3 Positional Accuracy

Critical to any accuracy assessment project is the determination of precisely the same location on both the reference data and on the imagery or map being assessed. If this correspondence is not achieved, then the resulting poor positional accuracy may cause a false thematic error to occur. For example, it is possible to be in the correct location and mislabel (incorrectly measure or classify) the attribute. It is also possible to correctly label the attribute, but be in the wrong location. In either case, error is introduced into the map or spatial data set. These two factors are not independent of each other, and great care needs to be taken to not only assess each of these factors but also control them to minimize the errors.

As we learned in Chapter 1, accuracy assessment is characterized by two measures: positional and thematic accuracy. This chapter reviews the concepts of positional accuracy and is organized into the following sections. The first section introduces positional accuracy and briefly reviews the causes of positional accuracy error. The second section compares and contrasts the seven common standards for positional accuracy. The next section reviews basic statistics and positional accuracy sample design and collection within the overall framework set forth in the most commonly used standard, the *National Standard for Spatial Data Accuracy* (NSSDA) (FGDC, 1998). The fourth section explains how to analyze the accuracy assessment sample data to estimate positional accuracy under each standard. Finally, the last section compares the standards to one another and outlines a recommendation for positional accuracy that incorporates the concepts of existing standards, yet avoids some of the assumptions required by them.

A major goal of this chapter is to bring clarity to the language and equations of positional accuracy assessment. Since the development of the first standards in 1942, each new standard has introduced new concepts and interpreted old concepts in new ways. As a result, the language of positional accuracy assessment is often confusing, and the equations that comprise the accuracy assessment standards are, unfortunately, sometimes incorrect.

WHAT IS POSITIONAL ACCURACY?

The *Glossary of the Mapping Sciences* (ASPRS and ASCE, 1994) defines positional accuracy as "the degree of compliance with which the coordinates of points determined from a map agree with the coordinates determined by survey or other independent means accepted as accurate." All locations on maps and georeferenced images are expressed by a set of values: x- and y-coordinates for horizontal location. Many data sets also include elevations, which are represented by the letter z.

Positional accuracy uses sampling to estimate the discrepancy between a map or image feature's coordinates or elevations and their "true" location on the earth's surface. Positional accuracy can refer to either horizontal (planimetric) or vertical (elevational) accuracy, and this chapter discusses both.

Several factors can affect the positional accuracy of a map or georeferenced image. For example, the sensor lens may be distorted, or the aircraft carrying the sensor may suddenly tilt or yaw, changing the relationship of the sensor's image plane to the ground. However, the most important cause of positional error arises from the impact of topography on remotely sensed imagery. Because the sensor image plane is flat and the earth has relief such as hills and ravines, the scale of the remotely sensed imagery relative to the earth varies with topographic changes, requiring that some sort of adjustment be made to "terrain-correct" the image. This correction is a complex process that is highly prone to error.

Figure 3.1 presents an example of horizontal positional inaccuracy in which an inaccurate road layer is displayed over the top of an ortho-corrected digital image. The reference data, which have been "accepted as accurate," are survey points indicated by a crosshair on the figure. As you can see, the road layer does not exactly align with the points (i.e., there are positional errors)—the roads are shifted to the north and west of their "true" location, as determined by the survey points. While we can clearly see that

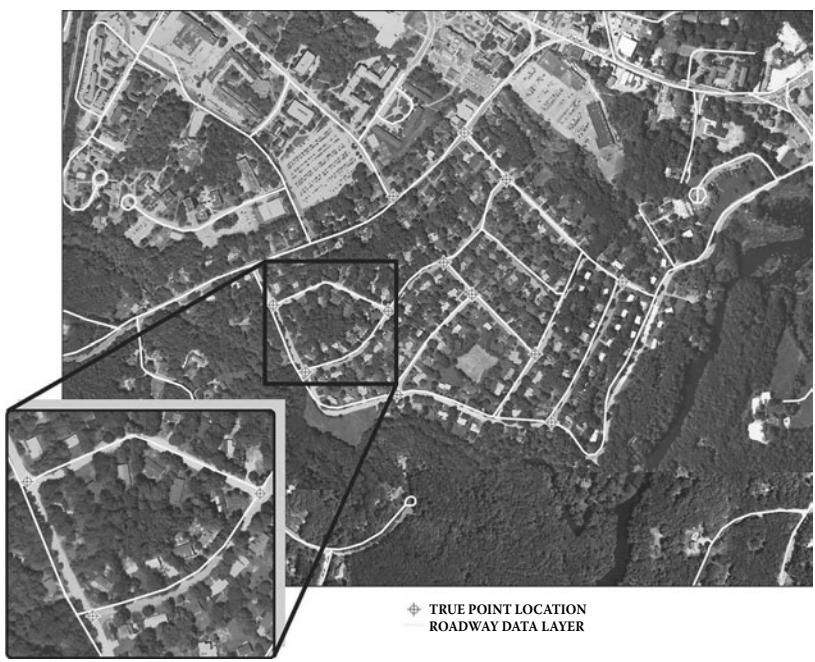

FIGURE 3.1 Illustration of positional errors in a road map (in white) compared to the image, which is assumed to be accurate.

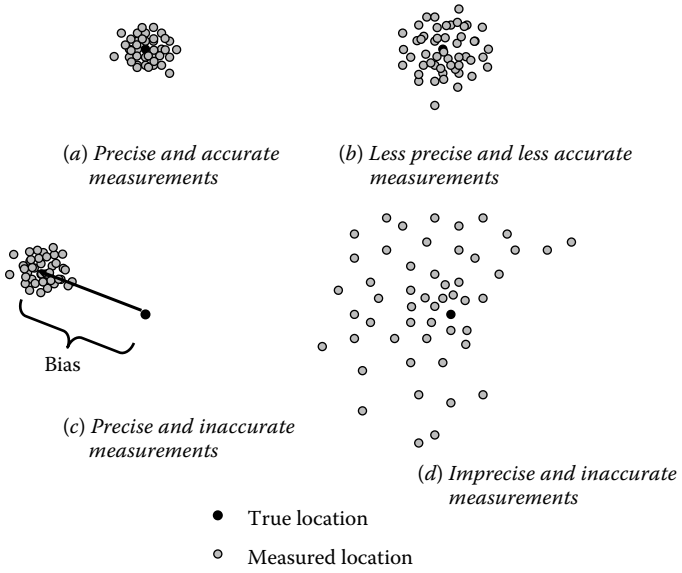

FIGURE 3.2 Illustrations of precision versus accuracy.

the position of the road is inaccurate, we need to use quantitative accuracy assessment to estimate the mean error in the accuracy of the position of the road layer.

In statistics and accuracy assessment, there are two terms that are commonly used and confused that need clarification. Accuracy and precision are often thought of as synonymous, but actually have very different meanings. Accuracy refers to the bias of an estimator. It measures how close an estimated or calculated value is to its true value. Precision refers to the variability in an estimator. It quantifies how repeated measures of the same estimator will vary. Inaccurate measurements can be very precise, and accurate measurements can be imprecise. Figure 3.2 illustrates the concepts of accuracy and precision with an example of multiple measurements made of one location.

In positional accuracy assessment, we are interested in characterizing the accuracy of a geospatial data set. We take samples to determine if a bias (systematic inaccuracy) exists in the data set, and we estimate the magnitude and precision of the bias. We also strive to ensure that our measurements of each sample's reference and geospatial data set's locations are themselves accurate, and we must take enough samples so that our estimate of the bias (if it exists) is precise.

WHAT ARE THE COMMON STANDARDS FOR POSITIONAL ACCURACY?

The *National Standard for Spatial Data Accuracy* (NSSDA) (FGDC, 1998) is the most widely used positional accuracy standard. However, new standards have been developed and several earlier standards are still in use. In addition, it is not

uncommon to have two standards applied to the same project. This section compares and contrasts the seven primary positional accuracy standards:

1. *United States National Map Accuracy Standards* (NMAS) (U.S. Bureau of the Budget, 1947),
2. *Principles of Error Theory and Cartographic Applications* (Greenwalt and Schultz, 1962, 1968), which is cited by all subsequent standards,
3. *ASPRS Interim Accuracy Standards for Large-Scale Maps* (ASPRS, 1989),
4. The Federal Geographic Data Committee's *National Standard for Spatial Data Accuracy* (FGDC, 1998),
5. The Federal Emergency Management Agency's (FEMA) *Guidelines and Specifications for Flood Hazard Mapping Partners* (FEMA, 2003),
6. *ASPRS Guidelines, Vertical Accuracy Reporting for Lidar Data* (ASPRS, 2004), and
7. The National Digital Elevation Program (NDEP) *Guidelines for Digital Elevation Data* (NDEP, 2004).

National Map Accuracy Standards

NMAS (reproduced in its entirety in Chapter 2) stipulates that:

- For horizontal accuracy, not more than 10% of the points tested may be in error by more than 1/30th of an inch (at map scale) for maps larger than 1:20,000 scale, or by more than 1/50th of an inch for maps of 1:20,000 scale or smaller, and
- For vertical accuracy, not more than 10% of the elevation tested may be in error by more than one half the contour interval.

The standard is very straightforward and simple and does not require any assumptions about the distribution of error. Using what is later termed the "percentile method," NMAS merely states that no more than 10% of the samples may exceed the maximum error allowed. However, because it relies on map versus ground units, and because it provides no guidance for creating statistically valid bounds on the estimated error, NMAS is rarely used today.

Principles of Error Theory and Cartographic Applications

The *Principles of Error Theory and Cartographic Applications* (Greenwalt and Schultz, 1962, 1968) report (hereafter referred to as Greenwalt and Schultz) approaches positional accuracy from a diametrically opposite standpoint compared to NMAS by proposing equations that should be applied to estimate the maximum error interval that would occur at various probabilities. The report interprets NMAS' (1947) "10% of the points taken" to limit the size of errors to that within "which 90% of the well defined points will not exceed" (Greenwalt and Schultz, 1962, 1968), which it terms

the "map accuracy standard" (MAS). The report uses probability theory to develop equations for calculating one-dimensional elevation (z) "map accuracy standard" and two-dimensional (x and y) "circular map accuracy standard" (CMAS) statistics by assuming that map errors are normally distributed.† MAS is the estimated interval around the mean vertical error, and CMAS is the estimated interval around the horizontal mean error within which 90% of the errors are predicted to occur.

While seemingly similar, the two standards—NMAS and Greenwalt and Schultz—are very different. NMAS *stipulates* the maximum size of error that 10% of the samples may not exceed. The Greenwalt and Schultz standard does not stipulate a maximum error. Rather, it *calculates* the probable maximum error interval around the mean error from the sample data.

Additionally, Greenwalt and Schultz does not specify 90% as the only probability level to be employed. Instead, it shows how to estimate the distribution of errors under various probability levels and provides tables for converting from one probability level to another.

ASPRS Interim Accuracy Standards for Large-Scale Maps

Similar to NMAS, ASPRS (ASPRS, 1989) standards stipulate a maximum distance beyond which errors may not exceed. However, ASPRS differs from NMAS in stating how to determine if the errors have exceeded the maximum acceptable error. Rather than stipulating that no more than 10% of the errors may exceed the stipulated maximum, ASPRS states that the mean error estimated from the samples may not exceed the stipulated maximum distance. Most importantly, the ASPRS standards migrate the units of measurement of error from map units to ground units. The ASPRS standards also restate the Greenwalt and Schultz CMAS equations, but do not imply that the equations should necessarily be used.

National Standard for Spatial Data Accuracy

As mentioned in Chapter 2, the *National Standard for Spatial Data Accuracy* (FGDC, 1998) established much-needed guidelines for measuring, analyzing, and reporting positional accuracy of both maps and georeferenced imagery such as orthophotos or orthoimages. While developed for federal agencies, the NSSDA standards have been widely accepted by many local and state government agencies, as well as by the private sector. Because of the importance of the NSSDA in establishing positional accuracy assessment procedures, we highly recommend that the reader download the NSSDA at http://www.fgdc.gov/standards/projects/FGDC-standards-projects/accuracy/part3/chapter3.

NSSDA explicitly rejects setting a maximum allowable error at any scale and suggests instead that the maximal allowable error threshold be determined as needed. Instead, accuracy is to be reported "in ground distances at the 95% confidence level," which is interpreted as allowing "one point to fail the threshold given in the product

† We will examine the implications of the assumption of normality later in this chapter.

specification"† when a sample of 20 points is used. It is not unusual for positional accuracy projects to use the equations of NSSDA to calculate accuracy statistics *and* to require that those statistics not exceed the distances established in the ASPRS (1989) standards. Similar to the ASPRS standards, NSSDA relies on ground rather than map units and uses the mean error as an accuracy statistic. NSSDA increases the probability level to 95%, an increase of 5% above the Greenwalt and Schultz-interpreted NMAS level of 90%. NSSDA also incorporates the approach of Greenwalt and Schultz by referencing its equations and defining accuracy as a measure of the maximum error expected at a specific probability level. However, as we will learn later, NSSDA incorrectly implements the Greenwalt and Schultz equations.

Guidelines and Specifications for Flood Hazard Mapping Partners

FEMA's *Guidelines and Specifications for Flood Hazard Mapping Partners* (FEMA, 2003) adds a new dimension to positional accuracy assessment by requiring that a minimum of 20 samples be collected for each major vegetation type of which there may be a minimum of 3, resulting in a minimum of 60 total sites sampled. The vegetation types specified are:

1. Bare-earth and low grass,
2. High grass, weeks, and crops,
3. Brush lands and low trees,
4. Forested, fully covered by trees,
5. Urban areas,
6. Sawgrass, and
7. Mangrove.

ASPRS Guidelines: Vertical Accuracy Reporting for Lidar Data

ASPRS Guidelines for Reporting Vertical Accuracy of Lidar Data (ASPRS, 2004) ratify the FEMA guidance to stratify the landscape into different land cover classes. The ASPRS classes differ slightly from the FEMA classes and are:

1. Open terrain,
2. Tall weeds and crops,
3. Brush lands and low trees,
4. Forested areas fully covered by trees, and
5. Urban areas with dense human-made structures.

The ASPRS guidelines also call for vertical accuracy to be reported in three different ways depending on the ground cover of the area being mapped or imaged:

† While NSSDA assumes that the two quotes in this sentence refer to the same statistic, they, in fact, imply two different statistics. The first quote refers to a "confidence level," which in statistics is the measure of reliability of the parameter being estimated, in this case the RMSE. The second quote refers to the estimated distribution of errors. The difference between these two statistics will be discussed in more detail later in this chapter.

1. "Fundamental vertical accuracy" is computed only from samples measured in open terrain and relies on the NSSDA equations for calculating accuracy.
2. "Supplemental accuracy" is measured from samples taken in nonopen terrain cover types and is determined using the "95th percentile error" method, which is defined as the "absolute value in a data set of errors. It is determined by dividing the distribution of the individual sample errors in the data set into 100 groups of equal frequency. " By definition, 95% of the sampled errors will be less than the 95th percentile value.
3. "Consolidated vertical accuracy" is a combination of the samples from both open terrain and other ground cover classes and is reported as a 95th percentile error.

Guidelines for Digital Elevation Data

The National Digital Elevation Program (NDEP) *Guidelines for Digital Elevation Data* (NDEP, 2004) essentially mirror the ASPRS (2004) lidar guidelines for vertical accuracy reporting in calling for the computation of Fundamental Vertical Accuracy, Supplemental Vertical Accuracy, and Consolidated Vertical Accuracy. Both documents also mandate that errors higher than the 95th percentile be documented in the metadata. NDEP refers to this aspect of its standard as its "truth in advertising approach."

POSITIONAL ACCURACY ASSESSMENT DESIGN AND SAMPLE SELECTION

Positional accuracy assessment requires the appropriate selection of samples to estimate the statistical parameters of the population of errors (e_i) occurring in the spatial data being assessed. Parameters such as the mean (μ), standard deviation (σ), and standard error (σ_μ) characterize the distribution of the population of errors and the reliability of estimators. The mean (μ) is the expected value of a random variable. In the case of positional accuracy, the mean is the expected error, which is usually estimated by the root-mean-square error, or RMSE. The standard deviation (σ) is the square root of the population variance. The variance measures how much the variables of a population deviate from the population mean. The standard error (σ_μ) is the square root of the variance of the estimate of the mean. It measures how estimates of the population mean will deviate from the true mean and is used to create a confidence interval around an estimate of the mean. Equations for calculating these variables and their estimators are presented in the following text.

Estimating positional error parameters requires the comparison of coordinates and/or elevations of identical sample locations from:

- The spatial data set to be assessed (map or imagery) and
- The reference data, which must be an "independent source of higher accuracy" (FGDC, 1998).

We rely on samples because measuring every point in the geospatial data set being assessed would be prohibitively expensive, and sampling can provide highly reliable estimates of the error population's parameters.

NSSDA (FGDC, 1998) outlines several requirements that govern positional accuracy sampling design and collection. They are:

Data independence. To ensure the objectivity and rigor of the assessment, it is critically important that the reference data be independent from the data being tested. In other words, the reference data cannot have been relied upon during the creation of the map or image being assessed. Thus, control points or digital elevation models used to create the spatial products being tested are unsuitable sources of reference data.

Source of reference data. The source of the reference points depends on a number of factors. In some cases, a map of larger scale than the map or image being assessed may provide sufficiently detailed reference coordinates. This is especially true if the map/image to be tested is small in scale and covers a large area. In other cases, such as engineering site drawings, much more precision is required for the reference data points; a field survey or use of a high-precision GPS may be required. NSSDA (FGDC, 1998) stipulates that the reference source data "be of the highest accuracy feasible and practicable." Other handbooks suggest that the reference data be from one to three times more accurate than the anticipated accuracy of the data being tested (Ager, 2004; MPLMIC, 1999; NDEP, 2004; ASPRS, 2004).

Number of samples. The NSSDA (FGDC, 1998) requires a minimum of 20 sample points. Other standards require a minimum of 20 samples per ground cover class and suggest that at least 30 sample points per class are preferred (NDEP, 2004; ASPRS 2004). For statistical rigor, more than 20 sample locations should be chosen. Fewer than 20 points do not provide sufficient samples for a statistically valid estimate. If the population of errors is normally distributed, as illustrated in Figure 3.3, then taking more than 30 samples results in the effort or cost required to collect additional samples exceeding the additional sample's contribution to the precision of the accuracy estimate. If the distribution of the population of errors is skewed or flat, then the sample size should be increased. Lopez et al. (2005) argue that at least 100 samples points are required to achieve a 95% confidence level. However, collecting reference samples, especially ground survey locations, can be extremely expensive and most positional accuracy assessments rely on the NSSDA minimum of 20 samples.

As an alternative, the number of samples required to meet a specified probability level can be calculated as long as reliable approximations of the mean and standard deviation are available. The equations for this calculation can be found in Appendix 3.1 of this chapter.

Identification of samples. The samples must consist of "well-defined points" that "represent a feature for which the horizontal position is known to a high degree of accuracy and position with respect to the geodetic datum" (FGDC, 1998). What constitutes a "well-defined point" will vary with the scale of

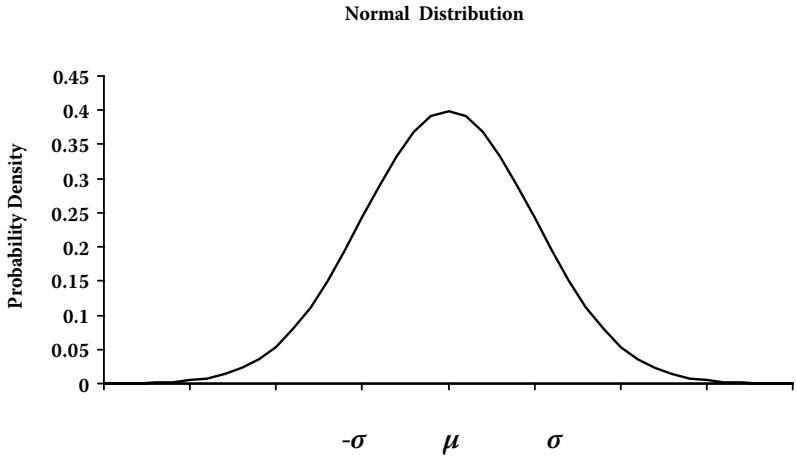

FIGURE 3.3 Shape of the normal distribution.

the map or imagery being assessed. Each point must be clearly identifiable on the spatial data being assessed and in the reference data set. Any error in locating the test points can significantly impact the positional accuracy results. NSSDA (FGDC, 1998) suggests that, "For graphic maps and vector data, suitable well-defined points represent right-angle intersections of roads, railroads, or other linear mapped features, such as canals, ditches, trails, fence lines, and pipelines. For orthoimagery, suitable well-defined points may represent features such as small isolated shrubs or bushes, in addition to right-angle intersections of linear features. For map products at scales of 1:5000 or larger, such as engineering plats or property maps, suitable well-defined points may represent additional features such as utility access covers, and intersections of sidewalks, curbs, or gutters."

Distribution of samples. The sample points must also be well distributed across the project area, and represent the full variety of topography, as topography has the largest impact on positional accuracy. Several options are available for distributing samples across the map or image being assessed:

- The points may be randomly selected using a random number generator. However, the sample points must be identifiable on the imagery or map being assessed, as well as on the reference data. Only a subset of the total population of map or image points will be identifiable. Additionally, locating random points in the field can be problematic if the points fall on private property or inaccessible terrain.
- NSSDA suggests that samples "may be distributed more densely in the vicinity of important features and more sparsely in areas that are of little or no interest." However, emphasis on "important features" will most likely result in a biased sample that may produce biased estimates of the error population parameters.

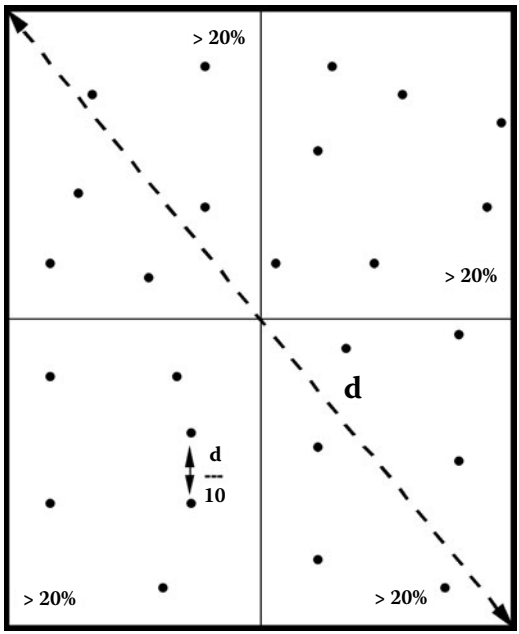

FIGURE 3.4 ASPRS (1989) suggested distribution of positional accuracy assessment sample locations.

- Figure 3.4 illustrates the ASPRS-suggested (1989) systematic sampling method, which ensures that the sample points are well distributed throughout the map or image being assessed. To implement the ASPRS sample distribution, first, the map or image is divided into quadrants. Next, a minimum of 20% of the sample points are allocated to each quadrant. To ensure adequate spacing between the sample points, no two points should be closer than $d/10$ distance from each other, where d is the diagonal dimension of the map or image. This spacing will minimize spatial autocorrelation (a topic that will be discussed in detail in later chapters). In addition, using the ASPRS systematic sample distribution requires assuming that the sample distribution is not correlated with map or image error. This is a reasonable assumption because most positional error is correlated with topography, and topography is rarely distributed on a grid pattern.

In summary, the design and collection of positional accuracy assessment requires the simultaneous consideration of several factors. Often there is a trade-off between well-distributed sample points and easily identifiable sample points. It is not uncommon for some of the desired sample points to fall on private land which may be inaccessible if a ground survey is being used as the reference data. Often, easily identifiable points are concentrated in small areas or are not evenly distributed throughout the map. Care must be taken to obtain the best possible combination of good test points that are appropriately distributed throughout the map or image being assessed.

HOW IS POSITIONAL ACCURACY ANALYZED?

Analyzing positional accuracy involves using sample data to estimate the fit of the spatial data layer (map or image) being assessed to the reference layer, which is assumed to be correct. The accuracy of the fit is depicted by characterizing the distribution of error using the error population's estimated mean, standard deviation, and standard error. Because much confusion exists between the commonly used accuracy standards, we will begin this section with a review of the basic statistics and then move on to the specific equations for depicting positional accuracy.

Review of Basic Statistics

The concepts in this section may be found in any standard statistics textbook. Documents directly relied upon for this text include *Principles of Error Theory and Cartographic Applications* (Greenwalt and Schultz, 1962, 1968), *Biostatistical Analysis (*Zar, 1974), and *Analysis and Adjustment of Survey Measurements (*Mikhail and Gracie, 1981).

This section first provides the equations for calculating and estimating the parameters of a population of values. Next, it discusses the assumptions and equations required to estimate the dispersal of values around the mean. Finally, it provides the equations for calculating a confidence interval around the estimate of the mean.

Parameters and Statistics

The arithmetic mean (μ) of a population of random variables (X_i) is the expected value of any random variable and is calculated by

$$\mu_{X_i} = \sum_i^N X_i / N \tag{3.1}$$

where

X_i = the value of the ith individual in the population, and
N = the total number of individuals in the population.

The mean is estimated from a sample by the variable $\overline{X}$ and is calculated by

$$\overline{X} = \sum_i^n x_i / n \tag{3.2}$$

where

x_i = the value of the ith sample unit chosen from the population, and
n = the total number of sample units chosen.

The standard deviation (σ) is the square root of the population variance, which measures how much the variables of a population deviate from their expected value (i.e., the population mean). The standard deviation is calculated by

$$\sigma = \sqrt{\sum_i^N (X_i - \mu)^2 / (N-1)} \tag{3.3}$$

where X_i, μ, and N are defined as before.

The standard deviation is estimated from a sample by the variable S and is calculated by

$$S = \sqrt{\sum_{i}^{n} (x_i - \bar{X})^2/(n-1)} \tag{3.4}$$

where x_i, $\bar{X}$, and n are defined as before.

A final key parameter in statistics is the standard error $(\sigma_{\bar{X}})$, which helps characterize the spread in the distribution of the possible means, which could be derived from a single *sample* of a population (rather than the entire population itself). According to the central limit theorem, the standard error,† which is the square root of the variance of the population of estimated means, is a valuable parameter because it allows us to estimate our confidence in our estimate of the mean. There is a population of possible estimated means (instead of just one) because there are many possible values of $\bar{X}$, each resulting from a different selection of samples of size n from the population.

The standard error is calculated by

$$\sigma_{\bar{X}} = \sigma/\sqrt{n} \tag{3.5}$$

where σ and n are defined as before.

The standard error is estimated from a sample by the variable $S_{\bar{X}}$ and is calculated by

$$S_{\bar{X}} = S/\sqrt{n} \tag{3.6}$$

where S and n are defined as before.

Estimating the Dispersal of Variables

Assuming that the frequency of the values of variables is normally distributed about the mean as depicted in Figure 3.3, the normal or Gaussian distribution can be used to approximate the distribution of population variables. Additionally, the standard normal distribution can be used to estimate an interval of X_i at specified probabilities within which the mean of the population (μ) will fall. To do so, the distribution of the population variables must be standardized by transforming the scale of the standard normal distribution to the scale of the population being studied.

Figure 3.3 illustrates the shape of the normal distribution. All normal distributions are shaped like the curve in Figure 3.3, with the area underneath the curve equal to 1. The *standard* normal distribution represents the distribution of the standard normal

† The term *standard error* is unfortunately used to denote different parameters in different professions. While most statistics texts define the standard error as the square root of the variance of the population of means $(\sigma_{\bar{X}})$, many mapping texts define the standard error as the square root of the variance of the population signified by σ, which statisticians call the *standard deviation*.

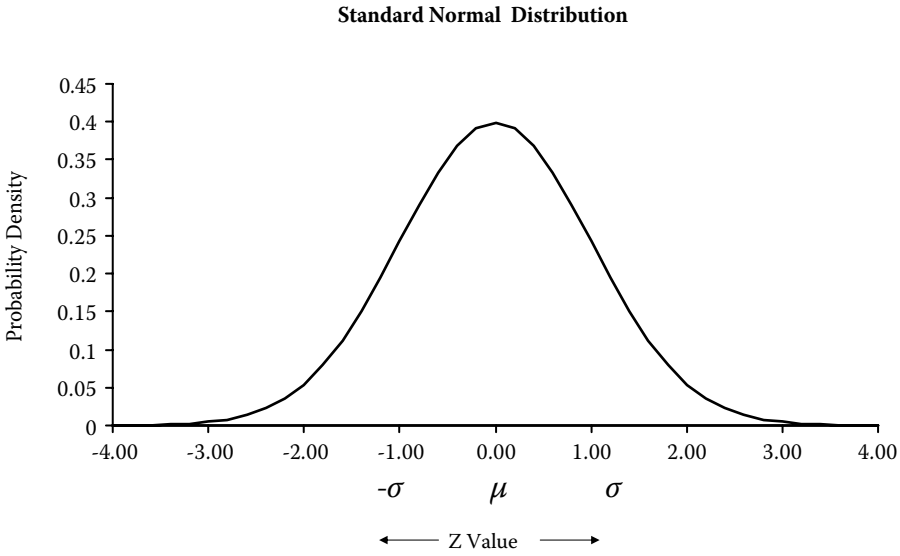

FIGURE 3.5 The standard normal distribution.

variable (Z_i) and is unique because it has a mean of 0 and a standard deviation of 1, as illustrated in Figure 3.5.

The standard normal variable, Z_i, is defined as

$$Z_i = (X_i - \mu)/\sigma \tag{3.7}$$

where

Z_i is the value from the *x*-axis of the standard normal distribution at the *i*th probability level,

X_i is the corresponding value from the *x*-axis of the population of interest, and

μ and σ are defined as before.

Using algebra, we can transform the *x*-axis scale of the normal distribution to that of our population by solving for values of X_i such that

$$Z_i * \sigma = (X_i - \mu) \text{ and}$$

$$X_i = Z_i * \sigma - \mu \tag{3.8}$$

With this formula, we could transform every Z_i value of the standard normal distribution into an X_i value of our population. More commonly, the transformation is used to calculate an interval at a specified probability level within which values of X_i will occur such that:

$$X_i < \mu < X_i, \text{ or using Equation 3.8, the interval becomes}$$

$$[\mu - Z_i * \sigma, \mu + Z_i * \sigma] \tag{3.9}$$

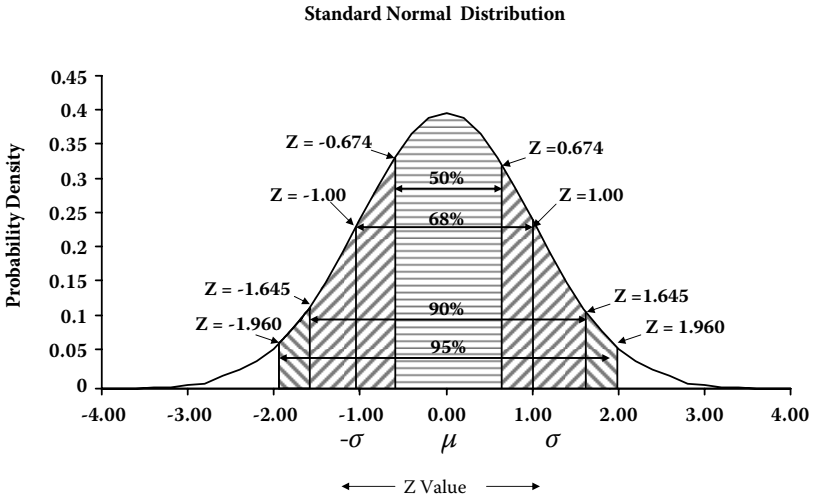

FIGURE 3.6 Probability areas and corresponding Z_i values of the standard normal distribution.

Transforming the values of the normal standard distribution into that of our population of interest requires that the distributions of variables of the normal distribution and of the population of interest be almost identical. This is not a big leap of faith, as the normal distribution characterizes a multitude of natural phenomena ranging from organism population dynamics to human polling behavior. However, it is always important to fully understand whether or not the population you are studying is actually normally distributed or not. Equation 3.9 expresses the dispersal around the mean of the variable X_i at the stipulated probability level if and only if, the population of X_i's is normally distributed. Figure 3.6 illustrates the portions of the normal distribution and the corresponding Z_i values that match various levels of probability.

To summarize, determining the interval at a specific probability within which the mean (μ) of our population of interest will fall requires simply:

1. Assuming that the population is normally distributed,
2. Looking up the Z_i value for the specified probability level in a standard normal table (which may be found in the back of any statistics text or by searching on the Internet),
3. Multiplying the Z_i value times the standard deviation (σ) of the population of interest, and
4. Adding and subtracting the resulting $Z_i * \sigma$ value from the mean (μ).

For example, the interval within which 90% of the values of a normally distributed population with a mean (μ) of 20 and a standard deviation (σ) of 4 can be determined by:

1. Looking up the Z_i value for 90% probability in a Z table or from Figure 3.6. At 90% probability, Z_i is equal to 1.645.

2. Calculating $Z_i * \sigma$ by multiplying 1.645 times the standard deviation of 4, which equals 6.58.
3. Adding and subtracting 6.58 from the mean to determine the interval at 90% probability:

$$= 20 - 6.58, \quad 20 + 6.58$$

which results in the interval ranging from 13.42 to 26.58.

Therefore, we know that 90% of the values of our population will fall within a range between 13.42 and 26.58. Figure 3.7 shows how the *x*-axis scale of the standard normal distribution transforms to that of our example.

Usually we do not know the true mean and the standard deviation of the population. However, because $\bar{X}$ and S are unbiased estimators of μ and σ, we can use the sample estimates of the mean $(\bar{X})$ and the standard deviation (S) to calculate the interval, which becomes

$$\bar{X} - Z_i * S < \mu < \bar{X} + Z_i * S \tag{3.10}$$

Estimating the Reliability of the Estimate of the Mean

Often we want to understand how reliable our estimate of the mean is. To do so requires using the sample data to develop a "confidence interval" around the estimated mean. Estimating the confidence interval once again employs the standard normal variable (Z_i), which, for the population of sample means is defined as

$$Z_i = (\bar{X}_i - \mu)/\sigma_{\bar{X}} \tag{3.11}$$

and can be estimated by

$$Z_i = (\bar{X}_i - \bar{X})/S_{\bar{X}} \tag{3.12}$$

where $\bar{X}_i$ is the value from the population of estimated means that corresponds to the Z_i value from the normal distribution and $\mu, \sigma, \bar{X}$, and $S_{\bar{X}}$ are defined as before.

The confidence interval on the estimate of the mean is calculated as

$$\bar{X} - Z_i(S_{\bar{X}}) < \mu < \bar{X} + Z_i(S_{\bar{X}}) \tag{3.13}$$

when sample sizes are large, and by

$$\bar{X} - t_i(S_{\bar{X}}) < \mu < \bar{X} + t_i(S_{\bar{X}}) \tag{3.14}$$

when sample sizes are small, where t_i is the value from the *x*-axis of the Student's *t* distribution† at the *i*th probability level.

† The Student's *t* distribution should be used instead of the *Z* distribution when sample sizes are below 30. Similar to the *Z* distribution, the values of the Student's *t* distribution can be found at the back of any statistics text or on the Web.

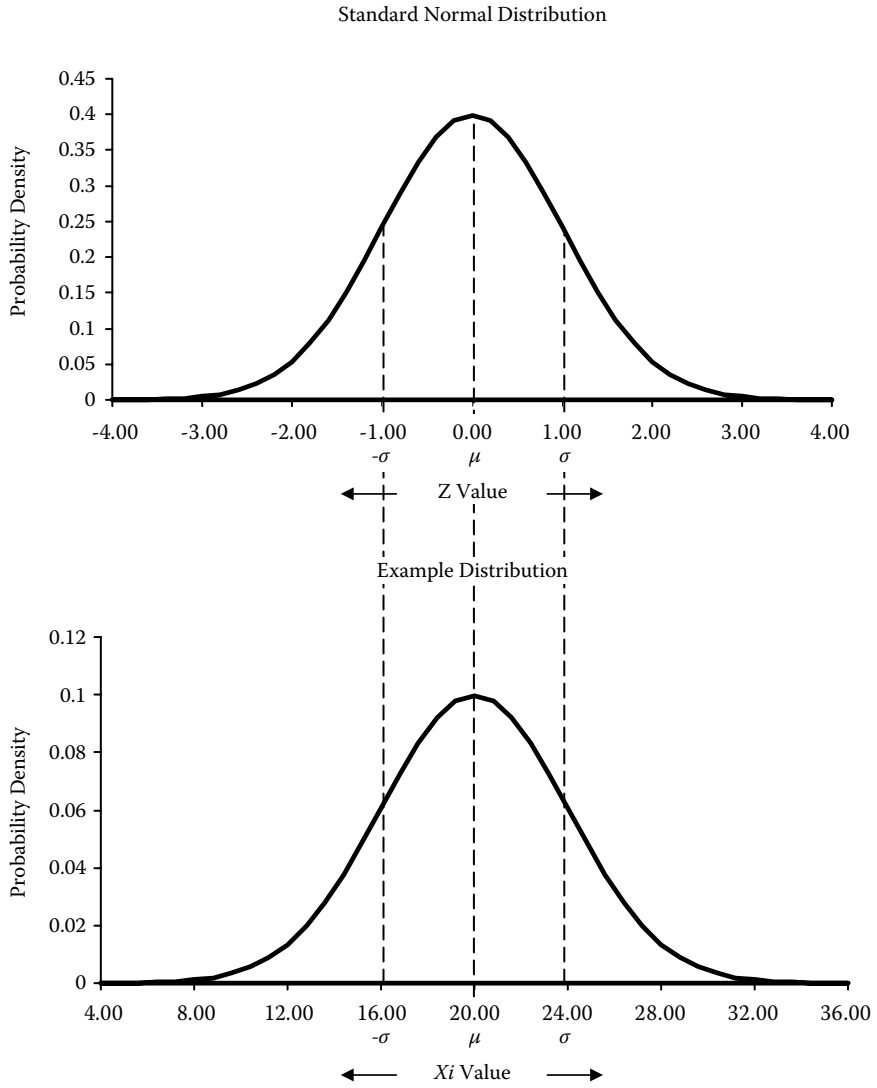

FIGURE 3.7 Transformation of the standard normal distribution *x*-axis scale to the *x*-axis scale of the example.

There is a subtle but very important distinction between estimating the *dispersal* interval of population values around the mean at a specified probability as calculated in Equation 3.10, and the calculation of a *confidence interval* around the estimate of the mean as shown in Equations 3.13 and 3.14. The former expresses the dispersal of a population of values around the mean at specified probabilities. The latter expresses the reliability of the estimate of the mean at specified probabilities.

An interesting aspect of the population of sample means is that it will be normally distributed even when the underlying population of variables is not. This important concept, which is derived from the *central limit theorem*, tells us that when the sample size is large enough and the samples are chosen without bias, then the distribution of the population of means *will* be normally distributed even when the population distribution from which the samples were chosen to estimate the means *is not* normally distributed. The central limit theorem permits us to state our confidence in our estimate of $\bar{X}$ regardless of the distribution of X_i, which allows us to rely on, rather than assume, that the shape of the standard normal distribution is the shape of the distribution of sample means.

Statistics in Positional Accuracy Assessment

In positional accuracy assessment, the NSSDA-specified and accepted measure of accuracy is the mean square root of squared differences between the map and the reference points. This term is called the root-mean-square error, or RMSE. RMSE is estimated from a sample of map and reference points. The mean square root of the square of the differences is used instead of the mean of the simple arithmetic differences to compensate for the fact that the errors can have both positive and negative values. An alternative estimator that would also deal with negative values would be to take the absolute value of the arithmetic mean of the errors.

The estimate of the standard deviation (S) of the squared differences is also an important parameter in many positional accuracy assessment standards (Greenwalt and Schultz, 1962, 1968; ASPRS, 1989). This chapter also suggests the use of the estimated standard error of the RMSE (S_{RMSE}) to build a confidence interval around the estimate of RMSE.

All positional accuracy parameters are estimated by comparing reference coordinates or elevations to the map or image coordinates or elevations of the data set being assessed at each sample location. Unfortunately, positional accuracy standards often confuse the estimate of the mean error (RMSE) with the estimate of the standard deviation (S), and the estimate of the standard deviation with that of the standard error (S_{RMSE}). In addition, terms that are commonly used in other disciplines are often applied slightly differently in positional accuracy assessment, which also adds to the confusion.

For example, statisticians and mapping professionals use the term "root-mean-square error," or RMSE, to refer to different error population parameters. "Error" in positional accuracy is the difference between the reference location and that of the geospatial data set being assessed. It is a measure of accuracy and measures the magnitude of an inaccurately estimated or calculated value. "Error" in statistics is the difference between an observed value and its statistical estimator, and is a measure of precision.

As a result, the equations for RMSE differ between the two applications:

- Mapping professionals define RMSE as the square root of the mean squared differences between the sample reference sample locations and

the corresponding locations on the geospatial data set being assessed. The equation for calculating RMSE in mapping applications is

$$\text{RMSE} = \sqrt{\sum_i^n (e_i)^2/n} \tag{3.15}$$

where

$$e_i = e_{ri} - e_{mi} \tag{3.16}$$

and

e_{ri} equals the reference elevation at the ith sample point,
e_{mi} equals the map or image elevation at the ith sample point, and
n is the number of samples.

- Statisticians define RMSE as the square root of the mean squared differences between a statistical estimator of a parameter and the value actually observed. The equation for calculating RMSE in statistics is

$$\sqrt{\sum_i^n (e_i - \bar{e})^2/(n-1)} \tag{3.17}$$

where $\bar{e}$ is the unbiased estimator of the mean or average difference, which is calculated by

$$\bar{e} = \sqrt{\sum_i^n (e_i)^2/n} \tag{3.18}$$

The only time that the mapping and the statistical RMSE are equal to one another is when the average error $(\bar{e})$ equals zero, which is a condition that is rare and which should always be tested for by calculating $(\bar{e})$ and determining if it is significantly different from zero. Unfortunately, many mapping standards make the assumption that $(\bar{e})$ equals zero, which results in misleading characterizations of map error.

This use of the same term to mean different things has, understandably, led to much confusion. The RMSE used in mapping is the square root of the estimated mean of the squares of the geospatial data set's positional errors. The RMSE used in statistics is the square root of the variance of the errors and characterizes how errors differ from the mean error. Because mapping professionals rely on statistics to characterize the frequency distribution of positional errors, it is critical that confusion be minimized and that statistics be properly applied to mapping applications.

This chapter attempts to eliminate the confusion surrounding positional accuracy assessment analysis by:

1. Detailing what equations should be used to characterize positional accuracy, and
2. Correcting the mistakes in currently used standards.

First, one-dimensional vertical accuracy assessment is discussed. Next, two-dimensional horizontal accuracy assessment is reviewed.

Vertical Accuracy

Statistical Parameters

The mean vertical positional error (μ_v)† is depicted in mapping applications by the vertical root-mean-square error ($RMSE_v$) of the sample of vertical errors (e_{vi}) and is estimated by

$$RMSE_v = \sqrt{\sum_i^n (e_{vi})^2/n} \tag{3.19}$$

where

$$e_{vi} = v_{ri} - v_{mi} \text{ and} \tag{3.20}$$

v_{ri} equals the reference elevation at the ith sample point,
v_{mi} equals the map or image elevation at the ith sample point, and
n is the number of samples.

An alternative estimator is the arithmetic mean of the absolute error values and is calculated by

$$|\bar{e}_v| = \sum_i^n |e_{vi}|/n \tag{3.21}$$

The standard deviation (σ_v) of the population of vertical errors is estimated by

$$S_v = \sqrt{\sum_i^n (e_{vi} - RMSE_v)^2/(n-1)} \tag{3.22}$$

and the standard error of estimates of $RMSE_v$ is estimated by

$$S_{RMSE_v} = S_v/\sqrt{n} \tag{3.23}$$

Assuming that the vertical errors are normally distributed, the estimated interval of errors at a specific probability can be expressed as

$$RMSE_v \pm Z_i(S_v) \tag{3.24}$$

At a 95% probability level, the equation becomes

$$RMSE_v \pm 1.96\,(S_v) \tag{3.25}$$

If RMSE equals zero, then the factor $\pm Z_i(S_v)$ will express the interval of error at the probability level specified by the Z_i variable and the interval at 95% will equal ±1.96 (S_v). A 90% interval, with $RMSE_v$ to zero, will be 1.645 (S_v).

† Some mapping texts use the subscript "z" to denote vertical error. Because this text (and most statistics texts) uses the variable Z_i to denote the standard normal variable, we use the subscript "v" to denote vertical error.

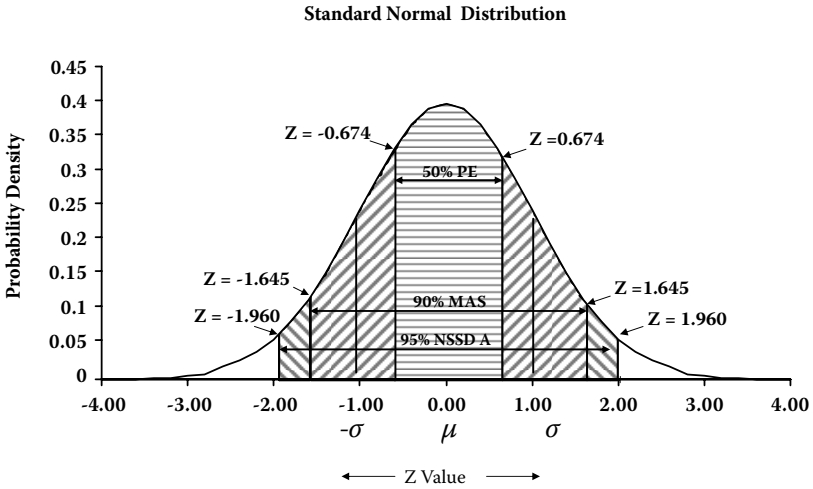

FIGURE 3.8 Areas and Z_i values of the standard normal distribution corresponding to the probability levels of various map accuracy standards.

The *Principles of Error Theory and Cartographic Applications* report is the first report to propose use of the $Z_i(S_v)$ interval as a standard in estimating positional accuracy (Greenwalt and Schultz, 1962, 1968). The report relies on estimating the interval $Z_i(S_v)$ at various probability levels where it is referred to as the probable error at 50% and the map accuracy standard at 90% (Greenwalt and Schultz, 1962, 1968). Figure 3.8 illustrates the portions of the normal distribution that correspond to the probable error at 50%, the map accuracy standard at 90%, and the NSSDA standard at 95%.

The Greenwalt and Schultz (1962, 1968) equations estimate the interval of errors around the mean error at different probability levels. Derived from the military science of ballistics, the equations result in an estimate of the probable dispersal of error around the mean error ($RMSE_z$) at specified probabilities.† Use of the interval as an accuracy standard was confirmed by subsequent Defense Mapping Agency reports (DMA, 1991), and is used in ASPRS's large-scale mapping standards (ASPRS, 1989), as well as NSSDA (FGDC, 1998).

Note that the $Z_i(S_v)$ interval is not a confidence interval around the estimate of $RMSE_v$, nor is it the range of expected errors at a given probability. Rather, it is an estimate of the maximum interval of error that will exist at a specified probability assuming that mean error equals zero and the errors are normally distributed. Unfortunately, spatial errors are often biased and interrelated, bringing the assumption of normality into question.

† Greenwalt and Schultz (1962, 1968) define the vertical map accuracy standard as "the size of error which 90% of the elevations will not exceed." However, the interval $Z_i(S_z)$ meets this definition only when $RMSE_v$ equals zero. The estimated size of elevation errors which will not be exceeded at a probability level specified by Z_i is $RMSE_v \pm Z_i(S_z)$.

To measure the reliability (or our confidence in) the estimate of $RMSE_v$, we calculate a confidence interval around $RMSE_v$, by converting the general confidence interval equation (Equation 3.13)

$$\bar{X} - Z_i(S_{\bar{X}}) < \mu < \bar{X} + Z_i(S_{\bar{X}})$$

to our mapping application terminology such that

$$RMSE_v - Z_i(S_{RMSE}) < \mu < RMSE_v + Z_i(S_{RMSE}) \tag{3.26}$$

for large sample sizes, and

$$RMSE_v - t_i(S_{RMSE}) < \mu < RMSE_v + t_i(S_{RMSE}) \tag{3.27}$$

for small samples sizes, where all variables are defined as before.

In most situations, if we have more than 30 samples, at a 95% confidence level the equation becomes:

$$RMSE_v - 1.96(S_{RMSE}) < \mu < RMSE_v + 1.96(S_{RMSE}). \tag{3.28}$$

That means that we are 95% certain that the interval contains the true, but unknown, population average error.

Table 3.1 displays the map and reference elevations for a hypothetical digital elevation data set. The errors at each sample point are calculated as well as the estimated $RMSE_v$, $|\bar{e}_v|$, S_v, S_{RMSE_v}, NSSDA accuracy statistic, MAS, and a 95% confidence interval around the estimate of $RMSE_v$. All the equations for these calculations are shown in Table 3.2.

NSSDA

The NSSDA (FGDC, 1998) requires that accuracy be reported at the 95% level, which is defined by NSSDA as meaning "that 95% of the positions in the data set will have an error with respect to true ground positions that is equal to or smaller than the reported accuracy." NSSDA references the Greenwalt and Schultz (1962, 1968) equations, but *mistakenly* stipulates that the vertical accuracy interval at the 95% probability be computed by multiplying the appropriate Z_i statistic times the *estimated mean* ($RMSE_v$) instead of the *estimated standard deviation* (S_v). The resulting NSSDA equation for calculating the NSSDA vertical accuracy statistic is†

$$NSSDA\ Vertical\ Accuracy_v = 1.96\ (RMSE_v) \tag{3.29}$$

rather than the Greenwalt and Schultz (1962, 1968) equation, which is

$$Accuracy_v = 1.96\ (S_v). \tag{3.30}$$

Estimating the interval within which 95% of the errors will fall requires assuming that our errors are normally distributed and converting the scale of the standard

† 1.96 is the standard normal distribution Z statistic (the value from the x-axis of the standard normal distribution) for an interval with a probability of 95%.

TABLE 3.1
Vertical Accuracy Example

Point ID	v_{ri} Reference	v_{mi} Map	Error = e_{vi} = Reference − Map ($v_{ri} - v_{mi}$) = e_{vi}	Error Squared $(v_{ri} - v_{mi})^2 = e_{vi}^2$	Absolute Error	(Absolute e_{vi} − $RMSE_v$)^2
1202	2362.2075	2361.3100	−0.8975	0.8055	0.8975	0.3502
1230	2421.5855	2420.9000	−0.6855	0.4699	0.6855	0.1442
1229	2701.6110	2701.1700	−0.4410	0.1945	0.4410	0.0183
125	705.3117	705.0190	−0.2927	0.0857	0.2927	0.0002
316	1009.2344	1009.0300	−0.2044	0.0418	0.2044	0.0103
369	920.0574	919.8740	−0.1834	0.0336	0.1834	0.0150
292	586.3659	586.2400	−0.1259	0.0159	0.1259	0.0323
143	761.4684	761.3910	−0.0774	0.0060	0.0774	0.0521
132	712.1791	712.1320	−0.0471	0.0022	0.0471	0.0669
1005	1190.4284	1190.4000	−0.0284	0.0008	0.0284	0.0769
274	809.0433	809.0500	0.0067	0.0000	0.0067	0.0894
112	387.2611	387.2960	0.0349	0.0012	0.0349	0.0734
339	965.6910	965.7480	0.0570	0.0032	0.0570	0.0619
130	1059.1342	1059.2300	0.0958	0.0092	0.0958	0.0441
113	428.7700	428.9630	0.1930	0.0372	0.1930	0.0127
122	1012.0117	1012.3100	0.2983	0.0890	0.2983	0.0001
136	308.7100	309.0110	0.3010	0.0906	0.3010	0.0000
104	529.4721	529.8260	0.3539	0.1252	0.3539	0.0023
101	427.1653	427.5840	0.4187	0.1753	0.4187	0.0128
1221	2690.1380	2689.5200	−0.6180	0.3819	0.6180	0.0975
129	483.4317	483.0480	−0.3837	0.1472	0.3837	0.0061
128	492.7014	492.5810	−0.1204	0.0145	0.1204	0.0344
114	799.9452	799.8560	−0.0892	0.0080	0.0892	0.0469
367	1273.0857	1273.0300	−0.0557	0.0031	0.0557	0.0625
108	1235.0128	1235.0300	0.0172	0.0003	0.0172	0.0833
325	1040.9078	1040.9700	0.0622	0.0039	0.0622	0.0593
250	211.4375	211.5230	0.0855	0.0073	0.0855	0.0485
1010	1189.4876	1189.6200	0.1324	0.0175	0.1324	0.0300
Sum				**2.77**	**6.31**	**1.60**

normal distribution to that of the error population using the estimated standard deviation as detailed in Equation 3.10 below.

$$[\bar{X} - Z_i(S), \quad \bar{X} + Z_i(S)]$$

If the mean error is equal to zero, then the interval becomes the Greenwalt and Schultz (1962, 1968) statistic of $Z_i(S)$. While the NSSDA standard is applied ubiquitously, it is valid only when $RMSE_v = S_v$. If S_v is less than $RMSE_v$ then the NSSDA statistic will overestimate the error interval, and if S_v is greater than $RMSE_v$, the NSSDA statistic will underestimate the error interval.

TABLE 3.2
Vertical Accuracy Example Equations and Statistics

Definition	Equation	Value
Estimated root-mean-square error of the population of vertical errors	$\text{RSME}_v = \sqrt{\sum_i^n (e_{vi})^2/n}$	0.320
Estimated absolute arithmetic mean of the population of vertical errors	$\lvert\bar{e}_v\rvert = \sum_i^n \lvert e_{vi}\rvert/n$	0.234
Estimated variance of the population of vertical errors	$S_v^2 = \sum_1^n (e_{vi} - \text{RMSE}_v)/(n-1)$	0.059
Estimated standard deviation of the population of vertical errors	$S_v = \sqrt{\sum_i^n (e_{vi} - \text{RMSE}_v)^2/(n-1)}$	0.244
Estimated standard error of the population of RMSEs	$S_{\text{RMSE}_v} = S_v/\sqrt{n}$	0.047
Greenwalt and Schultz MAS standard normal interval of e_{vi} at 90% probability	$1.645 * S_v$	0.401
Greenwalt and Schultz standard normal interval of e_{vi} at 95% probability	$1.96 * S_v$	0.478
NSSDA statistic	$1.96 * \text{RMSE}_v$	0.628
95% confidence interval around the estimate of RMSE	$\text{RMSE} \pm 1.96 * S_{\text{RMSE}}$	RMSE ± 0.092 which results in a range from 0.228 to 0.412

FEMA, ASPRS LIDAR, and NDEP Standards

The most recent standards and guidelines for elevation data all require that the landscape be stratified by vegetative cover class, that a minimum of 60 samples be chosen, and that the different types of cover classes be evaluated using different methods. While open terrain is evaluated using NSSDA equations, other cover classes are evaluated using the 95th percentile method, as is the consolidated vertical accuracy, which combines open terrain samples with other ground cover types.

Figure 3.9 shows an example of positional accuracy analysis with each sample point sorted by ground cover type and charted by the error measured between each points reference and map data. Both the NSSDA statistic at ±0.82 ft, and the consolidated vertical accuracy at ±0.91 ft are displayed.

Horizontal Accuracy

Statistical Parameters

Horizontal accuracy is more complex than vertical accuracy because the error is distributed in two dimensions (both the *x* and *y* dimensions), requiring the calculation of the radial error and reliance on the bivariate normal distribution to estimate

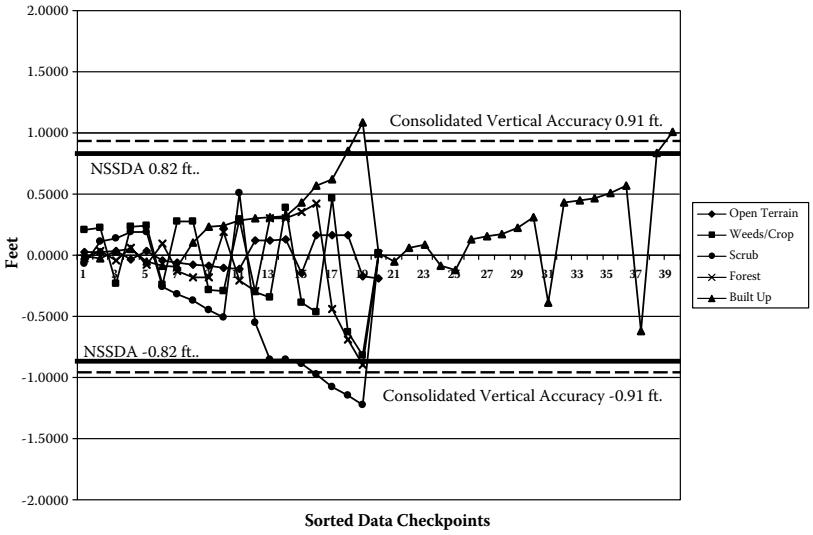

FIGURE 3.9 Application of 95th percentile criteria versus NSSDA for reporting positional accuracy. (Courtesy of Dewberry.)

probabilities. To calculate the horizontal root-mean-square error ($RMSE_h$),† first, the *x*-coordinate from the reference data is recorded followed by the *x*-coordinate from the spatial data set being assessed. Then the difference between the two locations is computed, followed by a squaring of this difference. The same process is used for the *y*-coordinate. Each test point then has an associated error distance, e_i, defined by the following equation:

$$e_h = \sqrt{(x_{ri} - x_{mi})^2 + (y_{ri} - y_{mi})^2} \tag{3.31a}$$

and

$$e_h^2 = (x_{ri} - x_{mi})^2 + (y_{ri} - y_{mi})^2 \tag{3.31b}$$

where x_r and y_r are the reference coordinates and x_m and y_m are the map or image coordinates for the *i*th sample point in the spatial data set being assessed.

The equation for the average horizontal error or horizontal root-mean-square error ($RMSE_h$) is calculated from the errors of the individual test sample points using the following equation:

$$\text{RMSE}_h = \sqrt{\sum_i^n ((x_{ri} - x_{mi})^2 + (y_{ri} - y_{mi})^2)/n} = \sqrt{\left(\text{RMSE}_x^2 + \text{RMSE}_y^2\right)/n} \tag{3.32}$$

† Greenwalt and Schultz (1962, 1968) refer to horizontal error as *circular error,* which they designate with the subscript "c." NSSDA (FGDC, 1998) refers to horizontal error as radial, designated by the subscript "r." Because the errors are usually elliptical rather than circular, and because we have already designated the subscript "r" to indicate a reference value of an accuracy assessment sample, this text uses the subscript "h" to designate horizontal error.

or

$$\text{RMSE}_h = \sqrt{\frac{\sum_i^n e_{hi}^2}{n}} \tag{3.33}$$

where e_{hi} is defined in the preceding equation and n is the number of test sample points.

An alternative estimator is the arithmetic mean of the absolute error values and is calculated by

$$|\bar{e}| = \sum_1^n |e_{hi}|/n \tag{3.34}$$

Once RMSE_h has been estimated, the standard deviation (S_h) of the population of horizontal errors can also be approximated from the samples by calculating the average standard deviation, using the Greenwalt and Schultz (1962, 1968) equation below:

$$S_h = (S_x + S_y)/2 \tag{3.35}$$

where

$$S_x = \sqrt{\sum_i^n ((x_{ri} - x_{mi}) - \text{RMSE}_x)^2/n - 1} \tag{3.36}$$

and

$$S_y = \sqrt{\sum_i^n ((y_{ri} - y_{mi}) - \text{RMSE}_y)^2/n - 1} \tag{3.37}$$

The estimated standard error of the population of RMSE_h's is

$$S_{\text{RMSE}_h} = S_h/\sqrt{n} \tag{3.38}$$

Assuming that the errors are normally distributed, the estimated interval of errors at a specified probability can be expressed as

$$\text{RMSE}_h \pm Z_i(S_h) \tag{3.39}$$

If RMSE_h is equal to zero, then the error interval becomes the Greenwalt and Schultz (1962, 1968) specified Z_i (S).

The confidence interval around the estimate of the mean horizontal error can be calculated as follows:

$$\text{RMSE}_h - Z_i(S_{\text{RMSE}_h}) < \mu < \text{RMSE}_h + Z_i(S_{\text{RMSE}_h}) \tag{3.40}$$

for large sample sizes, and

$$\text{RMSE}_h - t_i(S_{\text{RMSE}_h}) < \mu < \text{RMSE}_h + t_i(S_{\text{RMSE}_h}) \tag{3.41}$$

for small samples sizes.

Because horizontal error is measured in two dimensions, the *bivariate* standard normal distribution must be used to characterize the distribution of errors. Figure 3.10

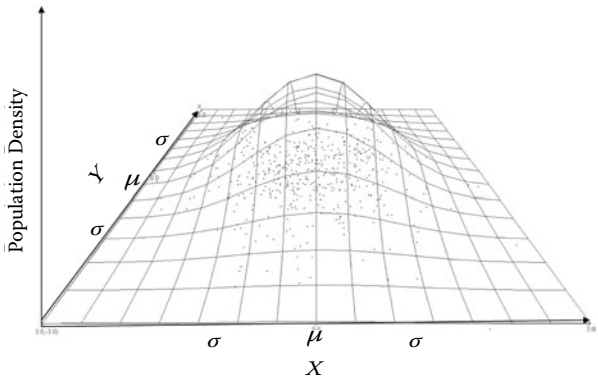

FIGURE 3.10 Three-dimensional representation of the standard normal bivariate distribution.

provides a three-dimensional illustration of the bivariate normal distribution. Figure 3.11 is an overhead view of the bivariate standard normal probability distribution for the commonly used map standards of the circular error probable (CEP) at 50%, the circular map accuracy standard (CMAS) at 90%, and NSSDA at 95%.

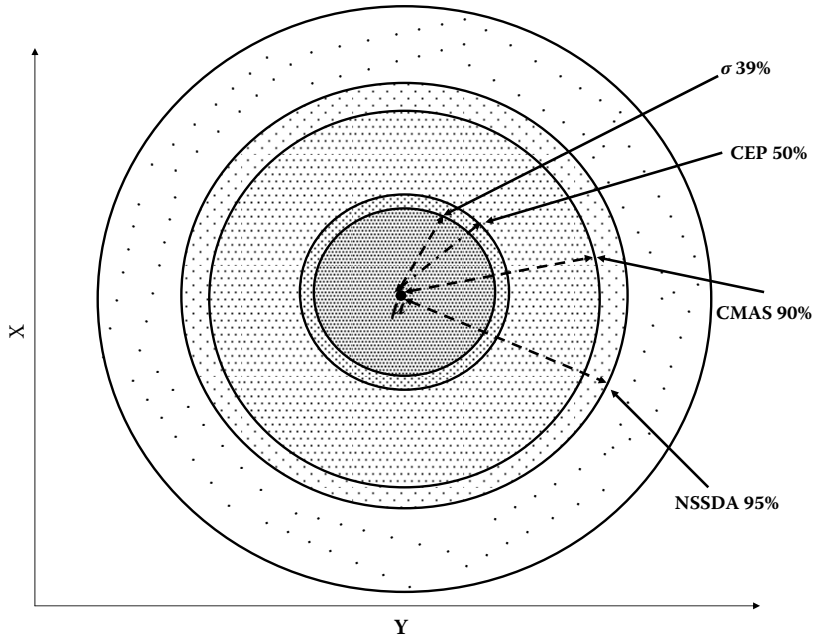

FIGURE 3.11 Two-dimensional representation of the normal standard bivariate or circular distribution with the probabilities of common horizontal map standards (From Greenwalt, C. and M. Schultz. 1962, 1968. *Principles of Error Theory and Cartographic Applications.* United States Air Force. Aeronautical Chart and Information Center. ACIC Technical Report Number 96. St. Louis, MO. 60 pages plus appendices).

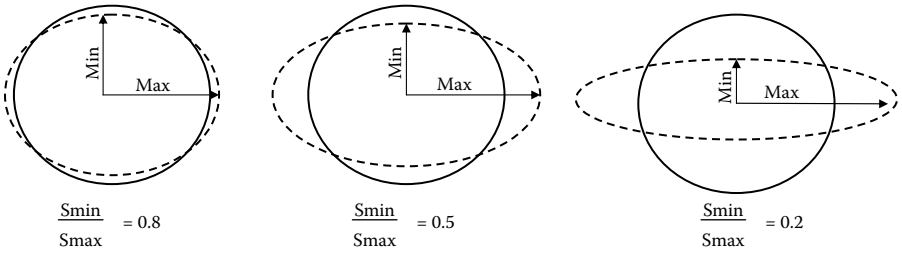

FIGURE 3.12 Comparison of circular to elliptical distributions for various ratios of S_{min}/S_{max} (From Defense Mapping Agency. 1991. *Error Theory as Applied to Mapping, Charting, and Geodesy*. Defense Mapping Agency Technical Report 8400.1. Fairfax, Virginia).

Relying on the bivariate standard normal distribution to characterize the distribution of horizontal errors requires that we assume that the horizontal errors are distributed in a circle with S_x equal to S_y. We can test for circularity by calculating the ratio of the S_{min} to S_{max} (where S_{min} is the lower of S_x or S_y, and S_{max} is the larger of S_x or S_y). Figure 3.12 shows how differences in S_x and S_y affect the shape of the distribution of errors. If the ratio of S_{min} to S_{max} is 0.2 or greater, Greenwalt and Schultz (1962, 1968) state that the circular distribution can be assumed.

As with vertical accuracy, many standards rely on the Z_i (S_h) as the statistic to estimate horizontal accuracy (DMA, 1991; NSSDA, 1998). The statistic estimates the maximum interval of error on either side of RMSE_h that will exist at a specified probability. The bivariate standard normal distribution Z_i statistic at 95% probability is 2.4477 (Greenwalt and Schultz, 1962, 1968) and the resulting interval of errors at 95% probability is

$$2.4477\ ((S_x + S_y)/2) \tag{3.42}$$

or

$$2.4477\ S_h \tag{3.43}$$

The interval of errors within which 95% of the errors will occur (assuming the errors are normally distributed) is

$$[\text{RMSE}_h - 2.447 * S_h,\ \text{RMSE}_h - 2.447 * S_h] \tag{3.44}$$

If RMSE_h is equal to zero, the estimated interval reduces to the Greenwalt and Schultz (1962, 1968) and ASPRS (1989) accuracy statistic of 2.447(S_h).

Because the distribution of RMSE_h is one-dimensional (even though the distribution of errors is two-dimensional), a confidence interval on RMSE_h at the 95% level is expressed by

$$\text{RMSE}_h - 1.96 S_{\text{RMSE}_h} > \mu > \text{RMSE}_h - 1.96 S_{\text{RMSE}_h} \tag{3.45}$$

for large samples and

$$\text{RMSE}_h - t_{95\%,\ n-1\ degrees\ of\ freedom}\, S_{\text{RMSE}_h} > \mu > \text{RMSE}_h + t_{95\%,\ n-1\ degrees\ of\ freedom}\, S_{\text{RMSE}_h} \tag{3.46}$$

for small samples.

NSSDA

As with vertical accuracy, the NSSDA accuracy statistic incorrectly applies the $RMSE_h$ rather than S_h to calculate the NSSDA accuracy statistic under two different conditions described in the following text: when the $RMSE_y$ and $RMSE_x$ are equal and when they are not equal.

When the errors are circular. NSSDA defines errors as circular if $RMSE_y = RMSE_x$ (rather than when $S_x = S_y$). Under NSSDA

$$\text{if } RMSE_y = RMSE_x, \text{ then}$$

$$RMSE_h = \sqrt{(2RMSE_x)^2} = \sqrt{(2RMSE_y)^2}$$

$$= 1.4142 * RMSE_x = 1.4142 * RMSE_y \tag{3.47}$$

Applying the circular error normal distribution Z statistic at 95% probability of 2.4477 results in

NSSDA Horizontal Accuracy $= 2.4477\ RMSE_h$ or

$$= 2.4477 * RMSE_h/1.4142$$

$$= 1.7308 * RMSE_h \tag{3.48}$$

Most organizations use this simplified equation regardless of whether the errors are distributed circularly or not.† However, as with elevational accuracy, the NSSDA horizontal accuracy value has no statistical basis. The population parameter that should be used to determine the interval of error at a specific probability level is the standard deviation of the horizontal errors (S_h), and not $RMSE_h$ (Ager, 2004).

When the errors are not circular. If $RMSE_y \neq RMSE_x$, then the NSSDA stipulates that the NSSDA accuracy statistic is

$$= 2.4477\ ((RMSE_x + RMSE_y)/2) \tag{3.49}$$

Tables 3.3 and 3.4 present sample reference and map coordinates from our earlier example and calculate the $RMSE_h$, S_h, S_{RMSE_h}, $Zi * S$ at 95% probability, as well as the circular map accuracy standard (CMAS) accuracy interval at 90%, the NSSDA statistic, and a 95% confidence interval around $RMSE_h$.

There is one final, very important issue that must be understood when implementing positional accuracy. A different RMSE value is often calculated as part of the spatial data set registration process. We will call this $RMSE_{reg}$. The calculation of $RMSE_{reg}$ during the registration process is a test of the goodness of fit of the registered data set to its control points. Because of its lack of independence from the data set being assessed, $RMSE_{reg}$ is not a valid measure of positional accuracy, and will almost always be lower (i.e., better) than $RMSE_z$ or $RMSE_h$. Independent positional

† Circularity is defined by Greenwalt and Schultz as S_{min}/S_{max} greater than or equal to 0.2. However, NSSDA restricts application of the circular distribution to those situations in which S_{min}/S_{max} is between 0.6 and 1.0.

TABLE 3.3
Horizontal Accuracy Example

Point ID	x_{ri} Reference	x_{mi} Map	Error in x Dimension = Reference − Map $(x_{ri} - x_{mi}) = e_{xi}$	Absolute Error $\lvert e_x \rvert$	Error in x Dimension Squared $(x_{ri} - x_{mi})^2 = e_{xi}^2$	(Absolute e_{xi} − $RMSE_x)^2$	y_{ri} Reference	y_{mi} (Map)	Error in y Dimension = Reference − Map $(y_{ri} - y_{mi}) = e_{yi}$	Absolute Error $\lvert e_y \rvert$	Error in y Dimension Squared $(y_{ri} - y_{mi})^2 = e_{yi}^2$	(Absolute y_i − $RMSE_y)^2$	Sum of Squared Errors $e_{xi}^2 + e_{yi}^2$	Sum of Absolute Error
107	6463928.4275	6463928.2891	0.1384	0.1384	0.0192	0.0388	1740487.9905	1740488.2089	−0.2184	0.2184	0.0477	0.0912	0.0669	0.3569
108	6478942.9446	6478942.9707	−0.0261	0.0261	0.0007	0.0956	1757945.7986	1757945.5996	0.1991	0.1991	0.0396	0.1033	0.0403	0.2251
110	6498179.1383	6498179.2172	−0.0789	0.0789	0.0062	0.0657	1736983.2778	1736983.7799	−0.5021	0.5021	0.2521	0.0003	0.2583	0.5810
111	6500864.5792	6500866.2526	−1.6734	1.6734	2.8004	1.7906	1758833.2498	1758830.8834	2.3664	2.3664	5.5999	3.4075	8.4003	4.0398
116	6527762.0733	6527762.1410	−0.0677	0.0677	0.0046	0.0716	1731210.4027	1731210.7259	−0.3232	0.3232	0.1045	0.0389	0.1091	0.3910
117	6539890.0536	6539890.2650	−0.2113	0.2113	0.0447	0.0154	1755842.1176	1755841.9103	0.2073	0.2073	0.0430	0.0981	0.0876	0.4186
122	6452053.8265	6452053.8601	−0.0336	0.0336	0.0011	0.0910	1728034.3838	1728034.6916	−0.3078	0.3078	0.0948	0.0452	0.0959	0.3414
123	6435447.0261	6435446.7694	0.2567	0.2567	0.0659	0.0062	1737489.6870	1737489.9830	−0.2960	0.2960	0.0876	0.0504	0.1535	0.5527
124	6445012.8528	6445012.7143	0.1385	0.1385	0.0192	0.0387	1757524.8057	1757524.7919	0.0138	0.0138	0.0002	0.2567	0.0194	0.1524
206	6523662.6628	6523662.7526	−0.0898	0.0898	0.0081	0.0603	1753217.8809	1753218.0854	−0.2045	0.2045	0.0418	0.0998	0.0499	0.2944
216	6503988.9073	6503989.0881	−0.1808	0.1808	0.0327	0.0239	1728652.7232	1728653.2982	−0.5750	0.5750	0.3306	0.0030	0.3633	0.7558
222	6497217.5322	6497217.6331	−0.1009	0.1009	0.0102	0.0549	1751316.3332	1751316.3331	0.0001	0.0001	0.0000	0.2708	0.0102	0.1010
227	6532154.2998	6532154.2726	0.0272	0.0272	0.0007	0.0949	1740450.9200	1740451.2630	−0.3430	0.3430	0.1177	0.0315	0.1184	0.3702
228	6514726.6170	6514726.6231	−0.0061	0.0061	0.0000	0.1084	1748724.1696	1748724.4427	−0.2731	0.2731	0.0746	0.0612	0.0746	0.2792
229	6480333.2958	6480333.3200	−0.0242	0.0242	0.0006	0.0968	1742388.0615	1742388.2686	−0.2071	0.2071	0.0429	0.0982	0.0435	0.2313
283	6510536.2705	6510536.4059	−0.1354	0.1354	0.0183	0.0400	1757706.5081	1757706.5519	−0.0438	0.0438	0.0019	0.2272	0.0202	0.1791
200	6509030.6018	6509030.5422	0.0596	0.0596	0.0036	0.0760	1746587.3294	1746587.4407	−0.1113	0.1113	0.0124	0.1674	0.0159	0.1709
112	6502026.5461	6502026.5552	−0.0091	0.0091	0.0001	0.1064	1779378.9142	1779378.7511	0.1631	0.1631	0.0266	0.1277	0.0267	0.1723

(*Continued*)

TABLE 3.3 (CONTINUED)
Horizontal Accuracy Example

Point ID	x_{ri} Reference	x_{mi} Map	Error in x Dimension = Reference – Map $(x_{ri} - x_{mi}) = e_{xi}$	Absolute Error $\lvert e_x \rvert$	Error in x Dimension Squared $(x_{ri} - x_{mi})^2 = e_{xi}^2$	(Absolute e_{xi} $-RMSE_x)^2$	y_{ri} Reference	y_{mi} (Map)	Error in y Dimension = Reference – Map $(y_{ri} - y_{mi}) = e_{yi}$	Absolute Error $\lvert e_y \rvert$	Error in y Dimension Squared $(y_{ri} - y_{mi})^2 = e_{yi}^2$	(Absolute y_i–$RMSE_y)^2$	Sum of Squared Errors $e_{xi}^2 + e_{yi}^2$	Sum of Absolute Error
232	6509030.6018	6509030.5422	0.0596	0.0596	0.0036	0.0760	1793670.4405	1793670.5122	–0.0717	0.0717	0.0051	0.2014	0.0087	0.1313
125	6436524.8263	6436525.0783	–0.2520	0.2520	0.0635	0.0069	1782491.7176	1782490.8878	0.8298	0.8298	0.6886	0.0957	0.7521	1.0818
126	6464717.9797	6464718.1392	–0.1595	0.1595	0.0254	0.0309	1778968.4568	1778968.1097	0.3472	0.3472	0.1205	0.0300	0.1459	0.5066
128	6536017.6536	6536017.6690	–0.0154	0.0154	0.0002	0.1023	1791341.5236	1791341.5819	–0.0583	0.0583	0.0034	0.2136	0.0036	0.0737
207	6523447.4186	6523447.4146	0.0040	0.0040	0.0000	0.1098	1781813.7690	1781813.5640	0.2050	0.2050	0.0420	0.0995	0.0420	0.2090
208	6458661.4231	6458661.4229	0.0002	0.0002	0.0000	0.1123	1763512.1326	1763512.0318	0.1008	0.1008	0.0102	0.1761	0.0102	0.1010
210	6432704.3136	6432704.3469	–0.0333	0.0333	0.0011	0.0912	1797681.4156	1797681.5817	–0.1661	0.1661	0.0276	0.1256	0.0287	0.1994
214	6524150.2574	6524150.1541	0.1033	0.1033	0.0107	0.0538	1766691.1359	1766691.0337	0.1022	0.1022	0.0105	0.1749	0.0211	0.2055
221	6490159.7535	6490159.5953	0.1582	0.1582	0.0250	0.0314	1774521.0769	1774520.9524	0.1245	0.1245	0.0155	0.1568	0.0405	0.2827
223	6464915.1344	6464915.4456	–0.3112	0.3112	0.0968	0.0006	1795190.2232	1795190.4426	–0.2194	0.2194	0.0481	0.0906	0.1450	0.5306
224	6446211.1711	6446211.4834	–0.3123	0.3123	0.0975	0.0005	1776288.7842	1776289.2253	–0.4411	0.4411	0.1946	0.0063	0.2921	0.7534
226	6513283.3804	6513283.4917	–0.1113	0.1113	0.0124	0.0502	1771237.4139	1771237.6203	–0.2064	0.2064	0.0426	0.0986	0.0550	0.3177
		Sum		**4.7780**	**3.3724**	**3.5409**				**9.2277**	**8.1266**	**6.6477**	**11.4990**	**14.0057**

TABLE 3.4
Horizontal Accuracy Example Equations and Statistics

Definitions	*X* Dimension Equations	*X* Dimension Values
Estimated root-mean-square of the population of errors	$\text{RMSE}_x = \sqrt{\sum_i^n (e_{xi})^2/n}$	0.3353
Estimated absolute arithmetic mean of the population of errors	$\lvert\bar{e}_x\rvert = \sum_1^n \lvert\bar{e}_{xi}\rvert/n$	0.1593
Estimated variance of the population of errors	$S_x^2 = \sum_i^n (\lvert e_{xi}\rvert - \text{RMSE}_x)^2/(n-1)$	0.1221
Estimated standard deviation of the population of errors	$S_x = \sqrt{\sum_i^n (\lvert e_{xi}\rvert - \text{RMSE}_x)^2/(n-1)}$	0.3494
Estimated standard deviation of the population of RMSEs	$S_{\text{RMSE}_x} = \sqrt{S_x^2/n}$	0.0638
Greenwalt and Schultz CMAS standard normal (*Z*) interval of the population of errors at 90% probability	$1.645 * S_x$	0.5748
Greenwalt and Schultz standard normal (*Z*) interval of the population of errors at 95% probability	$1.96 * S_x$	0.6849
NSSDA statistic	$1.96 * \text{RMSE}_x$	0.6572
Confidence interval on the estimate of RMSE_x at 95% probability	$\text{RMSE}_x \pm 1.96 * S_{\text{RMSE}}$	0.3353 ± 0.1250 which results in a range from 0.2102 to 0.4603

Definitions	*Y* Dimension Equations	*Y* Dimension Values
Estimated root-mean-square of the population of errors	$\text{RMSE}_y = \sqrt{\sum_i^n (e_{yi})^2/n}$	0.5205
Estimated absolute arithmetic mean of the population of errors	$\lvert\bar{e}_y\rvert = \sum_1^n \lvert e_{yi}\rvert/n$	0.3076
Estimated variance of the population of errors	$S_y^2 = \sum_1^n (\lvert e_{yi}\rvert - \text{RMSE}_y)^2/(n-1)$	0.2292
Estimated standard deviation of the population of errors	$S_y = \sqrt{\sum_1^n (\lvert e_{yi}\rvert - \text{RMSE}_y)^2/(n-1)}$	0.4788
Estimated standard deviation of the population of RMSEs	$S_{\text{RMSE}_y} = \sqrt{S_y^2/n}$	0.0874
Greenwalt and Schultz CMAS standard normal (*Z*) interval of the population of errors at 90% probability	$1.645 * S_y$	0.7876
Greenwalt and Schultz standard normal (*Z*) interval of the population of errors at 95% probabilty	$1.96 * S_y$	0.9384
NSSDA statistic	$1.96 * \text{RMSE}_y$	1.0201
Confidence interval on the estimate of RMSE_y at 95% probability	$\text{RMSE}_y \pm 1.96 * S_{\text{RMSE}}$	0.5202 ± .01713 which results in a range from 0.3491 to 0.6918

(*Continued*)

TABLE 3.4 (CONTINUED)
Horizontal Accuracy Example Equations and Statistics

Definitions	Circular Equations	Circular Values
Estimated root-mean-square of the populations of errors	$\text{RMSE}_h = \sqrt{\sum_i^n (e_{hi})^2/n}$	0.6191
Estimated absolute arithmetic mean of the population of errors	$\lvert\bar{e}_h\rvert = \sum_1^n \lvert e_{hi}\rvert/n$	0.4669
Estimated standard deviation of the population of errors	$S_h = (S_x + S_y)/2$	0.4141
Estimated standard deviation of the population of RMSEs	$S_{\text{RMSE}_h} = S_h/\sqrt{n}$	0.0756
Greenwalt and Schultz CMAS standard normal (Z) interval of the population of errors at 90% probability	$2.1460 * S_h$	0.8887
Greenwalt and Schultz standard normal (Z) interval of the population of errors at 95% probability	$2.4477 * S_h$	1.0136
Test for circularity	S_{min}/S_{max}	0.7298
$\text{NSSDA}_{\text{circular}}$ statistic	$1.7308 * \text{RMSE}_h$	1.0716
$\text{NSSDA}_{\text{elliptical}}$ statistic	$2.4477 * .5 * (\text{RMSE}_x + \text{RMSE}_y)$	1.0473
Confidence interval on the estimate of RMSE_h at 95% probability	$\text{RMSE}_h \pm 1.96 * S_{\text{RMSE}}$	0.6191 ± 0.1482 which results in a range from 0.4709 to 0.7673

accuracy assessment requires the collection of a separate and independent set of test sample points that were not used as control points in the registration process.

SUMMARY

A major tenet of most positional accuracy assessment standards is to report accuracy at a specified "confidence level" (FGDC, 1998; MPLMIC, 1999; NDEP, 2004). However, none of the existing standards provide equations for producing a confidence interval on the estimate of error. Table 3.5 provides a comparison of the commonly used positional accuracy standards. Each standard has both advantages and disadvantages:

1. NMAS is simple to implement, but it is based on map units instead of ground units, making it unusable for digital data. It also omits any guidance for estimating the range of errors at a given probability or for estimating a confidence interval around RMSE.
2. The Greenwalt and Schultz standard requires the assumption that the errors are normally distributed. The report states that the assumption of normality is "valid because positional error components generally follow a normal distribution pattern when sufficient data is available." However,

TABLE 3.5
Comparison of Commonly Used Positional Accuracy Standards to Each Other and to the Suggested New Standard

Positional Accuracy Standard	Uses a Maximum Distance of Error Allowed as the Standard	Provides Equations for Estimating Error Population Statistics	Requires the Assumption that the Errors Are Normally Distributed	Uses RMSE and 95% Percentile	Requires Stratification of the Landscape into Ground Cover Classes	Units
NMAS	Yes	No	Not required†	No, but uses 90th percentile	No	Map units
Greenwalt and Schultz	No	Yes	Yes	No	No	Unstated
ASPRS, 1989	Yes	Yes	Not required	No	No	Ground units
NSSDA	No	Yes	Yes	No	No	Ground units
FEMA, 2003	No	Yes	Yes	No	Yes	Ground units
ASPRS, 2004	No	Yes	Yes for RMSE, but not for 95th percentile	Yes	Yes	Ground units
NDEP, 2004	No	Yes	Yes for RMSE, but not for 95th percentile	Yes	Yes	Ground units
Combined standard: Z_i*S and a confidence interval on RMSE	No	Yes	Yes for calculating the probable range of errors, but no for calculating the confidence interval around the estimate of the mean error	No	Yes if desired	Ground units

† The assumption of normality is not required because probabilities of error distributions are not considered or calculated for the standard.

many practitioners are uncomfortable with this assumption and believe that positional errors are usually biased. Additionally, the Greenwalt and Schultz equations are often misinterpreted to calculate a level of confidence in the estimate of RMSE. However, a confidence level is a measure of the *reliability* of an estimation of a population parameter and is calculated

using the standard error, not the standard deviation. The expression used in Greenwalt and Schultz is

$$Z_i S$$

and the range of errors around either vertical or horizontal *RMSE* at a specified probability is

$$\text{RMSE} \pm Z_i S$$

where RMSE and S were defined earlier and Zi is the Z statistic for the specified probability. The Greenwalt and Schultz equation calculates the distance on either side of the RMSE beyond which errors will not occur at a specified probability, and does not indicate the confidence of the estimate of RMSE. Contrary to statements in the NSSDA, the equations in Greenwalt and Schultz do not, nor do they pretend to, provide a confidence level for the estimate of the RMSE.

3. As an improvement to NMAS, the ASPRS standards use ground units, but do not provide for any guidance on estimating either the range of errors at given probabilities or a confidence level on the estimate of RMSE.
4. NSSDA provides excellent guidance on positional accuracy sample design and collection methods. It also attempts to provide a means for estimating the range of errors at specified probabilities (and not, as it states, the "confidence level" of the RMSE estimate). However, the NSSDA equation is incorrect because it applies the RMSE variable in its equations where the estimate of the standard deviation (*S*) should be used instead.

An alternative clarifying standard would be to require calculation of both

- the interval of errors around RMSE that captures 95% of the map errors using Greenwalt and Schultz's equations and assumptions: $\text{RMSE} \pm Z_i S$, and
- a confidence interval around the estimate of RMSE at 95% probability $\text{RMSE} \pm Z_i S_{RMSE}$ for large sample sizes, and $\text{RMSE} \pm t_i S_{\text{RMSE}}$ for small sample sizes.

This standard has several advantages:

1. It relies on widely accepted statistical theory and equations to characterize geospatial positional error.
2. It corrects for the equation mistakes in NSSDA.
3. It clarifies the difference between estimating the range of errors at a certain probability versus calculating a confidence interval on the estimate of RMSE.
4. Use of the confidence interval does not require the assumption that the errors be normally distributed, because the population of possible RMSE values is normally distributed even if the population of the errors is not, and
5. The equations incorporate all of the critical concepts of NSSDA accuracy standards, including:

- reporting the estimated accuracy of a spatial data layer rather than specifying a standard to be met,
- use of estimated RMSE to estimate positional accuracy, and
- the ability to express a confidence level in the estimate of RMSE.

APPENDIX 3.1

Determining the Required Sample Size

If we have a prior estimate of the mean and standard deviation of our population of errors, we can determine how many samples we have to take to provide a specified confidence interval around our estimate of the mean error.

If d is the interval on either side of the mean that we want to estimate, then the confidence interval is

$$\bar{X} \pm d$$

and

$$d = tS_{\text{RMSE}}$$

Because $S_{\text{RMSE}} = S/\sqrt{n}$, we can solve for n because

$$n = (t^2 S^2)/d^2$$

For example, let us assume that we want our confidence interval on our estimate of the mean vertical error in Table 3.1 to be no more than ±20% of the mean at the 95% confidence level. Using the value of RMSE of 0.320 and the variance (S^2) of 0.059, we can calculate how many samples we would need to take as

$$n = \frac{(1.96)^2(0.059)}{(20\%(0.320))^2}$$

$$= \frac{3.842\,(0.059)}{.0041}$$

or about 55 samples.