Positional Accuracy Handbook

Using the National Standard for Spatial Data Accuracy to measure and report geographic data quality

OCTOBER 1999

The Governor's Council on Geographic Information was created in 1991 to provide leadership in the development, management and use of geographic information in Minnesota. With assistance from Minnesota Planning, the council provides policy advice to all levels of government and makes recommendations regarding investments, management practices, institutional arrangements, education, stewardship and standards.

The Council's GIS Standards Committee was created in 1993 to help GIS users learn about and use data standards that can help them be more productive. The committee's Internet home page is at http://www.lmic.state.mn.us/gc/committe/stand/index.htm. This handbook was designed, researched and written by the committee's Positional Accuracy Working Group: Christopher Cialek, chair, Land Management Information Center at Minnesota Planning; Don Elwood, City of Minneapolis; Ken Johnson, Minnesota Department of Transportation; Mark Kotz, assistant chair, Minnesota Pollution Control Agency; Jay Krafthefer, Washington County; Jim Maxwell, The Lawrence Group; Glenn Radde, Minnesota Department of Natural Resources; Mike Schadauer, Minnesota Department of Transportation; Ron Wencl, U.S. Geological Survey.

Minnesota Planning is charged with developing a long-range plan for the state, stimulating public participation in Minnesota's future and coordinating activities with state agencies, the Legislature and other units of government.

Upon request, the *Positional Accuracy Handbook* will be made available in an alternative format, such as Braille, large print or audio tape. For TTY, contact Minnesota Relay Service at 800-627-3529 and ask for Minnesota Planning.

Copies of the *Positional Accuracy Handbook* can be downloaded at: http://www.mnplan.state.mn.us/press/accurate.html. For additional information or printed copies of this handbook, contact the Land Management Information Center, 651-297-2488; e-mail gc@mnplan.state.mn.us. The council's Internet home page is at http://www.lmic.state.mn.us/gc/gc.htm. Copies of the National Standard for Spatial Data Accuracy can be downloaded at: http://www.fgdc.gov/standards/status/sub1_3.html.

October 1999

MINNESOTA PLANNING LAND MANAGEMENT INFORMATION CENTER

658 Cedar St.
St. Paul, MN 55155
651-296-1211
www.lmic.state.mn.us

Positional Accuracy Handbook

CASE STUDIES

APPENDIX

APPLYING THE NSSDA

This example demonstrates how the Positional Accuracy Handbook *helped the Minnesota Department of Transportation make a prudent business decision.*

Keeping track of the state's tens of thousands of road signs is no simple task. When speed limits change or a sign gets knocked down or simply gets old, the Minnesota Department of Transportation must install, update, repair or replace those signs. To efficiently manage this substantial resource, the department needs to accurately identify where signs are located and ultimately, to develop a GIS system for Facilities Management.

Traditional survey methods for collecting sign locations can be time consuming and costly. This is particularly true when dealing with large numbers of signs spread out over a sizeable area. With thousands of signs to survey, mainly situated near highway traffic, Mn/DOT looked to desktop surveying to provide a safe, quick and cost-effective way to collect sign location information. Desktop surveying is the process of calculating coordinate information from images on a computer. The images are collected using a van equipped with multiple cameras and geo-referenced with ground coordinates.

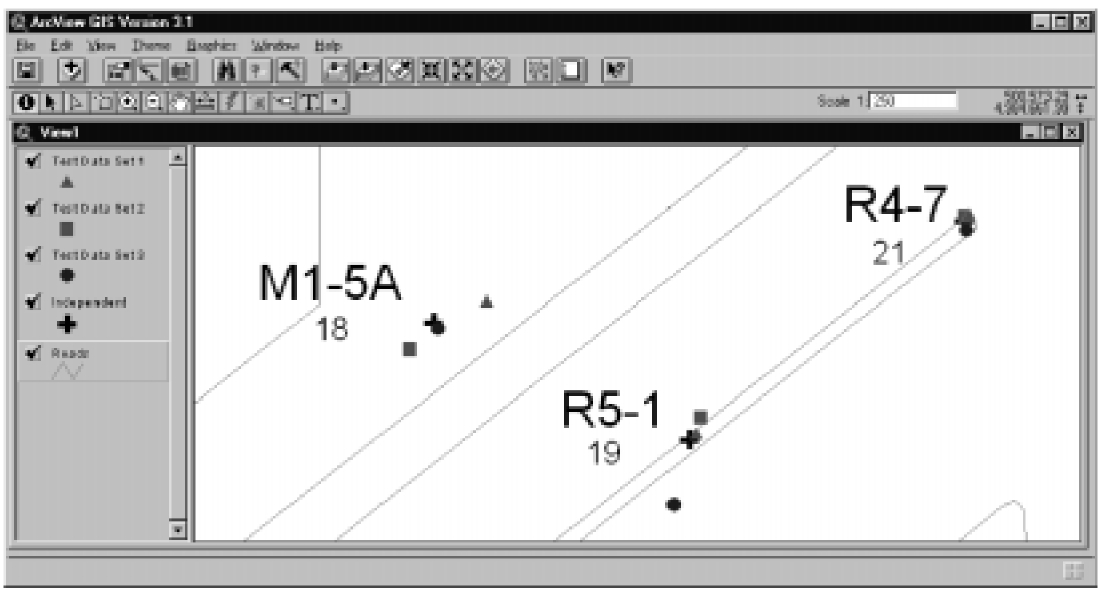
To evaluate this technology, Mn/DOT chose a short segment of State Trunk Highway 36 and collected x and y coordinates for all westbound signs with desktop surveying software packages from three vendors. A Mn/DOT survey crew was also sent out to collect the same signs with traditional survey equipment. The task of trying to figure out just how accurate the sign locations were for each desktop surveying package called for a standardized method; one with proven statistical merit.

Traditional methods of calculating accuracy are based on paper maps and would not work for this data. Mn/DOT turned to the draft *Positional Accuracy Handbook*, for a step-by-step approach and sound statistical methodology. The NSSDA recognizes the growing need for evaluating digital spatial data and provides a common language for reporting accuracy. Mn/DOT used the draft handbook to complete an accuracy evaluation and to critique this new data collection method.

After the results of Mn/DOT's *Mobile Mapping Accuracy Assessment* were released in May 1999, the department made the decision to use desktop surveying to collect locations for all signs in the Twin Cities metropolitan area, about 8,000 signs along 500 miles of roadway. Confidence in the accuracy and results of this new data collection method will save the state valuable time and resources.

Joella M. Givens
GIS Coordinator, Mn/DOT
joella.givens@dot.state.mn.us

Sign location comparisons for this section of Trunk Highway 36 in Ramsey County indicate errors ranging from 20 centimeters to 4 meters.



Positional Accuracy Handbook

Using the National Standard for Spatial Data Accuracy to measure and report geographic data quality

This handbook explains a national standard for data quality. The National Standard for Spatial Data Accuracy describes a way to measure and report positional accuracy of features found within a geographic data set. Approved in 1998, the NSSDA recognizes the growing need for digital spatial data and provides a common language for reporting accuracy.

The *Positional Accuracy Handbook* offers practical information on how to apply the standard to a variety of data used in geographic information systems. It is designed to help interpret the NSSDA more quickly, use the standard more confidently and relay information about the accuracy of data sets more clearly. It is also intended to help data users better understand the meaning of accuracy statistics reported in data sets. Case studies in this handbook demonstrate how the NSSDA can be applied to a wide range of data sets.

The risk of unknown accuracy

Consider this increasingly common spatial data processing dilemma. An important project requires that the locations of certain public facilities be plotted onto road maps so service providers may quickly and easily drive to each point. Global Positioning System receivers use state-of-the-art satellite technology to pinpoint the required locations. To provide context, these facility locations are then laid over a digital base map containing roads, lakes and rivers. A plot of the results reveals a disturbing problem: some facilities appear to be located in the middle of lakes (see figure 1).

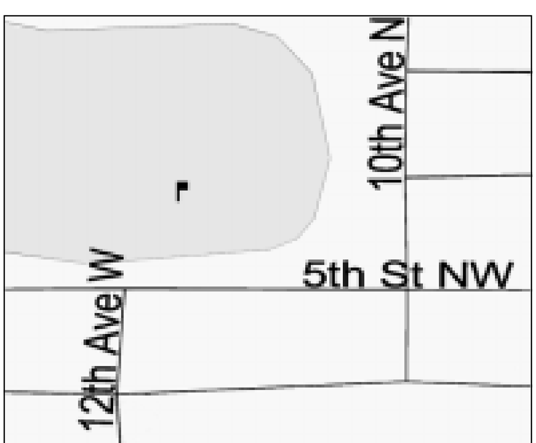


Figure 1. Variations in data accuracy are apparent when the two data sets are merged as shown here. The black flag marks the reported location of a building with a 5th Street address collected from a GPS receiver. Lake and road data come from U.S. Bureau of the Census.

Which data set is correct: the base map or the facility locations? No information about positional accuracy was provided for either data set, but intuition would lead us to believe that GPS points are much more accurate than information collected from a 1:100,000-scale paper map. Right?

In this case, *wrong*. The GPS receivers used for this study were only accurate to within 300 feet. The base map was assumed to be accurate to within 167 feet because it complied with the 1947 National Map Accuracy Standards. In reality, the base map may be almost twice as accurate as the information gathered from a state-of-the-art network of satellites. But, how would a project manager ever be able to know this simply by looking at a display on a computer screen?

Five components of data quality

This example illustrates an important principle of geographic information systems. The value of any geographic data set depends less on its cost, and more on its fitness for a particular purpose. A critical measure of that fitness is data quality. When used in GIS analysis, a data set's quality significantly affects confidence in the results. Unknown data quality leads to tentative decisions, increased liability and loss of productivity. Decisions based on data of known quality are made with greater confidence and are more easily explained and defended. Federal standards that assist in documenting and transferring data sets recognize five important components of data quality:

- **Positional accuracy.** How closely the coordinate descriptions of features compare to their actual location.
- **Attribute accuracy.** How thoroughly and correctly the features in the data set are described.

■ **Logical consistency.** The extent to which geometric problems and drafting inconsistencies exist within the data set.

■ **Completeness.** The decisions that determine what is contained in the data set.

■ **Lineage.** What sources are used to construct the data set and what steps are taken to process the data.

Considered together, these characteristics indicate the overall quality of a geographic database. The information contained in this handbook focuses on the first characteristic, **positional accuracy**.

Why a new standard is needed

How the positional accuracy of map features is best estimated has been debated since the early days of cartography. The question remains a significant concern today with the proliferating use of computers, geographic information systems and digital spatial data. Until recently, existing accuracy standards such as the National Map Accuracy Standards (described in the appendix) focused on testing paper maps, not digital data.

Today, use of digital GIS is replacing traditional paper maps in more and more applications. Digital geographic data sets are being generated by federal, state and local government agencies, utilities, businesses and even private citizens. Determining the positional accuracy of digital data is difficult using existing standards.

A variety of factors affect the positional accuracy of digital spatial data. Error can be introduced by: digitizing methods, source material, generalization, symbol interpretation, the specifications of aerial photography, aerotriangulation technique, ground control reliability, photogrammetric characteristics, scribing precision, resolution, processing algorithms and printing limitations. Individual errors derived from any one of these sources is often small; but collectively, they can significantly affect data accuracy, impacting how the data can be appropriately used.

The NSSDA helps to overcome this obstacle by providing a method for estimating positional accuracy of geographic data, in both digital and printed form.

The *National Standard for Spatial Data Accuracy* is one in a suite of standards dealing with the accuracy of geographic data sets and is one of the most recent standards to be issued by the Federal Geographic Data Committee. Minnesota was represented in the latter stages of the standard's development through the Governor's Council on Geographic Information and the state's Department of Transportation.

The role of the NSSDA in data documentation

The descriptive information that accompanies a data set is often referred to as *metadata*. Practically speaking, a well-documented data set is one that has a metadata record, including a standard report of positional accuracy based on NSSDA methods. Well documented and tested data sets provide an organization with a clear understanding of its investment in information resources. Trustworthy documentation also provides data users with an important tool when evaluating data from other sources. More information about metadata can be found in this handbook on page 7.

How the NSSDA works

There are seven steps in applying the NSSDA:

1. Determine if the test involves ***horizontal accuracy***, ***vertical accuracy*** or both.

2. Select a set of ***test points*** from the data set being evaluated.

3. Select an ***independent data set*** of higher accuracy that corresponds to the data set being tested.

4. Collect ***measurements*** from identical points from each of those two sources.

5. Calculate a positional accuracy ***statistic*** using either the horizontal or vertical accuracy statistic worksheet.

6. Prepare an accuracy statement in a standardized ***report*** form.

7. Include that report in a comprehensive description of the data set called ***metadata***.

FEATURES OF THE NSSDA

■ Identifies a well-defined **statistic** used to describe accuracy test results

■ Describes a **method** to test spatial data for positional accuracy

■ Provides a common language to **report** accuracy that makes it easier to evaluate the "fitness for use" of a database

Steps in detail

1. Determining which test to use. The first step in applying the NSSDA is to identify the spatial characteristics of the data set being tested. If planimetric accuracy — the *x,y* accuracy — of the data set is being evaluated, use the horizontal accuracy statistic worksheet (see figure 4). If elevation accuracy — *z* accuracy — is being evaluated, use the vertical accuracy worksheet (see figure 5).

2. Selecting test points. A data set's accuracy is tested by comparing the coordinates of several points within the data set to the coordinates of the same points from an independent data set of greater accuracy. Points used for this comparison must be well-defined. They must be easy to find and measure in both the data set being tested and in the independent data set.

For data derived from maps at a scale of 1:5,000 or smaller, points found at right-angle intersections of linear features work well. These could be right-angle intersections of roads, railroads, canals, ditches, trails, fences and pipelines. For data derived from maps at scales larger than 1:5,000 — plats or property maps, for example — features like utility access covers, intersections of sidewalks, curbs or gutters make suitable test points. For survey data sets, survey monuments or other well-marked survey points provide excellent test points.

Twenty or more test points are required to conduct a statistically significant accuracy evaluation regardless of the size of the data set or area of coverage. Twenty points make a computation at the 95 percent confidence level reasonable. The 95 percent confidence level means that when 20 points are tested, it is acceptable that one point may exceed the computed accuracy.

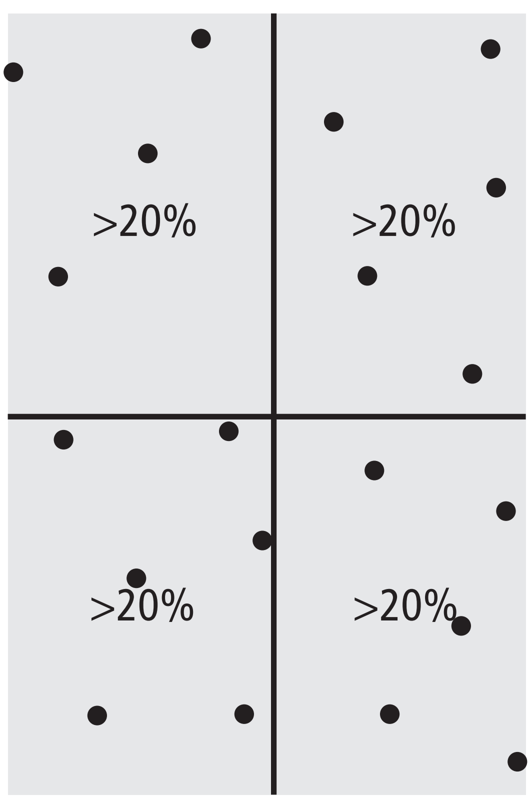
If fewer than 20 test points are available, another Federal Geographic Data Committee standard, the Spatial Data Transfer Standard, describes three alternatives for determining positional accuracy: 1) deductive estimate, 2) internal evidence and 3) comparison to source. For more information on this federal standard, point your browser to mcmcweb.er.usgs.gov/sdts/

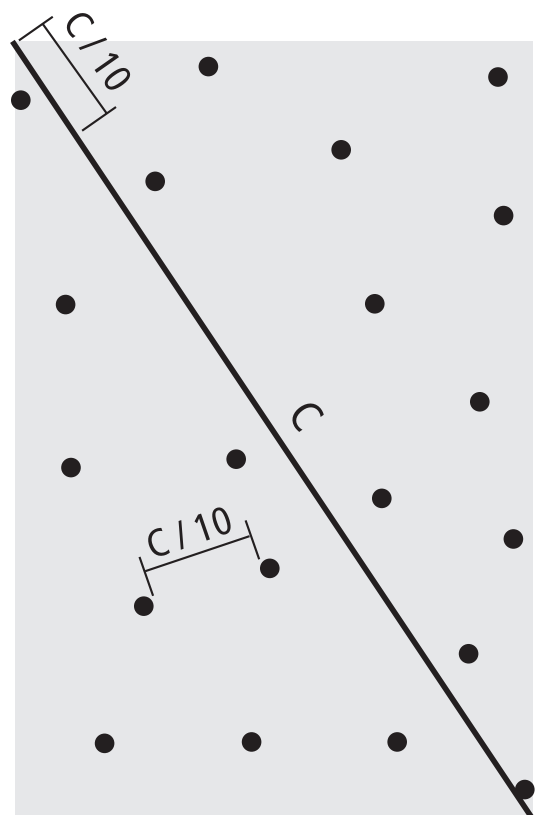
3. Selecting an independent data set. The independent data set must be acquired separately from the data set being tested. It should be of the highest accuracy available.

In general, the independent data set should be three times more accurate than the expected accuracy of the test data set. Unfortunately, this is not always possible or practical. If an independent data set that meets this criterion cannot be found, a data set of the highest accuracy feasible should be used. The accuracy of the independent data set should always be reported in the metadata.

Figure 2 (left). Ideal test point distribution.

Figure 3 (right). Ideal test point spacing.





The areal extent of the independent data set should approximate that of the original data set. When the tested data set covers a rectangular area and is believed to be uniformly accurate, an ideal distribution of test points allows for at least 20 percent to be located in each quadrant (see figure 2). Test points should be spaced at intervals of at least 10 percent of the diagonal distance across the rectangular data set; the test points shown in figure 3 comply with both these conditions.

It is not always possible to find test points that are evenly distributed. When an independent data set covers only a portion of a tested data set, it can still be used to test the accuracy of the overlapping area. The goal in selecting an independent data set is to try to achieve a balance between one that is more accurate than the data set being tested and one which covers the same region.

Independent data sets can come from a variety of sources. It is most convenient to use a data set that already exists, however, an entirely new data set may have to be created to serve as control for the data set being tested. In all cases, the independent and test data sets must have common points. Always report the specific characteristics of the independent data set, including its origin, in the metadata.

4. Recording measurement values. The next step is to collect test point coordinate values from both the test data set and the independent data set. When collecting these numbers, it is important to record them in an appropriate and similar numeric format. For example, if testing a digital database with an expected accuracy of about 10 meters, it would be overkill to record the coordinate values to the sixth decimal place; the nearest meter would be adequate. Use similar common sense when recording the computed accuracy statistic.

5. Calculating the accuracy statistic. Once the coordinate values for each test point from the test data set and the independent data set have been determined, the positional accuracy statistic can be computed using the appropriate accuracy statistic worksheet. Illustrations of filled out worksheets can be found in the handbook's case studies.

The NSSDA statistic is calculated by first filling out the information requested in the appropriate table and then computing three values:

- the **sum** of the set of squared differences between the test data set coordinate values and the independent data set coordinate values,
- the **average** of the sum by dividing the sum by the number of test points being evaluated, and
- the **root mean square error** statistic, which is simply the square root of the average.

The NSSDA statistic is determined by multiplying the RMSE by a value that represents the standard error of the mean at the 95 percent confidence level: 1.7308 when calculating horizontal accuracy, and 1.9600 when calculating vertical accuracy.

Accuracy statistic worksheets may be downloaded off the Internet from LMIC's positional accuracy web page ~~(www.mnplan.state.mn.us/press/accurate.html)~~ and clicking on *Download accuracy statistic worksheets.*

THE FEDERAL GEOGRAPHIC DATA COMMITTEE

The FGDC is a consortium of 16 federal agencies created in 1989 to better coordinate geographic data development across the nation. Additional stakeholders include: 28 states, the National Association of Counties and the National League of Cities, as well as other groups representing state and local government and the academic community. The Minnesota Governor's Council on Geographic Information represents the state on the FGDC. The committee has created a model for coordinating spatial data development and use. The National Spatial Data Infrastructure promotes efficient use of geographic information and GIS at all levels of government through three initiatives:

- ***Standards***. Developing common ways of organizing, describing and processing geographic data to ensure high quality and efficient sharing.
- ***Clearinghouse***. Providing Internet access to information about data resources available for sharing.
- ***Framework data.*** Defining the basic data layers needed for nearly all GIS analysis; better design of framework data layers promises easier data sharing.

For more information about the FGDC, visits its web site at www.fgdc.gov. The committee has set ambitious goals to identify areas where standards are needed and to develop those standards together with its partners. Details about committee-sponsored standards, both under development and completed, can be found on the Governor's Council web site at: www.lmic.state.mn.us/gc/standards.htm and at www.fgdc.gov/standards

Figure 4. Horizontal accuracy statistic worksheet.

A Point number	B Point description	C x (independent)	D x (test)	E diff in x	F (diff in x) ^{2}	G y (independent)	H y (test)	I diff in y	J (diff in y) ^{2}	K (diff in x) ^{2} + (diff in y) ^{2}
								sum	
								average	
								$RMSE_r$	
								NSSDA	

Column	Title	Content
A	Point number	Designator of test point
B	Point description	Description of test point
C	x (independent)	x coordinate of point from independent data set
D	x (test)	x coordinate of point from test data set
E	diff in x	x (independent) - x (test)
F	(diff in x) ^{2}	Squared difference in x = (x (independent) - x (test)) ^{2}
G	y (independent)	y coordinate of point from independent data set
H	y (test)	y coordinate of point from test data set
I	diff in y	y (independent) - y (test)
J	(diff in y) ^{2}	Squared difference in y = (y (independent) - y (test)) ^{2}
K	(diff in x) ^{2} + (diff in y) ^{2}	Squared difference in x plus squared difference in y = (error radius)^2
	sum	$\sum$ [(diff in x) ^{2} + (diff in y) ^{2}]
	average	sum / number of points
	$RMSE_r$	Root Mean Square Error (radial) = average$^{1/2}$
	NSSDA	National Standard for Spatial Data Accuracy statistic = 1.7308 * $RMSE_r$

Figure 5. Vertical accuracy statistic worksheet.

A Point number	B Point description	C z (independent)	D z (test)	E diff in z	F (diff in z) ^{2}
				sum	
				average	
				$RMSE_z$	
				NSSDA	

Column	Title	Contents
A	Point number	Designator of test point
B	Point description	Description of test point
C	z (independent)	z coordinate of point from independent data set
D	z (test)	z coordinate of point from test data set
E	diff in z	z (independent) - z (test)
F	(diff in z) ^{2}	Squared difference in z = (z (independent) - z (test)) ^{2}
	sum	$\sum$ (diff in z) ^{2}
	average	sum / number of points
	RMSE	Root Mean Square Error (vertical) = average$^{1/2}$
	NSSDA	National Standard for Spatial Data Accuracy statistic = 1.9600 * RMSE

6. Preparing an accuracy statement. Once the positional accuracy of a test data set has been determined, it is important to report that value in a consistent and meaningful way. To do this one of two reporting statements can be used:

Tested ______ (meters, feet) (horizontal, vertical) accuracy at 95% confidence level

Compiled to meet ______ (meters, feet) (horizontal, vertical) accuracy at 95% confidence level

A data set's accuracy is reported with the *tested* statement when its accuracy was determined by comparison with an independent data set of greater accuracy as described in steps 2 through 5. For example, if after comparing horizontal test data points against those of an independent data set, the NSSDA statistic is calculated to be 34.8 feet, the proper form for the positional accuracy report is:

Positional Accuracy: Tested 34.8 feet horizontal accuracy at 95% confidence level

This means that a user of this data set can be confident that the horizontal position of a well-defined feature will be within 34.8 feet of its true location, as best as its true location has been determined, 95 percent of the time.

When the method of compiling data has been thoroughly tested and that method produces a consistent accuracy statistic, the *compiled to meet* reporting statement can be used. Expanding on the same example, suppose the method of data collection consistently yields a positional accuracy statistic that was no worse — that is, no less accurate — than 34.8 feet for eight data sets tested. It would be appropriate to skip the testing process for data set nine, and assume that its accuracy is consistent with previously tested data. Report this condition using the following format:

Positional Accuracy: Compiled to meet 34.8 feet horizontal accuracy at 95% confidence level

To appropriately use the *compiled to meet* reporting statement, it is imperative that the data set compilation method consists of standard, well-documented, repeatable procedures. It is also important that several data sets be produced and tested. Finally, the NSSDA statistics computed in each test must be consistent. Once all these criteria are met, future data sets compiled by the same method do not have to be tested. The largest — or worst case — NSSDA statistic from all tests is always reported in the *compiled to meet* statement.

7. Including the accuracy report in metadata. The final step is to report the positional accuracy in a complete description of the data set. Often described as data about data, metadata lists the content, quality, condition, history and other characteristics of a data set.

The Minnesota Governor's Council on Geographic Information has established a formal method for documenting geographic data sets called the ***Minnesota Geographic Metadata Guidelines***. The guidelines are a compatible subset of the federal *Content Standards for Digital Geospatial Metadata* intended to simplify the process of creating metadata. A software program called *DataLogr* eases the task of collecting metadata that adheres to the Minnesota guidelines.

To report the positional accuracy of a data set, complete the appropriate field in section 2 of the metadata guidelines (see figures 6 and 7). The horizontal and vertical positional accuracy reports are free text fields and can be filled out the same way. Write the entire accuracy report statement followed by an explanation of how the accuracy value was determined and any useful characteristics of the independent data set.

Potential users of the data set might find this type of additional information useful:

- Specifically stating that the National Standard for Spatial Data Accuracy was used to test the data set.

POSITIONAL ACCURACY DESCRIPTIONS IN THE FEDERAL METADATA STANDARD

Section 2.4 of the full federal *Content Standards for Digital Geospatial Metadata* contains a number of positional accuracy related fields:

The NSSDA statistic should be placed in field 2.4.1.2.1 for horizontal accuracy and in field 2.4.2.2.1 for vertical accuracy. The text string "National Standard for Spatial Data Accuracy" should be entered in field 2.4.1.2.2 for horizontal accuracy and in field 2.4.2.2.2 for vertical accuracy.

Finally, an explanation of how the accuracy value was determined can be included in the horizontal positional accuracy report fields: 2.4.1.1 for horizontal and 2.4.2.1 for vertical.

■ Describing what is known about the variability of accuracy across the data set.

■ Pointing users to other sections of the metadata for more information.

Testing the NSSDA

Case studies in this handbook offer practical examples of how the NSSDA was applied to a selection of widely varied data sets. Each example strives to employ the procedures described here, and each offers a unique approach in establishing accuracy measurements due to the distinctive conditions of the test data set, the independent source and other local characteristics. Positional accuracy in these examples ranges from specific to general, from 0.2 meter to 4,800 meters, providing NSSDA users with ideas of how to adapt the standard to their own data sets.

To find out more about standards, metadata guidelines and *DataLogr*, go to www.lmic.state.mn.us and look under Spatial Data Standards, or contact LMIC by e-mail at clearinghouse@mnplan.state.mn.us or call Christopher Cialek at 651-297-2488.

References

American National Standards Institute, 1998. Information Technology — Spatial Data Transfer Standard (SDTS) *ANSI-NCITS 320. New York, mcmcweb.er.usgs.gov/sdts/*

American Society for Photogrammetry and Remote Sensing, Specifications and Standards Committee, 1990. "ASPRS Accuracy Standards for Large-Scale Maps," Photogrammetric Engineering & Remote Sensing. *Washington, D.C., v56, no7, pp1068-1070.*

Federal Geographic Data Committee, 1998. Geospatial Positioning Accuracy Standards; Part 3: National Standard for Spatial Data Accuracy, *FGDC-STD-007.3. Federal Geographic Data Committee. Washington, D.C. www.fgdc.gov/standards/status/sub1_3.html*

Federal Geographic Data Committee, 1998. Content Standards for Digital Geospatial Metadata (version 2.0), *FGDC-STD-001. Federal Geographic Data Committee. Washington, D.C. www.fgdc.gov/standards/status/csdgmovr.html*

Greenwalt, C.R. and Shultz, M.E. Principles of error theory and cartographic applications; Technical Report No. 96, *1962. U.S. Air Force Aeronautical Chart and Information Center. St. Louis, Missouri.*

Minnesota Governor's Council on Geographic Information. Minnesota Geographic Metadata Guidelines (version 1.2)*, 1998. St. Paul, Minnesota. www.lmic.state.mn.us/gc/stds/metadata.htm*

U.S. Bureau of the Budget. U.S. National Map Accuracy Standards,*1947. Washington, D.C.*

Figure 6. Formal NSSDA accuracy statements reported in section 2 of the *Minnesota Geographic Metadata Guidelines.*

Horizontal positional accuracy	Tested 0.181 meters horizontal accuracy at 95% confidence level.
Vertical positional accuracy	Tested 0.134 meters vertical accuracy at 95% confidence level.

Figure 7. An example of a detailed positional accuracy statement as reported in metadata.

Horizontal positional accuracy	Digitized features outside areas of high vertical relief: tested 23 feet horizontal accuracy at the 95% confidence level using the NSSDA. Digitized features within areas of high vertical relief (such as major river valleys): tested 120 feet horizontal accuracy by other testing procedures. For a complete report of the testing procedures used, contact Washington County Surveyor's Office as noted in Section 6, Distribution Information. All other features are generated by coordinate geometry and are based on a framework of accurately located PLSS corner positions used with public information of record. Computed positions of parcel boundaries are not based on individual field surveys. Although tests of randomly selected points for comparison may show high accuracy between field and parcel map content, variations between boundary monumentation and legal descriptions of record can and do exist. Caution is necessary when using land boundary data shown. Contact the Washington County Surveyor's Office for more information.
Vertical positional accuracy	Not applicable

Case Study A

Minnesota Department of Transportation

Applying the NSSDA to large-scale data sets

PROJECT TEAM

Ken Johnson
Geodetic services engineer

Mike Lalla
Photogrammetry mapping supervisor

Mike Schadauer
Land information systems engineer

The project

This project evaluates the accuracy of topographic and digital terrain model data sets created using photogrammetric techniques. The Minnesota Department of Transportation's photogrammetric unit produces these data sets, which are used within the agency to plan and design roadways and roadway improvements.

The horizontal accuracy of the topographic data set was tested. Although the elevation contours of the topographic data set do record vertical data, a different data set, a Digital Terrain Model, was used to test vertical accuracy, as DTMs tend to be more accurate. DTMs are used to compute complex solutions dealing with design issues such as material quantities and hydraulics. To rely on these solutions, understanding the accuracy of the DTM is crucial.

The tested data set

The two data sets consist of digital elevation contours and a digital terrain model created from 457 meter altitude, 1:3000 scale aerial photography. They cover a corridor on Interstate Highway 94 from Earl Street to the junction of interstates 494 and 694, east of St. Paul. The mapping width varies from 250 meters to 862 meters and averages 475 meters. The horizontal accuracy, reflected in the digital topographic map, and the vertical accuracy, reflected in the digital terrain model, both were tested.

The independent data set

Mn/DOT's photogrammetric unit has historically assessed digital terrain models for **vertical accuracy** using National Map Accuracy Standards. The traditional method of evaluating vertical accuracy was to perform a test on every fourth model, or stereo pair, throughout the corridor. In each tested model, the x, y and z field coordinates of 20 random points were collected using Geotracer Dual Frequency GPS receivers with an expected accuracy of 10-15 mm (rms). The z coordinate from the field was compared to the z coordinate from the corresponding x and y coordinates in the digital terrain model to determine if they met the standard: 90 percent of the points fall within one half contour.

This project has 13 test models. With about 20 points in every test model there are 296 control points available for the entire corridor. Even though this is far more than the minimum suggested by the National Standard for Spatial Data Accuracy, the points were already measured so all were used.

The selection of control points for vertical accuracy testing was very simple. The survey field crew selected about 20 random points in every fourth model. These did not have to be well-defined points in the horizontal dimension because they were only intended to evaluate the vertical accuracy of the digital terrain model. As long as the control points were within the extent of the digital terrain model, they served to help evaluate the digital terrain model's vertical accuracy.

The digital topographic map made from the same aerial photography was not originally assessed for **horizontal accuracy**; however, it complied with National Map Accuracy Standards, which vary based on the scale of the map. It was assumed that, if horizontal problems existed, they would be uncovered and addressed when field crews conducted additional surveys to supplement the mapping.

This meant that there were no preconceived methods for assessing the horizontal accuracy. The method chosen for this example was to collect the coordinates of 40 well-defined points throughout the corridor. Geotracer Dual Frequency GPS receivers were used. Data was collected using a fast static method with an expected accuracy of 10-15 mm (rms).

In selecting the 40 control points used to assess the vertical accuracy, the project team chose points that were well defined both on the topographic map and in the field, and were fairly evenly distributed throughout the corridor. Examples of these include manholes, catch basins and right-angle intersections of objects such as sidewalks. Forty points were chosen rather than the minimum of 20 because they were fairly easy to collect and because of the long narrow shape of the corridor. Having the extra points opened the possibility of comparing a test of the 20 easternmost control points with a test of the 20 westernmost control points.

The worksheet

The completed worksheets for the vertical and horizontal accuracy testing are shown in tables A.1 and A.2, respectively. The columns listed as independent are the GPS collected points. The columns listed as test are the photogrammetrically derived points taken off the DTM for the vertical test and the topographic map for the horizontal test.

Table A.1. Vertical accuracy statistic worksheet.

Point number	z (test) Photo elev	z (independent) Field elev	diff in z Photo field	(diff in z)^2
100	293.755	293.79	-0.035	0.001202
101	293.671	293.71	-0.039	0.001515
102	293.87	293.9	-0.03	0.000913
103	293.815	293.85	-0.035	0.001241
104	294.609	294.62	-0.011	0.000113
105	295.238	295.3	-0.062	0.003834
106	295.54	295.56	-0.02	0.000394
107	295.28	295.3	-0.02	0.000385
108	294.933	295	-0.067	0.004465
109	294.431	294.46	-0.029	0.000847
110	293.994	294.02	-0.026	0.000664
111	293.736	293.77	-0.034	0.001131
112	293.537	293.58	-0.043	0.001886
113	293.478	293.55	-0.072	0.005164
114	293.671	293.7	-0.029	0.000858
115	293.949	293.97	-0.021	0.000425
116	294.427	294.49	-0.063	0.00402
117	294.837	294.88	-0.043	0.001881
118	295.19	295.28	-0.09	0.008062
119	295.318	295.31	0.008	0.000057
581	259.435	259.42	0.015	0.000213
582	258.766	258.74	0.026	0.000682
583	258.603	258.61	-0.007	0.000046
584	258.71	258.72	-0.01	0.000105
585	259.407	259.39	0.017	0.000275
586	259.285	259.28	0.005	0.000022
587	259.41	259.43	-0.02	0.000405
588	260.017	260.02	-0.003	0.000008
589	260.596	260.67	-0.074	0.005473
590	261.801	261.84	-0.039	0.001513
592	263.428	263.42	0.008	0.00007
593	256.949	256.93	0.019	0.000352
594	256.853	256.82	0.033	0.001105
595	256.766	256.72	0.046	0.002095
596	256.411	256.39	0.021	0.000441
597	256.12	256.14	-0.02	0.000414
598	258.249	258.3	-0.051	0.002645
599	258.395	258.46	-0.065	0.004181
600	258.414	258.46	-0.046	0.002159
			sum	1.375
			average	0.005
			RMSEz	0.068
			NSSDA	**0.134**

The positional accuracy statistic

The vertical root mean square error is shown as a linear error. In table A.1, the vertical RMSE is 0.068 m.

The horizontal RMSE deals with two dimensions giving x and y coordinates. Using the equation of a circle:

$$x^2 + y^2 = r^2$$

and modifying it slightly into

$$(x_{independent} - x_{test})^2 + (y_{independent} - y_{test})^2 = r_{error}^2$$

the error radius is found for each coordinate. The horizontal RMSE is calculated by adding up the radius errors, averaging them and taking the square root. This gives a circular error defined by the radius. The horizontal RMSE in table A.2 is a circle defined by a radius of 0.105 m.

Table A.2. Horizontal accuracy statistic worksheet.

Point number	Point description	x (independent)	x (test)	diff in x	(diff in x)^2	y (independent)	y (test)	diff in y	(diff in y)^2	(diff in x)^2 + (diff in y)^2
1	TP1A	178247.28	178247.37	-0.089	0.007921	48326.075	48326.135	-0.06	0.0036	0.011521
2	TP2	178249.23	178249.17	0.055	0.003025	48287.228	48287.171	0.057	0.003249	0.006274
3	TP3A	178456.79	178456.73	0.06	0.0036	48337.408	48337.283	0.125	0.015625	0.019225
4	TP4	178715.82	178715.88	-0.054	0.002916	48542.511	48542.543	-0.032	0.001024	0.00394
5	TP5	179047.54	179047.65	-0.104	0.010816	48657.388	48657.44	-0.052	0.002704	0.01352
6	TP6	179227.78	179227.8	-0.016	0.000256	48336.177	48336.147	0.03	0.0009	0.001156
7	TP7	179238.56	179238.69	-0.132	0.017424	48671.457	48671.48	-0.023	0.000529	0.017953
9	TP9	180257.36	180257.39	-0.033	0.001089	48337.972	48337.97	0.002	4E-06	0.001093
10	SWK	180426.36	180426.36	0.005	2.5E-05	48445.001	48444.917	0.084	0.007056	0.007081
11	DI	180568.35	180568.48	-0.128	0.016384	48523.693	48523.696	-0.003	9E-06	0.016393
12	TP12A	180680.73	180680.78	-0.053	0.002809	48275.075	48274.978	0.097	0.009409	0.012218
13	SW	180676.31	180676.38	-0.076	0.005776	48413.085	48413.154	-0.069	0.004761	0.010537
14	TP14	180654.46	180654.47	-0.01	0.0001	47955.055	47954.992	0.063	0.003969	0.004069
15	MH	180843.48	180843.56	-0.083	0.006889	48505.391	48505.548	-0.157	0.024649	0.031538
17	TP17	181338.97	181339.11	-0.141	0.019881	48313.103	48313.244	-0.141	0.019881	0.039762
18	TP18	181283.2	181283.25	-0.051	0.002601	48174.063	48174.057	0.006	3.6E-05	0.002637
19	TP19	181075.07	181075.09	-0.018	0.000324	48171.737	48171.637	0.1	0.01	0.010324
20	TP20A	181495.79	181495.85	-0.057	0.003249	48043.414	48043.497	-0.083	0.006889	0.010138
21	TP21	181679.58	181679.59	-0.009	8.1E-05	48242.779	48242.744	0.035	0.001225	0.001306
22	TP22	181673.86	181673.82	0.044	0.001936	48579.533	48579.693	-0.16	0.0256	0.027536
24	TP24	181937.26	181937.3	-0.045	0.002025	48136.264	48136.256	0.008	6.4E-05	0.002089
26	TP26	182085.95	182085.96	-0.004	1.6E-05	48127.717	48127.778	-0.061	0.003721	0.003737
27	TP27	182243.61	182243.57	0.041	0.001681	48032.915	48032.879	0.036	0.001296	0.002977
28	TP28	182289.49	182289.56	-0.065	0.004225	48729.272	48729.211	0.061	0.003721	0.007946
29	TP29	182259.51	182259.63	-0.122	0.014884	48630.614	48630.707	-0.093	0.008649	0.023533
30	TP30	182277.52	182277.57	-0.053	0.002809	48410.278	48410.398	-0.12	0.0144	0.017209
32	TP32	182590.79	182590.88	-0.095	0.009025	48437.482	48437.633	-0.151	0.022801	0.031826
33	TP33	182494.13	182494.22	-0.099	0.009801	48422.78	48422.862	-0.082	0.006724	0.016525
34	TP34	182410.21	182410.24	-0.027	0.000729	48672.544	48672.564	-0.02	0.0004	0.001129
35	TP35	182740.18	182740.15	0.027	0.000729	48307.436	48307.447	-0.011	0.000121	0.00085
36	TP36	182771.78	182771.77	0.015	0.000225	47967.3	47967.292	0.008	6.4E-05	0.000289
37	TP37	183067.28	183067.27	0.014	0.000196	48044.513	48044.539	-0.026	0.000676	0.000872
38	TP38	183242.23	183242.18	0.048	0.002304	47952.797	47952.798	-0.001	1E-06	0.002305
39	TP39	183458.2	183458.24	-0.035	0.001225	47885.194	47885.162	0.032	0.001024	0.002249
40	TP40	183778.2	183778.26	-0.059	0.003481	48230.799	48230.753	0.046	0.002116	0.005597
41	TP41	183886.38	183886.4	-0.019	0.000361	47924.349	47924.264	0.085	0.007225	0.007586
42	TP42	184394.5	184394.54	-0.031	0.000961	48083.648	48083.657	-0.009	8.1E-05	0.001042
43	TP43A	184644.38	184644.44	-0.059	0.003481	48068.904	48068.751	0.153	0.023409	0.02689
44	TP44	184804.19	184804.35	-0.16	0.0256	48192.963	48192.891	0.072	0.005184	0.030784
45	TP45	185120.62	185120.67	-0.054	0.002916	48201.523	48201.505	0.018	0.000324	0.00324
									sum	0.436896
									average	0.0109224
									RMSE	0.10451029
									NSSDA	0.1808864

The NSSDA requires a 95 percent confidence level. To attain this, the vertical RMSE is multiplied by 1.96 and the horizontal RMSE is multiplied by 1.7308, resulting in horizontal and vertical accuracies of 0.181 and 0.134 meters respectively.

The accuracy statement and metadata

Figure A.1 contains formal NSSDA reports for both the horizontal and vertical positional accuracy measured for this project.

Observations and comments

A couple of concerns were raised in applying the NSSDA to this data set. First, the field test shots need to be points that have been placed on the map. For this project, the field crew substituted a few points that had not been originally placed on the topographic maps, so these were not considered. The second concern dealt with map symbol placement, origin and scale. For example, with a map symbol such as a catch basin, the origin is the lower left corner. The field crew may have collected the control point using the center of the catch basin, and even if they used the correct corner, the scaled size may be different. This type of systematic error could have a major impact on the accuracy statement of this project.

Figure A.1. Positional accuracy statements as reported in metadata.

Horizontal positional accuracy	Using the National Standard for Spatial Data Accuracy, the data set tested 0.181 meters horizontal accuracy at 95% confidence level.
Vertical positional accuracy	Using the National Standard for Spatial Data Accuracy, the data set tested 0.134 meters vertical accuracy at 95% confidence level.

Case Study B

City of Minneapolis

Applying the NSSDA to contract service work

PROJECT TEAM

Don Elwood
Engineer, engineering design

Tara Mugane
Engineer, engineering design

Lisa Zick
Engineering graphics analyst, engineering design

The project

The city of Minneapolis uses its planimetric database for a variety of engineering and planning purposes. In this project, it provided a photo control to produce a digital orthophoto database for the city.

Presently, about two-thirds of Minneapolis is covered with high-resolution color digital orthophotographs. The primary use for this database is to identify changes over time and transfer them to the planimetric database. Updated planimetry is digitized from digital orthophotos into the planimetric database. A precise match between the orthophoto products and the planimetric database is critical as design crews use the planimetric database and orthophotos for street design plans. For this reason, it was a worthwhile investment to develop a positional accuracy estimate for digital orthophotos using the NSSDA.

The tested data set

Orthophotographs are being generated by contract from aerial photography of a flight height of approximately 3,000 feet. The photos are being scanned at a resolution of 25-30 microns creating an 80-megabyte file per quarter section area. Photo distortions are removed through a procedure that applies elevation and horizontal control to the scanned photos. This process results in a color orthophoto with a ground pixel resolution of one-half foot. Many of the images have a quality sufficient to identify cracks in pavement surface.

As with traditional aerial photography projects, photo control on painted targets placed at regular intervals around the city was required. In addition, the vendor was required to use the city's planimetric basemap for control and to supply the city with coordinate values for those targets.

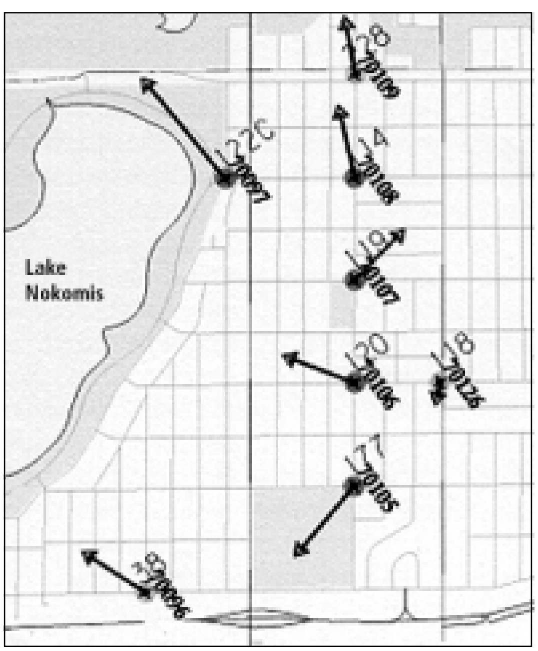


Figure B.1. Portion of a Minneapolis control quality map identifying the magnitude and direction of differences between city monuments and corresponding vendor-supplied data.

The independent data set

Survey monument locations from the planimetric database were used as the source of independent data. City survey crews painted targets on the existing monuments within the flight area, attempting to select targets on the corner of each quarter section. The city compared the coordinate results from the orthophoto vendor with the control data and a plot was made to show areas of "control quality." Figure B.1 illustrates the use of points to display the range between the independent data and vendors values while the arrows show the direction the vendor's coordinates are in comparison to the independent data. Errors are investigated and additional control is provided where needed before the vendor starts orthophoto production.

The worksheet

Table B.1 contains a partial list of coordinate values for the independent (monument) coordinates, test (vendor) coordinates, the differences between each of those values and the squared difference for 126 monuments. From these values a sum, average, root mean square error and National Standard for Spatial Data Accuracy statistic are calculated.

The positional accuracy statistic

The horizontal root mean square value is the sum error squared in both the x and y directions divided by the number of control points. The RMSE calculated value was 0.577 feet. This root mean square value multiplied by 1.7308 gives a 0.999 feet horizontal accuracy at the 95 percent confidence level.

The accuracy statement and metadata

Figure B.2 contains the formal NSSDA report for horizontal positional accuracy measured for this project.

Observations and comments

Once test data was received from the vendor, the results were mapped. The project team noticed large random errors where good control was expected. It investigated those points and found monuments incorrectly painted or monuments not properly labeled. In areas of significant elevation change, errors were larger. To compensate, some points were thrown out because adequate control was not available. In areas of large elevation change, additional control data was given to the vendor who then was able to return updated coordinate data with improved results.

Table B.1. Horizontal positional accuracy worksheet.

Point number	x (independent)	x (test)	diff in x	(diff in x)^2	y (independent)	y (test)	diff in y	(diff in y)^2	(diff in x)^2 + (diff in y)^2
3234	542850.895	542850.872	0.023	0.000529	152223.812	152223.840	-0.028	0.000784	0.001313
3062	522260.248	522260.211	0.037	0.001369	138937.691	138937.700	-0.009	8.1E-05	0.00145
118	541542.816	541542.781	0.035	0.001225	141704.350	141704.309	0.041	0.001681	0.002906
230A	540484.021	540484.057	-0.036	0.001296	148882.295	148882.347	-0.052	0.002704	0.004
441	535191.599	535191.550	0.049	0.002401	161295.030	161295.075	-0.045	0.002025	0.004426
811	539143.822	539143.898	-0.076	0.005776	173109.161	173109.161	0	0	0.005776
2215A	535233.433	535233.408	0.025	0.000625	148235.180	148235.075	0.105	0.011025	0.01165
334A	540460.246	540460.317	-0.071	0.005041	156140.475	156140.383	0.092	0.008464	0.013505
3325	542852.905	542852.843	0.062	0.003844	154853.051	154852.845	0.206	0.042436	0.04628
310	535172.307	535171.633	0.674	0.454276	157367.845	157367.914	-0.069	0.004761	0.459037
821	545478.707	545478.821	-0.114	0.012996	173183.409	173182.379	1.03	1.0609	1.073896
125	532263.602	532262.676	0.926	0.857476	141665.163	141665.682	-0.519	0.269361	1.126837
126	531934.210	531933.463	0.747	0.558009	141661.771	141662.617	-0.846	0.715716	1.273725
2117	542896.486	542897.629	-1.143	1.306449	141706.605	141706.675	-0.07	0.0049	1.311349
3035	545524.703	545523.756	0.947	0.896809	139073.949	139073.149	0.8	0.64	1.536809
431	527338.530	527339.579	-1.049	1.100401	162558.097	162558.762	-0.665	0.442225	1.542626
544	530126.997	530127.929	-0.932	0.868624	168401.063	168401.941	-0.878	0.770884	1.639508
2055	519318.325	519319.943	-1.618	2.617924	136413.556	136413.553	0.003	9E-06	2.617933
								sum	41.952161
								average	0.332953659
								RMSE	0.577021368
								NSSDA	**0.998708583**

Figure B.2. Positional accuracy statements as reported in metadata.

Horizontal positional accuracy	Using the National Standard for Spatial Data Accuracy, the data set tested 1 foot horizontal accuracy at 95% confidence level.
Vertical positional accuracy	Not applicable.

Case Study C

Washington County Surveyor's Office

Measuring horizontal positional accuracy in a county parcel database

PROJECT TEAM

Jay Krafthefer
GIS manager

Marc Senjem
Survey project coordinator

Mark Nieman
Survey/GIS specialist

GPS field team

The project

The objective was to analyze and determine the positional accuracy of Washington County's recently completed parcel base data set. Covering 425 square miles with approximately 82,000 parcels of land ownership, the data set has features typically found in half-section maps, including plat boundaries, lot lines, right-of-way lines, road centerlines, easements, lakes, rivers, ponds and other requirements of county land record management.

An estimation of the parcel data's positional accuracy has existed for some time, established through limits and standards used to develop the data set. The county wanted the ability to report positional accuracy in a standardized format for the following reasons:

Figure C.1. Map of Washington County, Minnesota, showing only digitized features of the parcel database.

■ To create complete and high quality metadata for the parcel base data set.

■ To improve communication of data accuracy in sales and exchange of digital data with customers.

■ To aid in decision-making when the parcel base is merged or combined with other data collections, a typical use for this data set.

Convenient and standardized positional accuracy information can help the formation of metadata for hybrid data sets and applications supported by the parcel base.

The tested data set

Features in Washington County's parcel map are derived from a variety of source documents. Most are fairly complete with angle and distance information describing their design, including subdivision plats, registered land surveys, condominium plats, certificate of surveys, right-of-way plats and auditor's metes and bounds parcel descriptions. Sources range in date from the 1850s through the present.

The locations of parcel boundaries were derived using coordinate geometry analysis. This work was primarily referenced to the Public Land Survey System. Field verifications were not performed on discrepancies found in the analysis of documents. Undefined features such as undocumented road locations were located by a digitizing process using partially rectified aerial photographs at a scale of 1-inch equals 200 feet. This group of digitized features is comprised primarily of hydrographic features with some roads. It is estimated that less than 5 percent of all features in the database were digitized (see figure C.1).

The potential exists for a significant difference in accuracy between digitized features and those created by a coordinate geometry process. Therefore, the accuracy of each group was computed and reported separately and described as follows: *Test Data/Digitized* and *Independent Data/Digitized* are used to reference the digitized elements.

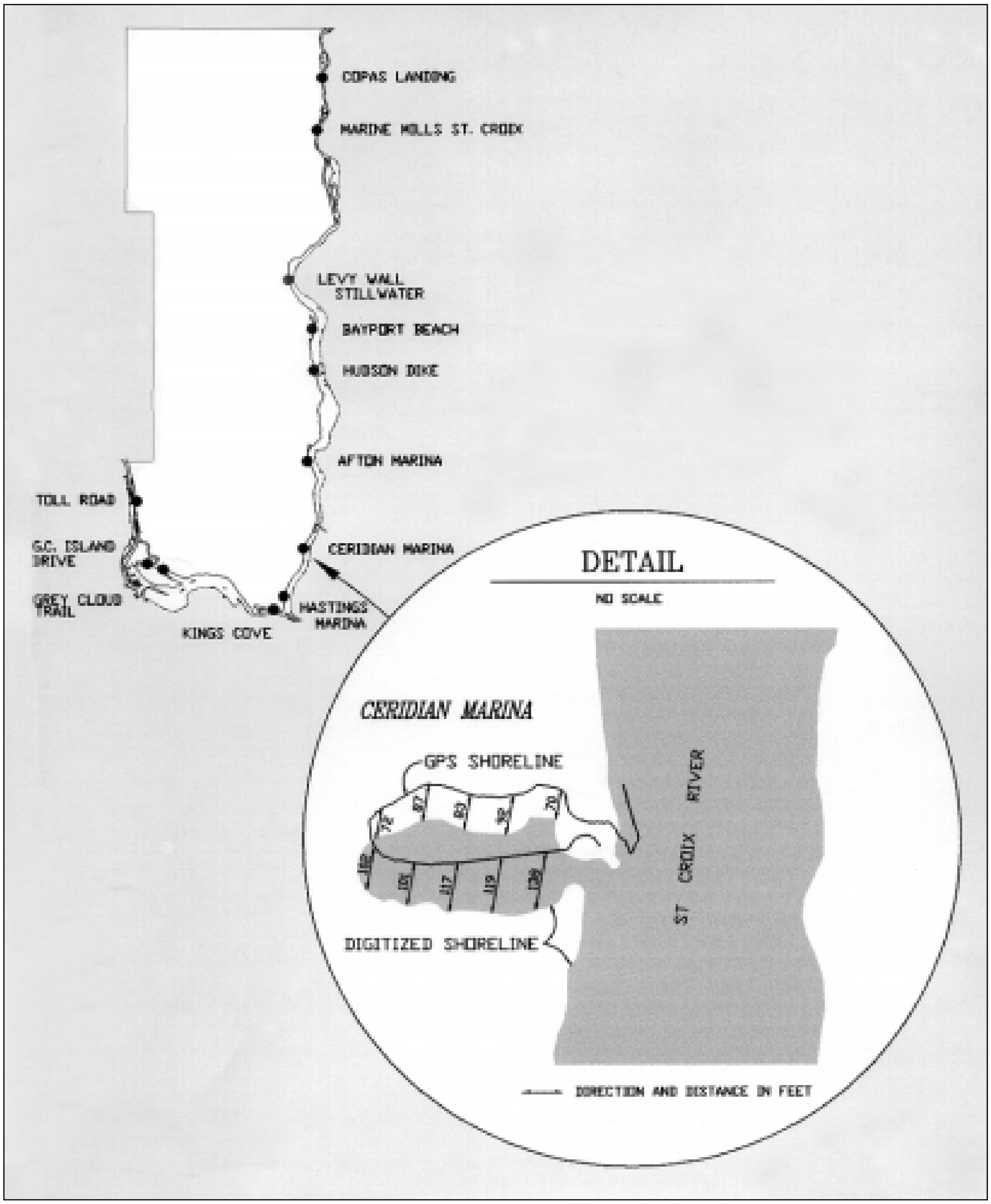
Test Data/COGO and *Independent Data/COGO* are used for referencing the nondigitized feature group.

The independent data set

The county was not aware of any appropriate, existing independent data set, and decided to create an independent set based on corresponding points identified in the test data. Readily available GPS equipment capable of producing sub-meter results prompted use of field measured locations for control. Such a level of accuracy would meet the NSSDA stipulation of using an independent source of data of the highest accuracy feasible and practical to evaluate the accuracy of the test data set.

Because of the availability of real time differential GPS equipment and its ease of use, more than the minimum of 20 test points were collected. To identify potential test points, a plot of the entire county parcel database was generated, with only digitized features shown. This was possible due to the unique design of the database where features are coded based on their quality. The selection set consisted of primarily water and road features. The overall number of easily identifiable right angle

Figure C.2. Vicinity map of 12 linear feature locations, Washington County, Minnesota.



intersections was small compared to what exists in the COGO data set. All of the following types of points were designated as potential control points:

- water line at bridge abutment
- intersection of road, creek and ditch (culvert)
- road intersection
- intersection of road and power line
- road and trail intersection
- stream outlet at lake or river
- ditch intersection
- stream and railroad intersection
- road and railroad intersection

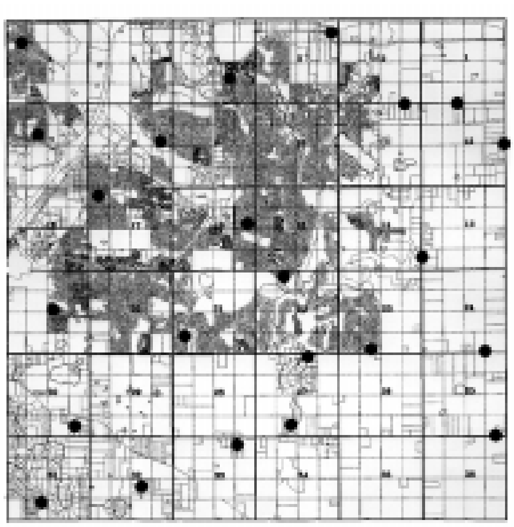
Of these, approximately 175 were identified as possible candidates on the map, giving the field crew freedom in making their selection. Significance was not given to precise spacing of points due to the high number to be collected.

Areas of high vertical relief within the county, such as the bluff areas along the Mississippi and St. Croix rivers, have less positional accuracy in the data set. This is due to a shortage of aerial targeting to support the partially rectified aerial photographs in these areas. Unfortunately, these areas contained few right angle points to collect from the digitized group. Even though points in the river valleys were given a higher priority for collection, few control points were actually collected there. Without considering accuracy in the river areas, data users could be seriously misguided.

To overcome this dilemma, a plan was developed to use the GPS equipment's ability to trace linear features by collecting points in continuous mode. In this mode a point was collected every second as a person walked the boundary line of a physical feature. These lines could then become a form of linear control and compared with corresponding linear features that had been digitized. Staff devised a method of comparing points at even intervals along the selected features and reporting a difference in their location as compared to their position on the map. Twelve evenly spaced areas along the river shoreline were identified in the digitized data set and could be easily found on the ground. Preference was given to using features that would not vary due to seasonal changes and would match the spring season conditions that existed at the time of the aerial photographs. (See figure C.2.)

As with the digitized data, a suitable independent data set was not known to exist for the COGO data. Again, GPS equipment owned by Washington County (capable of survey quality accuracy) was available for use in this study. This equipment was capable of generating points to an accuracy of two centimeters using a post-processed differential mode and constitutes the highest accuracy data set practical.

Like the digitized data, closer examination of the COGO data revealed a mix of data accuracy. First, the COGO data contained PLSS section lines that were mapped based on Public Land Survey Corners. The PLSS corners were located in a measurement process, which was based on GPS control. Secondly, parcel lines were mapped by interpreting property descriptions of record and applying coordinate geometry analysis to define their position. The Public Land Survey System was the supportive framework for these land descriptions. The situation did not permit actual boundary surveys or field verification of individual parcel



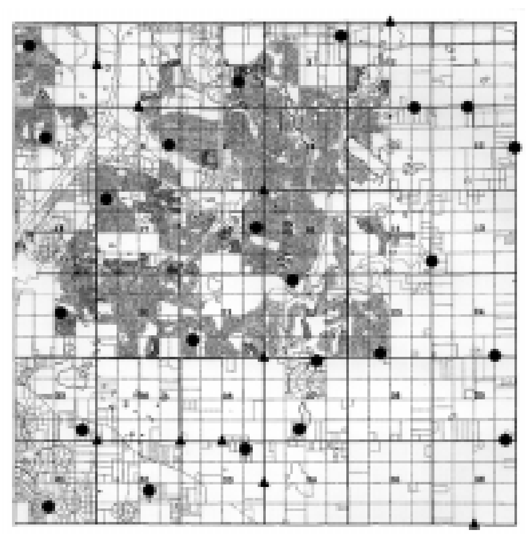


Figure C.3 (left). COGO test points, Washington County, Minnesota.

Figure C.4 (right). COGO points and PLSS corners, Washington County, Minnesota.

boundary lines, although that would have been extremely helpful. Therefore, this mixture of field positions (PLSS corners) with paper records (recorded legal descriptions) did not lend itself to the straightforward development of a single positional accuracy statement.

One might expect a pattern of higher levels of accuracy along PLSS section lines with lessening accuracy toward the interior of a PLSS section. Even if an inward rate of change could be predicted, more complications would arise due to the nature of land descriptions of public record. The position of features such as boundary corners described in these documents may not always match the position of their physical counterparts on the ground. (See observations section for more discussion on why these disparities exist.) These circumstances presented a situation of such significance that it was difficult to apply and use the NSSDA under its literal definition. Because the positional accuracy of the data set was not uniform, the project team doubted a single positional accuracy value could properly communicate the positional accuracy of the entire data set.

Due to this concern, Washington County chose to conduct a study of the COGO data set using some methods of the NSSDA. The result of this study would then be combined with an observation statement in the metadata.

County survey field crews selected and analyzed 21 random property corners in a study area of a single PLSS township. (See figure C.3.) Points were evenly spaced throughout the township and were of a mixture of platted and metes and bounds parcels. Post-processed differential GPS techniques were used to collect the points. A general comparison was made to see if any of the 21 points were near PLSS corner positions recovered in more recent years; that is, PLSS positions recovered subsequent to subdivision development and boundary monumentation in their vicinity (see figure C.4). It appears none of the test points were influenced by incorrectly assumed PLSS corner positions.

The worksheet

From the list of test points collected with GPS equipment, an AutoLISP script was created to import these points into the AutoCAD drawing containing the features. A second AutoLISP script was written to select the feature points that correspond to the GPS test points and import those coordinate values into a text file. This coordinate

Table C.1. Independent coordinate data and coordinate geometry test point comparison.

Point #	Point description	x (independent)	x (test)	diff in x	(diff in x)^2	y (independent)	y (test)	diff in y	(diff in y)^2	(diff in x)^2 + (diff in y)^2
10751	r/w & lot line (m&b)	486062.125	486061.709	0.4	0.2	168699.106	168698.974	0.1	0.0	0.2
1100	r/w & lot line (platted)	480383.263	480380.433	2.8	8.0	168103.428	168103.496	-0.1	0.0	8.0
11730	r/w & lot line (m&b)	491133.630	491133.362	0.3	0.1	153041.796	153041.828	0.0	0.0	0.1
1382	r/w & lot line (platted)	462816.265	462816.057	0.2	0.0	166767.786	166767.874	-0.1	0.0	0.1
1397	r/w & lot line (platted)	470589.879	470588.959	0.9	0.8	166326.072	166325.827	0.2	0.1	0.9
1490	r/w & lot line (m&b)	492381.275	492381.352	-0.1	0.0	166191.528	166191.305	0.2	0.0	0.1
2901	r/w & lot line (m&b)	487165.209	487165.039	0.2	0.0	159005.809	159005.818	0.0	0.0	0.0
6180	r/w & lot line (platted)	461796.422	461795.986	0.4	0.2	172592.941	172593.162	-0.2	0.0	0.2
7100	r/w & lot line (platted)	466652.141	466651.230	0.9	0.8	162901.920	162901.132	0.8	0.6	1.5
lot_1_2	r/w & lot line (platted)	481423.044	481422.194	0.8	0.7	173240.868	173240.547	0.3	0.1	0.8
11840	r/w & lot line (platted)	491813.966	491813.949	0.0	0.0	147708.306	147708.645	-0.3	0.1	0.1
3960	r/w & lot line (platted)	483922.111	483922.116	0.0	0.0	153178.492	153178.429	0.1	0.0	0.0
4041	r/w & lot line (platted)	479920.587	479920.492	0.1	0.0	152711.877	152711.858	0.0	0.0	0.0
5120	r/w & lot line (platted)	475454.065	475453.940	0.1	0.0	147133.085	147133.258	-0.2	0.0	0.0
5549	r/w & lot line (platted)	469407.975	469407.927	0.0	0.0	144480.696	144480.912	-0.2	0.0	0.0
6391	r/w & lot line (platted)	463062.352	463062.426	-0.1	0.0	143447.557	143447.761	-0.2	0.0	0.0
6576	r/w & lot line (platted)	463813.337	463813.443	-0.1	0.0	155699.943	155700.107	-0.2	0.0	0.0
8009	r/w & lot line (platted)	472135.343	472135.103	0.2	0.1	153996.576	153996.484	0.1	0.0	0.1
9336	r/w & lot line (platted)	478399.063	478399.053	0.0	0.0	157767.858	157767.940	-0.1	0.0	0.0
9378	r/w & lot line (platted)	478840.112	478839.711	0.4	0.2	148370.597	148370.816	-0.2	0.0	0.2
4786	r/w & lot line (platted)	465173.302	465173.120	0.2	0.0	148308.262	148308.520	-0.3	0.1	0.1
									sum	12.5
									average	0.6
									RMSE	0.8
									NSSDA	**1.3**

text file was then inserted into the spreadsheet table where calculations could be performed.

For linear features identified in the river valleys, points were selected at regular intervals from both the GPS control values and the check point data set. The number of points collected from each feature area ranged from four to 22 depending on the nature of the selected feature. These points were fed into an individual spreadsheet template. Two spreadsheet tables are provided as examples (see tables C.1 and C.2).

The positional accuracy statistic

A preliminary comparison was made of the digitized part of this data set. Several divisions of the overall 50 points were made. Separate spreadsheets comparing each were prepared. Points groupings were: north half of the county; south half of the county; 25 of 50 points selected at random; and all 50 points. The results were: 25 feet, 20 feet, 23 feet and 23 feet, respectively. This shows good uniformity.

Table C.2. Independent coordinate data and digitized test point comparison.

Point #	Point description	x (independent)	x (test)	diff in x	(diff in x)^2	y (independent)	y (test)	diff in y	(diff in y)^2	(diff in x)^2 + (diff in y)^2
34	152nd-stream 3	459897.8245	459900.2241	-2	6	254995.3250	254990.1862	5	26	32
35	132nd-Isleton 4	475603.3345	475602.9600	0	0	244363.6045	244371.4900	-8	62	62
36	155th-Manning 5	489350.1000	489350.1700	0	0	256106.3855	256110.1900	-4	14	14
37	180th-Keystone 6	483572.5260	483572.5700	0	0	269361.1230	269357.1800	4	16	16
38	May-RR 7	494171.3170	494160.1307	11	125	238673.6400	238666.0810	8	57	182
40	Otchipwe-94th 9	505295.9165	505293.1600	3	8	223453.3670	223446.1800	7	52	59
41	Neal-BrwnsCr. 10	497444.6805	497461.9147	-17	297	218442.2285	218479.5031	-37	1389	1686
42	75th-Keats 11	481800.0900	481797.1300	3	9	213775.5375	213762.8200	13	162	170
43	Irish-RR 12	475144.8540	475146.2412	-1	2	233082.6265	233082.4062	0	0	2
44	Linc-Robert 13	466236.2490	466238.0000	-2	3	211022.1690	211022.2500	0	0	3
45	C.R. 6-Stream 14	475253.3275	475247.4079	6	35	189933.5615	189931.4246	2	5	40
46	4th-Grd.Ang. 15	472999.8705	473000.8200	-1	1	175394.1410	175391.2300	3	8	9
47	Lake-Century 16	461164.5183	461162.2000	2	5	163210.0978	163207.3600	3	7	13
49	65th-Geneva 18	460948.6040	460948.0200	1	0	140008.1990	140006.2700	2	4	4
50	50th-ditch 19	496582.6795	496567.2195	15	239	147523.4710	147536.9953	-14	183	422
51	Jama-EPDR 20	474434.7155	474434.5700	0	0	126212.9210	126207.5300	5	29	29
52	Pioneer-GCID 21	460963.7915	460964.1100	0	0	118776.1925	118775.1700	1	1	1
53	127th-NB10 22	493944.2820	493949.9100	-6	32	106859.0630	106859.2000	0	0	32
54	Wash-Frontg 23	500142.5240	500140.4300	2	4	206064.7665	206062.9800	2	3	8
55	Point-RR 24	513038.7305	513036.5199	2	5	203149.2675	203144.4737	5	23	28
56	30th-Norman 25	498843.6095	498848.6000	-5	25	189987.0710	189984.8500	2	5	30
57	Rivercrest-Riv 26	516059.2450	516059.0524	0	0	180143.2225	180136.5409	7	45	45
58	Ramp-S.B.15 27	492428.0940	492427.6400	0	0	173572.2310	173557.3200	15	222	223
59	Indian-Hud. 28	500207.7530	500207.0900	1	0	173314.2015	173312.5700	2	3	3
60	VllyCr.-Put. 29	512300.0370	512306.3396	-6	40	162002.5330	162005.9951	-3	12	52
62	87th-Quadrant 30	513787.7670	513805.4900	-18	314	128008.0970	128011.9900	-4	15	329
2	Road-RR	512838.5425	512832.1230	6	41	265305.5275	265304.6426	1	1	42
3	Road-Road	513804.5885	513779.1351	25	648	265288.9815	265292.0679	-3	10	657
4	Road-RR	506995.3440	506986.9698	8	70	259036.2875	259039.1469	-3	8	78
5	Road-Road	505890.0345	505900.1300	-10	102	267608.0790	267586.4100	22	470	571
6	Road-Road	499522.8775	499516.9900	6	35	268070.4880	268057.5600	13	167	202
7	Road-Road	500886.3235	500889.8827	-4	13	277084.7130	277076.8184	8	62	75
9	Road-Road	506832.2160	506832.9800	-1	1	284524.2900	284524.2300	0	0	1
15	Road-Road	512469.9380	512494.5300	-25	605	300556.4700	300550.1500	6	40	645
16	Road-Road	499541.7365	499542.9300	-1	1	295469.7610	295470.3500	-1	0	2
19	Stream-Road	495674.4090	495672.6012	2	3	295158.3380	295155.4875	3	8	11
20	Road-Road	493348.3880	493356.4200	-8	65	283897.9365	283893.6900	4	18	83
21	Road-Road	486511.0920	486512.7100	-2	3	275873.6795	275878.2400	-5	21	23
22	Road-Road	483617.1320	483617.8700	-1	1	275899.2500	275902.5300	-3	11	11
23	Stream-Road	479455.9240	479472.6850	-17	281	291709.0365	291683.7402	25	640	921
24	Road-Road	469037.3285	469025.2000	12	147	298365.8860	298366.1900	0	0	147
25	Road-Road	456160.0730	456172.3500	-12	151	300964.3090	300970.9900	-7	45	195
26	Road-Road	453048.3560	453051.8400	-3	12	300995.9335	301016.3400	-20	416	429
31	Road-Road	471995.9350	472007.9600	-12	145	289610.8105	289606.5200	4	18	163
32	Road-Shoreline	471828.5090	471845.6500	-17	294	289748.0805	289734.9000	13	174	468
33	Stream-Road	473084.1800	473083.8500	0	0	283256.8455	283250.2300	7	44	44
36	Stream-Road	467667.6425	467674.1085	-6	42	272602.0220	272610.5601	-9	73	115
37	Stream-Stream	452112.0170	452108.5500	3	12	277310.9820	277322.2800	-11	128	140
38	Road-Road	451973.4935	451977.1198	-4	13	269628.2735	269625.6395	3	7	20
41	Road-Road	473066.7460	473069.3294	-3	7	264154.8040	264153.7500	1	1	8
									sum	8545
									average	171
									RMSE	13
									NSSDA	**23**

Digitized linear features. Although unique, the result shown for the special linear features did produce a result matching estimates developed years earlier from experience in mapping these areas. A horizontal error of up to 120 feet can be expected for the digitized features in the high relief areas.

COGO features. The method chosen to compare values between control and data checkpoints does not entirely conform to the NSSDA. Limiting the scope of control to a single township was intentional due to the nature of the COGO data set. For this reason, the potential cost as compared to the final value could not be justified in locating control countywide. Apparently by chance, results of the study area seem to indicate that none of the 21 points selected for control are related to a recovered PLSS corner position type. From experience in building the parcel database, the 1.3 foot result (table C.1) meets expectations. This appeared a realistic representation of what exists over most of the county in areas not influenced by a corrected section corner position.

The accuracy statement and metadata

The project team thought it would be useful and informative for potential data users to better understand the methods used to derive the accuracy statements. The team developed a brief description of the test to fill out the positional accuracy portion of the metadata, in addition to pointing the reader to other sources of information.

In the case of COGO data, the project team believed a specialized summary statement can more appropriately communicate the positional accuracy of the data than can the accuracy reporting statement of the NSSDA. Although this is not as simple and standardized as the NSSDA statement, this method does provide a higher level of information to the user, hopefully increasing the user's confidence in the data and allowing the data to be used more appropriately.

Observations and comments

An optional method of collecting COGO control was considered but not used by Washington County, but it may be instructive to others attempting to implement the NSSDA.

The county was divided into quadrants. Five points were selected within each quadrant. Consideration was given to areas of greater feature density, occasionally concentrating more points in these areas. A buffer of 2 miles (the diameter of 4 miles is approximately equal to 10 percent of the diagonal distance across the data set) was generated around each point. The NSSDA calls for a minimum of 20 points. The following types of points were designated:

- railroad crossing with highway
- lot corner in subdivision plat

Figure C.5. Detailed positional accuracy statements as reported in metadata.

Horizontal positional accuracy	Digitized features of the parcel map database outside areas of high vertical relief tested 23 feet horizontal accuracy at the 95% confidence level using modified NSSDA testing procedures. See Section 5 for entity information of digitized feature groups. See also Lineage portion of Section 2 for additional background. For a complete report of the testing procedures used contact Washington County Surveyor's Office as noted in Section 6, Distribution Information. Digitized features of the parcel map database within areas of high vertical relief tested 119 feet horizontal accuracy by estimation as described in the complete report noted above. All other features are generated by coordinate geometry and are based on a framework of accurately located PLSS corners positions used with public information of record. Computed positions of parcel boundaries are not based on individual field survey. Although tests of randomly selected points for comparison may show high accuracy between field and parcel map content, variations between boundary monumentation and legal descriptions of record can and do exist. Caution is necessary in use of land boundary data shown. Contact the Washington County Surveyor's Office for more information.
Vertical positional accuracy	Not applicable.

- lot corner (old plat, metes or bounds) based on certificate of survey
- road intersection
- road intersection at PLSS corner
- intersection of projected right-of-way line and road centerline
- radius point on cul-de-sac
- road right-of-way limit at B corner

The parcel map was developed one PLSS section at a time. Typically the cartographer relied on the PLSS as the foundation for information created. As a result, the positioning of points at the section corners and along the outer edges was more reliable than within the interior. Because of the way sections are normally subdivided, the least reliable mapped parcels were located near the interior of each quarter and quarter/quarter section. Expecting exterior section points to be the most accurate, the project team focused on interior points to anticipate the worst case accuracy. Corners of property ownership make up an estimated 90 percent of the parcel database. For this reason it seemed appropriate to have a proportionate representation. The allotment of points was defined as follows:

- subdivision plats, 6 points, 30 percent
- metes and bounds parcels, 6 points, 30 percent
- right-of-way corners, road intersections, railroad/highway intersections, 6 points, 30 percent
- PLSS, 2 points, 10 percent

Where possible these three groups were further divided into categories by 50-year intervals, such as sources from 1850 to 1900; 1900 to 1950; and 1950 to 1998. Again, this option was not used, but may have merit in other situations.

Incorrectly used PLSS corner positions. The following discussion exemplifies only a single aspect of why land descriptions do not always match their positions on the ground and what impact this can have in trying to apply the NSSDA.

Increased activity in the monument maintenance of the Public Land Survey System to support GIS development over the past 15 to 30 years has provided for a high rate of consistency between land parcels and their descriptions of record. Actually older parcels dating back 100 to 150 years are also quite consistent in comparison of ground position and written documents of record. Unfortunately, inconsistencies do exist. The inconsistencies come from situations where subdivision plats and metes and bounds parcels were established based on an incorrect PLSS corner position. In areas where an ongoing maintenance program of the PLSS has not existed, the likelihood of this occurrence is much greater. Where an incorrectly assumed PLSS position has caused land occupation to be inconsistent with a property description of public record, laws exist that may protect the landowner and can sometimes help to remedy the situation. Unfortunately, the legal record is not always changed. In these areas, statements of expected positional accuracy using strict application of the NSSDA could mislead the digital data user. More information is required in the metadata to keep the data user properly informed.

A study of PLSS corners. A single PLSS corner can control the position of parcels in up to four PLSS sections. This is essentially the limit of potential impact for a single discrepancy in PLSS corner position. A study within a single township (36 PLSS sections) randomly selected from within Washington County estimated the frequency of these occurrences. Of the 138 PLSS corners in the township, 10 on record had been corrected from a previously established incorrect position. The length of positional adjustment varied from 0.5 feet to 34 feet. It is known that at least as many others have also existed but clear documentation of their details does not exist. Unfortunately it is possible that parcel boundaries were established on the ground based on these incorrect PLSS positions. The lack of information about the time period in which these incorrect positions were used further complicates the issue. Relative accuracy may be very high in these situations while absolute accuracy is significantly less. This situation can seriously affect the validity of applying the NSSDA to a parcel boundary data set.

Mixed meaning of positional accuracy. What do the results mean when random field monumented property corners are chosen for controlling position when compared against Washington County's method in establishing its base map? At every PLSS section corner positional accuracy is at its best. Here ground position was used as the starting base for digitally mapping the legally recorded documents. As one moves to the interior of a section, ground position compared to the legal record may or may not diminish. Mapping the interior of the section primarily follows more of a theoretical record. A blind comparison of a mapped parcel corner with a randomly selected corresponding monumented ground position can easily be performed. However, a discrepancy does not necessarily dictate that the relative positional accuracy of the parcel boundary is anything less than perfect when ground truth is in disagreement with the legal record.

With legal rights based many times on possession, errors in legal records or flaws in measurement methods may not actually reduce the accuracy of occupied ownership. Laws provide protection under certain circumstances. There are legal mechanisms to protect owners within a subdivision, for example, from all having to relocate their homes, physical improvements and land boundaries because an incorrect PLSS corner position was involved. When the NSSDA standard is applied to this situation, solutions to address some of the standard's components are not so straightforward. It is difficult to find a control point three or more times greater in accuracy than something that is theoretical. When interpretations of law are introduced, ambiguity can further cloud the situation. Although difficult to grasp for the nonprofessional who is not familiar with land records and surveying methods, it is vitally important for the common user of parcel based land information to have a certain confidence level when making decisions based on information from GIS analysis.

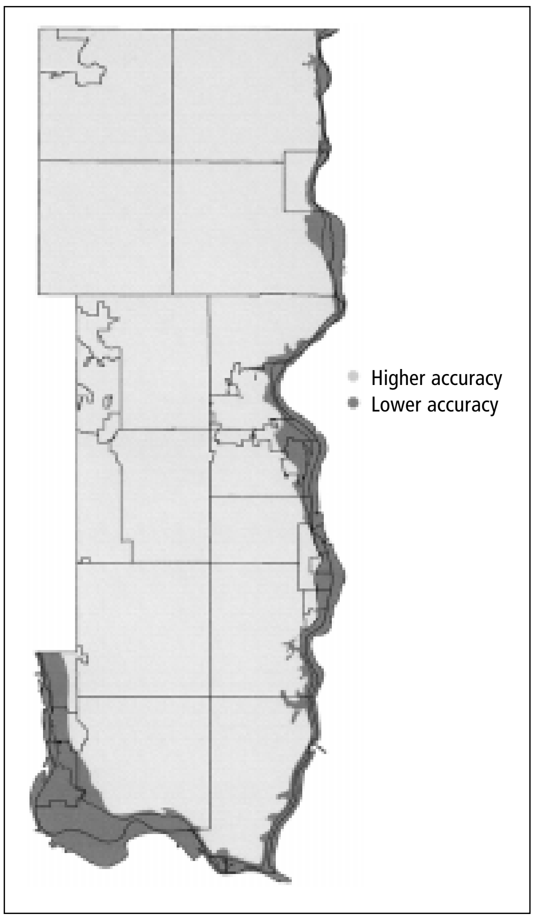


Figure C.6. Digitized features of Washington County grouped by accuracy.

General comments. The nature of a parcel data set presents some unique challenges. The idea of grouping data types or features of differing accuracy, especially if presented to the user as a graphic, can communicate the quantity and location of error quite effectively (see figure C.6). As land boundaries have evolved since the mid-1800s in Washington County, so have the measurement methods. Land settlement and corresponding boundary development have been random but some consistent patterns may be found. Groups of accuracy can perhaps be tied to parcels reflecting their original measurement quality — from Gunter's chain to steel tape to computerized electronics and satellites. Collections of parcels may owe their associated accuracy to whether the terrain is level or hilly. The added characteristics of place in time, or the nebulous facets of law make for additional complications.

This example project required good familiarity with the subject data set to properly apply the NSSDA according to its specifications. Spatial features of the database were known to be of varying accuracy and were selectively grouped in some cases. In other instances, applying the standard became difficult. A standard designed to address the vast range of GIS data types can, at times, have a "one size fits all" feeling. To avoid this impression and still accomplish the task, the underlying intent must be considered, with reasonable and straightforward approaches explored and applied. This was the intention when applying the NSSDA to the Washington County parcel base. The major obstacles encountered were a result of the nature of the features being represented in the parcel base and the implications of boundary law.

The Washington County project team believes the best solution is to provide as much information as possible by meeting metadata standards and showing thoroughness in critical background information. The NSSDA can provide important and needed information, but may not give the complete picture in all applications.

Case Study D

The Lawrence Group

Describing positional accuracy of a street centerline data set when NSSDA testing cannot be applied

PROJECT TEAM

Jim Maxwell
Vice president, GIS services

Dan Och
GIS specialist

The project

Although the intent was to use the National Standard for Spatial Data Accuracy to describe, measure and report the positional accuracy of a regional street centerline data set, the Lawrence Group was unable to implement the NSSDA due to time and budget constraints. This modified project, however, is a good example of how to report useful information about positional accuracy, even when NSSDA testing is not applied.

The tested data set

The Lawrence Group Street Centerline Database is a digital network of pavement and road right-of-way centerlines covering the Twin Cities metropolitan area. The seven-county regional data set consists of approximately 140,000 graphic elements as well as associated attribute information. It is generally created in state plane coordinates and provided to users in UTM coordinates. As is typical of regional data sets, the positional source documents used for its creation were numerous and varied in quality.

The initial centerline data set was created over a period of several years. During the last decade, a variety of additional sources of information have become available. Rather than using only hardcopy references such as half-section and plat maps, sources such as digital parcel base maps, digital orthophotos, global positioning references, enhanced control points and satellite imagery have become increasingly accessible. The variety of reference sources used is also due to the limited availability of certain source materials. For example, some counties have completed highly accurate digital source material, while other counties have not. Generally, the quality of source material has improved, as evidenced by the increasing availability of high resolution digital ortho imagery.

The independent data set

Several issues arose in the attempt to choose independent points applicable to a regional street centerline file. It was necessary to find control points that were *known* to be more accurate *and* to permit the data set to be tested at numerous, ideally random locations *and* to have point locations that could be identified on the street centerlines data set.

Since this regional data set was developed from a variety of sources, it was also necessary to have an understanding of the accuracy and availability of potential independent data sets in different parts of the region.

Evaluating existing independent data sets. If Public Land Survey corners were to be used, they would have to be associated with street centerlines. This may be possible at street intersections that lie at survey corners. However, this would essentially eliminate the random selection of control points, and force accuracy measurement only at intersections that are typically reliable and well established.

Since Public Land Survey corners were unable to provide a random and unbiased sampling, other independent reference points with reference information pertaining to the streets or street centerlines were needed. The Lawrence Group found such a set of information through the Minnesota Department of Transportation in the form of county maps showing various control marks as determined by different organizations. These marks are numbered and, although they do not provide for a random selection of measurable points, there were sufficient corresponding coordinate information and street centerline tie descriptions to develop a reasonable sampling of measurement locations. Unfortunately, it was determined that using these descriptive ties to reference the streets was insufficient for establishing a set of points that meet NSSDA requirements due to the potential margin of error in the descriptive variables.

Without a suitable independent data set, the only option to establish a new set of control points was by means of global positioning devices. Unfortunately, The Lawrence Group did not have the staff time, expertise, equipment and budget needed to collect such a data set. The result is that the data set remains untested using the NSSDA.

Providing other positional accuracy information. If fewer than 20 test points are available, the NSSDA recommends three alternatives for determining positional accuracy as described in another FGDC standard — the Spatial Data Transfer Standard. They are: 1) deductive estimate, 2) internal evidence and 3) comparison to source. In this case, a significant amount of internal evidence is available and should be provided in the metadata.

The worksheet

Not applicable for reasons specified above.

The positional accuracy statistic

Not applicable for reasons specified above.

The accuracy statement and metadata

See figure D.1. for an example of how a descriptive metadata record can be presented when NSSDA testing is not applied.

Observations and comments

While the Global Positioning System option for collecting independent data is an excellent means of providing a positional accuracy measuring solution, some practical obstacles should be noted. First, high quality GPS equipment is costly and not always available to the data provider. Second, the equipment and associated base station utilization requires expertise. Third, this method requires significant staff hours since base stations need to be attended and GPS receivers require a technician. The effort could require four to seven people to coordinate a project of this size.

Overcoming these obstacles may be difficult. For many organizations, funds and expertise are not readily available. It may require a coordinated effort by any or all parties with an interest in the positional accuracy of the data set. The coordinated effort may require the sharing of costs, staff and expertise. However, if these obstacles are overcome, this option can be an excellent method for measuring and conforming to NSSDA guidelines.

Figure D.1. Detailed positional accuracy statements as reported in metadata.

Horizontal positional accuracy	There are differing degrees of positional accuracy throughout the centerline data set. Positional accuracy of the data is adjusted to available sources known to have greater degrees of accuracy than those used to generate the existing data. However, in all cases (including instances where COGO line work has been the reference) street segments were generalized to at least 4 feet in order to reduce the number of vertices, thus reducing file size. The goal is to have 95% of roads located to within the digital right-of-way or pavement centerlines provided from counties, where such digital data is available. In other areas, 95% of roads are intended to be within ten meters of the road or right-of-way center. The centerlines of Dakota and Washington counties were adjusted by referencing the Washington and Dakota County Surveyors department geospatial data set. Positional adjustments were also made to data supplied by the Ramsey County Public Works department in May 1998. These adjustments were made relative to each county's centerline data set. Ramsey County was continuing to review its own data for positional accuracy; however, adjustment was necessary in order that users could more efficiently use both data sets concurrently. Due to the large number of changes, the Ramsey County adjustments are noted here in the metadata rather than being reflected in the quarterly change report as individual segment changes. Positional adjustments were made as of July 31, 1998 to street centerlines within the city of Minneapolis. These changes were made relative to street centerlines provided by the city of Minneapolis. Due to the large number of changes, these positional adjustments are noted here rather than in the quarterly change report. For all other areas, various resources were used. For more information please review the "Lineage" section of this metadata record.
Vertical positional accuracy	Not applicable.

Case Study E

Minnesota Department of Natural Resources

Attempting to apply the NSSDA to a statewide watershed data set containing nondiscrete boundaries

PROJECT TEAM

Glenn Radde
Wetlands GIS manager

Joe Gibson
IS manager, Department of Natural Resources, Waters

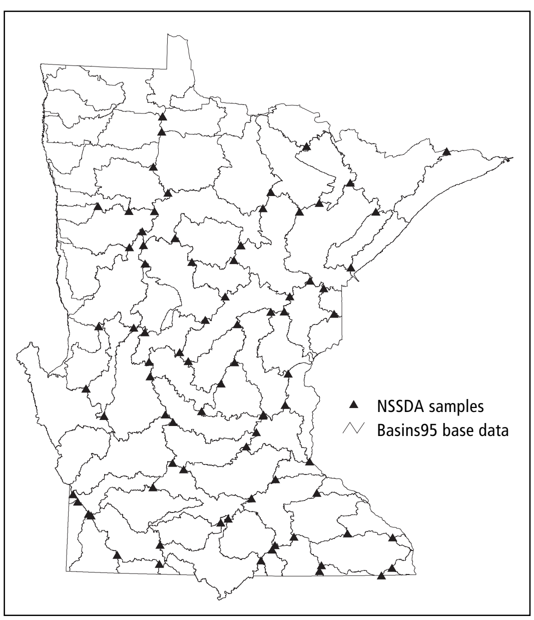
Jim Solstad
Surface water hydrologist, Department of Natural Resources, Waters

The project

The objective was to explore whether or not it is reasonable to develop positional accuracy assessments of height-of-land watershed delineations. Unlike parcel corners, township survey monuments and other similar features, it is conceptually difficult to meaningfully evaluate watershed delineations for positional accuracy. Professionals use detailed topographic maps and expertise to locate height-of-land boundaries for watersheds or drainage basins. In the end, what matters most is not the positional accuracy of individual points, but the overall fidelity a watershed delineation has with field experience regarding how a particular drainage basin actually functions.

However, since the state has a growing number of watershed delineations, from the major basin on down to lake- and ditch-shed, positional accuracy assessments might be useful as a way to evaluate the suitability of a particular delineation for a particular task. For example, the DNR was requested to provide a reasonably accurate delineation of the Red River of the North regional drainage for use in a climatological modeling project.

Figure E.1. NSSDA major watershed sample points



The most detailed data covers only Minnesota, with access to national Hydrologic Unit Code data of drainage basins at two scales: 1:250,000 and 1:2 million. A positional accuracy assessment might help clarify the use of these smaller scale coverages for more detailed studies. This project established sample points to compare the positional accuracy of these small scale coverages to the more detailed statewide coverage DNR Waters has developed in cooperation with various state and federal agencies.

The tested data set

The project team tested the positional accuracy of the U.S. Geological Survey 1:250,000 and 1:2 million Hydrologic Unit Code data sets. Both of these national data sets were derived from 1:500,000 and 1:250,000 source maps during the 1970s. Since these coverages are readily available on the Internet, they remain authoritative, national delineations. For this purpose, the federal eight-digit Hydrologic Unit Codes correspond to Minnesota's major and minor watershed identification numbers. These evaluations focused on comparing the delineations of major watersheds with the corresponding eight-digit HUCs.

The independent data set

The project team used the /basins95 coverage as an independent data set of higher accuracy. The /basins95 coverage was first created in 1977 by manually delineating height-of-land watersheds of 5-6 square miles in area on USGS 7.5-minute quadrangles. In the early 1990s, these mylars were assembled and scanned as 1:100,000-scale map sheets. A diverse group, including state and federal agencies and Mankato State University, worked on verifying and correcting known errors in this data set. In 1993, the Department of Natural Resources incorporated major and minor drainage basin data, at 1:24,000-scale, from both the USGS for the Minnesota River Basin and Mankato State University Water Resources Center for its 13-county service area in southwestern Minnesota. The result of these efforts has been the creation of the most authoritative and comprehensive data set of drainage basins for Minnesota.

It is important to note that there are discrepancies between the three data sets discussed here. For example, the /basins95 coverage contains delineations of 81 major watersheds, ranging from 12.54 square miles in size to 2,852.87 square miles. The USGS 1:250,000-scale HUC coverage delineates 82 major drainage basins ranging from 6.14 square miles to 2,885.94 square miles. Finally, the USGS 1:2 million-scale HUC coverage delineates 86 major drainage areas ranging in size from 16.08 square miles to 2,927.15 square miles, with some sliver polygons of 0.03 square miles.

The worksheet

Figure E.1 shows the location of 82 sample points from which to compare each of the three drainage basin delineations. The project team selected outlets, pour points or other edge or boundary features that provided clean sample points representing the intersection of one major watershed with another. The team thought that a positional accuracy assessment of these points would provide a reasonable comparison between the three coverages.

Tables E.1 and E.2 show the results of the comparisons. The table evaluating all 82 sample points across the three coverages is quite lengthy, so the 10 best and worst sample points are presented based on how they ranked as squared and summed differences.

The positional accuracy statistic

The NSSDA statistic for the 1:250,000-scale HUC coverage evaluated against the /basins95 data is 4,802.96 meters (or 4.8 kilometers), and the 1:2 million-scale HUC coverage statistics is 5,069.22 meters (or 5.1 kilometers). When the project team compared the 1:2 million-scale HUC coverage against the more detailed 1:250,000-scale HUC data, the resulting NSSDA statistic was 3,785.50 meters (or 3.8 kilometers).

While these numbers may seem large, the positional accuracy reported here does not preclude a

Table E.1 (left). Comparing 1:250,000 HUC to DNR's 1995 basins; 10 best and worst sample points.

Table E.2 (right). Comparing 1:2 million HUC to DNR's 1995 basins; 10 best and worst sample points.

Best Ten Sample Points					
ID	X-Coord	Y-Coord	X-Diff	Y-Diff	Squared/Summed
58	316246	5194439	22	16	740
9	515466	4954879	104	-27	11545
35	327887	5257696	-62	-106	15080
70	450940	4990434	-125	0	15625
63	489464	5061367	116	58	16820
80	322470	5057799	-148	42	23668
20	248112	4891460	27	-161	26650
71	438771	4973151	-177	0	31329
15	337579	5372432	69	-180	37161
16	337578	5372430	68	-183	38113

Worst Ten Sample Points					
ID	X-Coord	Y-Coord	X-Diff	Y-Diff	Squared/Summed
76	473392	4853699	2325	3452	17321929
77	470152	4848275	4438	2254	24776360
25	363075	4945478	2414	-5375	34718021
47	389175	5126315	4196	4416	37107472
61	484658	5136228	1198	-6197	39838013
13	564580	5292404	-5981	-3006	44808397
18	282652	4841907	-1668	6740	48209824
55	313005	5233367	-6482	2672	49155908
39	423720	5198685	-5635	5153	58306634
64	485839	5023201	1794	7600	60978436

SUM	631,448,624
AVG	7,700,592.98
RMSE	2,774.99
NSSDA	**4,802.96**

Best Ten Sample Points					
ID	X-Coor	Y-Coord	X-Diff	Y-Diff	Squared/Summed
66	408186	5049328	-8	-13	233
58	316246	5194439	409	-58	170,645
15	337579	5372432	409	-108	178,945
16	337578	5372430	409	-112	179,825
8	523951	4917059	-77	-473	229,658
9	515466	4954879	393	-285	235,674
51	459012	5261312	498	-231	301,365
69	459036	5010406	-241	-494	302,117
23	325843	4923739	70	-575	335,525
5	529633	4829235	143	-664	461,345

Worst Ten Sample Points					
ID	X-Coord	Y-Coord	X-Diff	Y-Diff	Squared/Summed
54	527131	5267895	-4166	-570	17,680,456
37	326526	5311468	-122	4268	18,230,708
24	349658	4953280	-4489	1653	22,883,530
18	282652	4841907	-1387	5143	28,374,218
47	389175	5126315	4172	3928	32,834,768
13	564580	5292404	-5292	-3006	37,041,300
61	484658	5136228	1372	-7048	51,556,688
64	485839	5023201	2005	7785	64,626,250
55	313005	5233367	-8810	2923	86,160,029
39	423720	5198685	-6918	6395	88,754,749

SUM	703,400,916
AVG	8.578,059.95
RMSE	2,928.83
NSSDA	**5,069.22**

number of valid uses of the data sets, as discussed in the observations section.

The accuracy statement and metadata

Figure E.2 contains formal NSSDA reports for both 1:250,000-scale and 1:2 million-scale HUC data sets.

Observations and comments

At first glance, it may appear that things are not all that well between the three height-of-land delineations of Minnesota's major watersheds. Federal efforts were begun and finished long before Minnesota did any of its more detailed studies. Minnesota has subsequently revised this more detailed data, and is currently cooperatively working on adding lakeshed delineations to the drainage basin hierarchy.

This small study does demonstrate the process of identifying, calculating and reporting positional accuracy statistics for watershed pour points, outlets and similar features. Most importantly, it validates that these smaller scale coverages still have value. The largest NSSDA statistic was slightly greater than 5 kilometers, which was well within the climatology project's 10 kilometer accuracy requirement. While any of these coverages would have worked well, the NSSDA statistic does provide useful insights as to when it is wise to use larger-scale data.

Figure E.2. Positional accuracy statements as reported in metadata.

Horizontal positional accuracy	Using the National Standard for Spatial Data Accuracy, the 1:250,000-scale HUC data tested 4,802.96 meters horizontal accuracy at 95% confidence level. Using the National Standard for Spatial Data Accuracy, the 1:2 million-scale HUC data tested 5,069.22 meters horizontal accuracy at 95% confidence level.
Vertical positional accuracy	Not applicable.

Appendix

National Map Accuracy Standards

In 1941, the U.S. Bureau of the Budget (now the Office of Management and Budget) developed positional accuracy specifications for federal maps known as the United States National Map Accuracy Standards. Although revised periodically, the standards are still in use today.

The standards were written for published paper maps at a time long before digital spatial data. Accuracy testing was generally applied to a map series by federal agencies using representative sampling. Positional accuracy in this and other standards is defined by two components: horizontal accuracy and vertical accuracy. Horizontal accuracy is determined by map scale. A threshold was established for maps with a published scale larger or smaller than 1:20,000. For maps with scales larger than 1:20,000, not more than 10 percent of the points tested can be in error by more than 1/30 inch. For maps with publication scales of 1:20,000 or smaller, the error distance decreases to 1/50 inch.

This standard relies upon testing of "well-defined points." Selected points identifiable on both the map and on the ground are measured and the difference between the two — the error between mapped and actual location — is recorded in *map inches* at the publication scale. Standards for a typical 1:24,000-scale U.S. Geological Survey quadrangle is 40 feet, or 1/50 inch at map scale. The statistic used is based on a 90 percent confidence level; 90 percent of the tested points must fall within the standard's threshold to comply.

Vertical accuracy involves a methodology similar to that used in horizontal accuracy testing. The vertical testing also references map publication scale and involves a 90 percent threshold regarding number of test points required. At all map scales, the maximum allowable vertical tolerance is one-half the published contour interval.

National Map Accuracy Standards do not require any description of the testing process to determine map accuracy. The standard does note that "surveys of higher accuracy" may be used for testing purposes. It is generally assumed that surveys of higher accuracy are acquired independently, although no specific criteria are provided.

In summary, general characteristics of the National Map Accuracy Standards include positional accuracy testing that is largely dependent on map publication scale with test results of selected points reported in inches at map scale. The standards specify that 90 percent of the well-defined points tested must fall within a specified tolerance based on the map publication scale. If the map meets accuracy standard requirements it can be labeled as complying with National Map Accuracy Standards.

Another map accuracy standard of potential interest, but not discussed in this handbook, is the American Society for Photogrammetry and Remote Sensing Accuracy Standards for Large-Scale Maps. These accuracy standards address both horizontal and vertical accuracy, allow for multiple "classes" of map accuracy, and provide specific detail regarding map accuracy testing requirements. Several aspects of these standards are included in NSSDA and referenced in that document.

UNITED STATES NATIONAL MAP ACCURACY STANDARDS

With a view to the utmost economy and expedition in producing maps which fulfill not only the broad needs for standard or principal maps, but also the reasonable particular needs of individual agencies, standards of accuracy for published maps are defined as follows:

1. Horizontal Accuracy. For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale; for maps on publication scales of 1:20,000 or smaller, 1/50 inch. These limits of accuracy shall apply in all cases to positions of well- defined points only. Well-defined points are those that are easily visible or recoverable on the ground, such as the following: monuments or markers, such as bench marks, property boundary monuments, intersections of roads, railroads, etc.; corners of large buildings or structures (or center points of small buildings); etc. In general what is well defined will also be determined by what is plottable on the scale of the map with 1/100 inch. Thus while the intersection of two road or property lines meeting at right angles would come within a sensible interpretation, identification of the intersection of such lines meeting at an acute angle would obviously not be practicable within 1/100 inch. Similarly, features not identifiable upon the ground within close limits are not to be considered as test points within the limits quoted, even though their positions may be scaled closely upon the map. In this class would come timber lines, soil boundaries, etc.

2. Vertical Accuracy, as applied to contour maps on all publication scales, shall be such that not more than 10 percent of the elevations tested shall be in error more than one-half the contour interval. In checking elevations taken from the map, the apparent vertical error may be decreased by assuming a horizontal displacement within the permissible horizontal error for a map of that scale.

3. The accuracy of any map may be tested by comparing the positions of points whose locations or elevations are shown upon it with corresponding positions as determined by surveys of a higher accuracy. Tests shall be made by the producing agency, which shall also determine which of its maps are to be tested, and the extent of such testing.

4. Published maps meeting these accuracy requirements shall note this fact on their legends, as follows: "This map complies with National Map Accuracy Standards."

5. Published maps whose errors exceed those aforestated shall omit from their legends all mention of standard accuracy.

6. When a published map is a considerable enlargement of a map drawing (manuscript) or of a published map, that fact shall be stated in the legend. For example, "This map is an enlargement of a 1:20,000-scale map drawing," or "This map is an enlargement of a 1:24,000-scale published map."

7. To facilitate ready interchange and use of basic information for map construction among all federal map making agencies, manuscript maps and published maps, wherever economically feasible and consistent with the uses to which the map is to be put, shall conform to latitude and longitude boundaries, being 15 minutes of latitude and longitude, or 7.5 minutes or 3-3/4 minutes in size.

US Bureau of the Budget

Issued June 10, 1941
Revised April 26, 1943
Revised June 17, 1947