data points than DEMs or DTMs, while capturing critical points that define terrain discontinuities and are topologically encoded so that adjacency and proximity analyses can be performed.

Vertical accuracy – Measure of the positional accuracy of a dataset with respect to a specified vertical datum.

Vertical error – The displacement of a feature's recorded elevation in a dataset from its true or more accurate elevation.

Appendix A: Cross-Reference to NDEP Guidelines for Elevation Data (V1.0)

ASPRS Guidelines Section Number	Corresponding NDEP Guidelines Section Number
1.1. Vertical Accuracy	1.5.1.1 Vertical Accuracy
1.2. Horizontal Accuracy	1.5.1.2 Horizontal Accuracy
2.0 Accuracy Assessment and Reporting	1.5.2 Accuracy Assessment
2.1 General Guidance	1.5.2.1 General Guidance
2.2 Designing Accuracy Tests	1.5.2.2 Designing Accuracy Tests
2.3 Selecting and Collecting Checkpoints	1.5.2.3 Selecting & Collecting Checkpoints
2.4 Deriving Dataset Elevations for Checkpoints	1.5.2.4 Deriving Dataset Elevations for Checkpoints
2.5 Computing Errors	1.5.2.5 Computing Errors
2.6 Analyzing Errors	1.5.2.6 Analyzing Errors
3.0 Calculating & Reporting Vertical Accuracy	1.5.3 Calculating and Reporting Vertical Accuracy
3.1 Fundamental Vertical Accuracy	1.5.3.1 Fundamental Accuracy
3.2 Supplemental and Consolidated Vertical Accuracies	1.5.3.2 Supplemental and Consolidated Vertical Accuracies
3.3 Reporting Vertical Accuracy of Untested Data	1.5.3.3 Reporting Vertical Accuracy of Untested Data
3.4 Testing and Reporting Horizontal Accuracy	1.5.3.4 Testing and Reporting Horizontal Accuracy
3.5 Accuracy Assessment Summary	1.5.3.5 Accuracy Assessment Summary
3.6 Relative Vertical Accuracy	1.5.3.6 Relative Vertical Accuracy