A Practical Global Spatial Data Model (GSDM) for the 21st Century

Earl F. Burkholder, PS, PE
Global COGO, Inc.
Circleville, Ohio 43113

Presented at Navigation 2000 - ION National Technical Meeting
Long Beach, California - January 23, 1998

BIOGRAPHY

Earl F. Burkholder is President of Global COGO, Inc., a firm formed to promote use of the 3-dimensional global spatial data model (GSDM). A former Editor of the ASCE Journal of Surveying Engineering, he taught upper division surveying courses including geodesy, cartography, adjustments, and programming at Oregon's Institute of Technology from 1980 through 1993. In 1990/91 he spent most of a sabbatical year at the University of Maine learning more about applications of new technology to modern surveying practice. Since then he has written on various aspects of the GSDM.

ABSTRACT

We often work with two dimensions, but spatial data are 3-dimensional (3-D). Modern measurement systems collect 3-D spatial data. Computer data bases store digital spatial data in 3-D. Human perception of spatial relationships is primarily visual. Mathematical models provide a conceptual connection between digital spatial data and its graphic representation - data visualization. Digital spatial data are also used to make analog products such as maps and other hardcopy diagrams. Numerical representation of spatial data includes listings of coordinates for individual points and relative point-pair relationships (such as bearings and distances) between points. Traditional models for spatial data include a simple flat earth model (used extensively in plane surveying) and the more complex ellipsoidal earth model (used in geodetic surveying). In each case, a natural distinction is made between horizontal and vertical due to the local perception of "up." Given the traditional use of horizontal and vertical datums, it follows that 3-D measurements are handled in terms of those models. However, except for local flat-earth problems, traditional equations for manipulating 3-D spatial data are more complex than equations for rectangular coordinates due to 1) using the ellipsoidal earth model, 2) using mixed units (angular units for latitude/longitude and meters for vertical), and 3) extensive use of 2-dimensional map projections to "flatten the earth."

A simple practical 3-dimensional global spatial data model (GSDM), which includes both functional and stochastic components, is presented. The GSDM 1) accommodates all modes of spatial data measurement, 2) utilizes a digital spatial data base containing earth-centered earth-fixed (ECEF) X/Y/Z coordinates, 3) does not distort physical measurements as does a 2-dimensional map projection, 4) uses one set of solid geometry equations world-wide, 5) portrays an accurate view of spatial data based upon any origin selected by the user, 6) preserves geometrical integrity by using local coordinate differences, 7) optionally stores stochastic model information in the X/Y/Z covariance matrix of each point, and 8) where covariances are stored, provides the standard deviation of each spatial data component and subsequently derived quantities such as distance, direction, area or volume.

INTRODUCTION

Models are a collection of defined rules used to make a conceptual connection between abstract ideas and human experience. A model is deemed "good" to the extent it is simple and appropriate, i.e. easy to understand and differences between modeled values and actual values are small. Judgement is required to evaluate the trade-off between the advantages of a model which is simple as compared to one which is more appropriate but more complex. For example, a simple plane triangle may be an appropriate model for a small tract of land. Due to earth's curvature, differences between computed values according to the model and corresponding values on the earth's surface become larger as distances become greater. At some point, the differences (also called systematic errors) become unacceptably large. The problem can be solved by using a more appropriate
model, in this case, a spherical triangle. The drawback is added complexity. Computations involving spherical trigonometry are more complex than computations based upon plane trigonometry. A popular alternate is to identify systematic error corrections which are applied to simple model computations to improve the agreement between computed values and corresponding real-world values. Two examples are spherical excess for triangles and scale factor corrections for map projections. Rather than focusing on corrections, the approach here is to reexamine fundamental issues and to identify a global spatial data model (GSDM) which is both better and simpler than existing models. The GSDM so identified also includes standard procedures for 3-dimensional error propagation and spatial data accuracy analysis world-wide.

HISTORICAL:

Existing spatial data models include a variety of 2- and 3-dimensional concepts ranging from simple to complex. Examples embodying various measurement units include:

- Flat earth - plane surveying & local mapping:
  1. Two-dimensional X/Y tangent plane coordinates (northing and eastings are also used).
  2. In the third dimension:
     a. Surveyors use profiles to show grades in terms of centerline stationing & elevation.
     b. Architects show elevation perspectives.
     c. Mappers use hachures and contour lines.

- Spherical earth - geography and navigation:
  1. Two-dimensional curvilinear latitude/longitude positions based upon spherical earth.
  2. Third dimension is elevation or altitude.

- Ellipsoidal earth - science and geodesy:
  1. Two-dimensional curvilinear latitude/longitude positions based upon ellipsoidal earth.
  2. Third dimension based upon:
     a. Distance from geoid, orthometric height.
     b. Distance from ellipsoid, ellipsoid height.

GLOBAL INFRASTRUCTURE:

The phrase being used by the Atlantic Institute and others to describe various facets of organizing and using spatial data is the "global spatial data infrastructure" (GSDI) which, according to Fritz Petersohn (1997), is all-inclusive and includes various models, e.g. economic, political, geometrical.... It is hoped the GSDM herein described is viewed with approbation by those addressing such broader global issues. One part of that challenge is examining each model carefully in terms of new technology, modes of operation, and integration with other models. Another part of the challenge is "looking forward" (Gibbons 1997) and avoiding the practice of retrofitting new technology to outdated models. The goal in all cases should be to identify and promote use of models which are both simple and appropriate. The GSDM meets those criteria and should serve as the geometrical foundation of the global spatial data infrastructure.

DEFINITION OF THE GSDM:

A formal definition of the GSDM (Burkholder 1997e) and issues associated with implementing the GSDM are considered separately. The GSDM has both functional and stochastic components and is essentially the geometrical portion of the 3-D Geodetic Model described by Leick (1995). As such, the GSDM is a collection of existing equations and contains little or no new science. However, the seminal contribution of the GSDM is defining an efficient global/local computational environment and delineating specific procedures for handling 3-D spatial data. Many computations, whether 2-D or 3-D, are easier and can be performed with greater integrity in terms of the GSDM. Furthermore, issues of error propagation and spatial data accuracy (both network and local) are handled with aplomb by the GSDM.

The Functional Model

The functional portion of the GSDM is an earth-centered earth-fixed (ECEF) right-handed rectangular coordinate system (DMA 1991). The GSDM:

- Has an origin located at the earth's center of mass.
- Is oriented in space by:
  1. a Z axis defined by the location of the Conventional Terrestrial Pole.
  2. an X axis defined by the zero longitude meridian in the equatorial plane.
  3. a Y axis in equatorial plane completing the right-handed system.
Uses meters as the unit of length.

Embodies rules of solid geometry (Burkholder 1993) by which:

1. Geocentric coordinates and differences are:

\[ \Delta X = X_j - X_i \quad \& \quad X_j = X_i + \Delta X \quad (1) \& (2) \]
\[ \Delta Y = Y_j - Y_i \quad \& \quad Y_j = Y_i + \Delta Y \quad (3) \& (4) \]
\[ \Delta Z = Z_j - Z_i \quad \& \quad Z_j = Z_i + \Delta Z \quad (5) \& (6) \]

2. Local coordinate differences (\( \Delta e/\Delta n/\Delta u \)) are:

\[
\begin{bmatrix}
\Delta e \\
\Delta n \\
\Delta u
\end{bmatrix} =
\begin{bmatrix}
-\sin \lambda & \cos \lambda & 0 \\
-\sin \phi \cos \lambda & -\sin \phi \sin \lambda & \cos \phi \cos \lambda \\
\cos \phi \cos \lambda & \cos \phi \sin \lambda & \sin \phi
\end{bmatrix}
\begin{bmatrix}
\Delta X \\
\Delta Y \\
\Delta Z
\end{bmatrix}
\] (7)

3. Geocentric coordinate differences are:

\[
\begin{bmatrix}
\Delta X \\
\Delta Y \\
\Delta Z
\end{bmatrix} =
\begin{bmatrix}
-\sin \lambda & -\sin \phi \cos \lambda & \cos \phi \cos \lambda \\
\cos \lambda & -\sin \phi \sin \lambda & \cos \phi \sin \lambda \\
0 & \cos \phi & \sin \phi
\end{bmatrix}
\begin{bmatrix}
\Delta e \\
\Delta n \\
\Delta u
\end{bmatrix}
\] (8)

4. The ground-level horizontal distance, \( HD(1) \), (Burkholder 1991) between Standpoint and Forepoint is:

\[ HD(1) = \sqrt{\Delta e^2 + \Delta n^2} \] (9)

5. The 3-D azimuth, \( \alpha \) (Burkholder 1997b) from Standpoint to Forepoint is:

\[ \alpha = \tan^{-1} \frac{\Delta e}{\Delta n} \] (10)

- Borrows equations from geodesy in which:

\[ a = \text{ellipsoid semi-major axis} = 6,378,137,000 \text{ meters (WGS84)} \]
\[ e^2 = \text{ellipsoid eccentricity squared} = 0.0066943799901 \text{ (WGS84)} \]
\[ \phi = \text{geodetic latitude (+N,-S)} \]
\[ \lambda = \text{geodetic longitude (+E,-W)} \]
\[ h = \text{ellipsoid height} \]
\[ N = \text{ellipsoid normal} = a/(1 - e^2 \sin^2 \phi)^{1/2} \]

1. Geocentric coordinates are computed from geodetic coordinates using:

\[ X = (N + h) \cos \phi \cos \lambda \] (11)
\[ Y = (N + h) \cos \phi \sin \lambda \] (12)
\[ Z = [N (1 - e^2) + h] \sin \phi \] (13)

2. Three-dimensional geodetic coordinate values are derived from geocentric ECEF coordinates:

\[ \lambda = \tan^{-1} \frac{Y}{X} \] (14)
\[ h = \frac{\sqrt{X^2 + Y^2}}{\cos \phi} - N \] (15)
\[ \tan \phi = \frac{Z}{\sqrt{X^2 + Y^2}} \left(1 + e^2 N \sin \phi \right) \] (16)

Equations (15) and (16) are solved by iteration or see Burkholder (1993) for algorithm by Vincenty.

- Treats elevation, orthometric height, as a derived quantity (see letter E in Figure 1) computed as:

\[ \text{Elevation} = \text{Ellipsoid Height} - \text{Geoid Height} \] (17)

The BURKORD™ 3-D Diagram

- True 3-D. Computations follow rules of solid geometry.
- Linear adjustment model.
- Meter length units.

Figure 1, Components of the Global Spatial Data Model
The Stochastic Model

The stochastic component of the GSDM represents an application of variance/covariance propagation laws as described in Chapter 4 by Mikhail (1976) and uses the following matrix formulation as applied to equations of the functional model:

$$\Sigma_{1y} = J_{1y} \Sigma_{xx} J_{1y}^T$$

(18)

where:

- $\Sigma_{yy}$ = Covariance matrix of computed result.
- $\Sigma_{xx}$ = Covariance matrix of variables.
- $J_{1y}$ = Jacobian matrix of partial derivatives of the result with respect to the variables.

The GSDM makes extensive use of two covariance matrices, the geocentric covariance matrix and the local covariance matrix. In each case:

1. The covariance matrix is $3 \times 3$ symmetric which means a maximum of six numbers are required to store upper (or lower) triangular values.
2. Units for each covariance matrix is meters squared.
3. Standard deviation is the square root of each diagonal element.

**Geocentric Covariance Matrix**

$$\Sigma_{x'y'z'} = \begin{bmatrix} \sigma_x^2 & \sigma_{x'y} & \sigma_{x'z'} \\ \sigma_{y'x} & \sigma_y^2 & \sigma_{y'z'} \\ \sigma_{z'x} & \sigma_{z'y} & \sigma_z^2 \end{bmatrix}$$

(19)

**Local Covariance Matrix**

$$\Sigma_{e'n'u'} = \begin{bmatrix} \sigma_e^2 & \sigma_{en} & \sigma_{eu} \\ \sigma_{en} & \sigma_n^2 & \sigma_{nu} \\ \sigma_{eu} & \sigma_{nu} & \sigma_u^2 \end{bmatrix}$$

(20)

Each covariance matrix can be obtained from the other using the following rotation matrix as shown in equations (22) and (23).

**Rotation Matrix**

$$R = \begin{bmatrix} -\sin \lambda & -\sin \phi \cos \lambda & \cos \phi \cos \lambda \\ \cos \lambda & -\sin \phi \sin \lambda & \cos \phi \sin \lambda \\ 0 & \cos \phi & \sin \phi \end{bmatrix}$$

(21)

Compute the geocentric covariance matrix from the local covariance matrix:

$$\Sigma_{x'y'z'} = R \Sigma_{e'n'u'} R^T$$

(22)

Compute the local covariance matrix from the geocentric covariance matrix:

$$\Sigma_{e'n'u'} = R^T \Sigma_{x'y'z'} R$$

(23)

Additional details regarding use of the stochastic model can be found in Appendix B of Burkholder (1997a) which contains a listing of results obtained by applying equation (18) to the functional model equations of the GSDM.

Prototype software for performing 3-D coordinate geometry and error propagation computations based upon the GSDM is called BURKORD™ and is described further at www.bright.net/~eburk/. A BURKORD™ data base stores the geocentric covariance matrix along with the $X/Y/Z$ coordinates for each point. This minimizes the electronic storage space required for a 3-D spatial data set because derived quantities such as the local covariance matrix, geodetic coordinates, and map projection (or state plane) coordinates are computed upon demand. For more details see also Burkholder (1997c,d, & 1998).

**IMPLEMENTATION**

Of many points to be made about implementing the GSDM, one is user choice. Even if the GSDM is not adopted as a computational standard, an individual user can enjoy immediate benefits based upon personal use. This is true because the GSDM is a geometrically consistent arrangement of existing equations and is compatible with existing computational practices. Furthermore, the GSDM makes it easy to get 2-D solutions from 3-D data. Understandably, the converse is not true. At the other end of the spectrum, benefits based upon corporate (or collective) use of the GSDM will accrue if and to the extent the GSDM is adopted, either formally or informally, as a computational standard and/or used as a standard for data sharing. Given the simplicity, universality, and world-wide applicability of the GSDM, the benefits can be enormous.

A second point is that the GSDM can be implemented over a wide spectrum of applications and levels of technical sophistication. At the simplest level, the GSDM can be used for 3-D coordinate geometry computations without using the stochastic model portion (all standard deviations are zero). For example, geocentric $X/Y/Z$ coordinates published by the National Geodetic Survey (NGS) for existing NAD83 control points can be input
directly. Or, if the 3-D data are given in terms of geodetic latitude/longitude/height, those coordinates can be converted to X/Y/Z values using equations (11), (12), & (13) and added to the data base. Should it be said these are NAD83 X/Y/Z coordinates? That might be appropriate if it becomes necessary to distinguish them from, say, ITRF X/Y/Z coordinates of some named epoch. See later section on "consequences."

Of course, the optional stochastic portion of the GSDM can also be used if one assigns appropriate standard deviations (in all three dimensions) to controlling points and if one includes appropriate statistics of subsequent network adjustments or other measurements added to the data base.

And, at the high-level end, since the GSDM enjoys full mathematical rigor (sans non-Euclidean considerations), it can also be a valuable tool for very sophisticated scientific applications. As a collection of consistent geometrical rules equally applicable world-wide, the GSDM can also be attached to reference frames such as the ITRF. In that case, the primary control values are not based upon published NAD83 X/Y/Z coordinates, but based upon appropriate epoch specific data published by the International Earth Rotation Service. Since the 3-D concepts are well documented in existing literature—see for example Abusali et al, (1995) and Malys/Slater (1994)—it is quite probable the GSDM is already being used, although not necessarily by that name, to monitor daily earth tide movements, long term continental drift, or to separate "noise" from "temporal migration" in data used to pinpoint the location of the earth's center of mass.

Practical implementation of the GSDM by the spatial data user community probably lies between the simple and complex extremes cited above. This paper has little, if any, bearing on methods or observations used by scientists to establish mean locations for the earth's center of mass, the CTP, or the reference meridian. But, once a reference frame such as the ITRF is defined for a given epoch and mean X/Y/Z positions are published for permanently monumented control sites such as the continuously operating reference stations (CORS), those 3-D control values and the GSDM should become the foundation for a world-wide geospatial data base.

Looking ahead into the 21st century, the following points should be considered:

- The GSDM provides a proven well defined mechanism for establishing, identifying, tracking, and evaluating 3-D spatial data accuracy component-by-component.
- Modern measurement systems collect spatial data in 3 dimensions. A 3-dimensional spatial data model which is both simple and appropriate will ultimately be adopted as a world-wide standard.
- Global implementation policies, including evaluation of the GSDM, should at all times be consistent with established scientific principles regarding the definition of a coordinate system and realization of a reference frame (Mueller 1985).
- Computation of elevation relies upon knowledge of geod height. Given past improvements in geoid modeling and anticipated refinement in the future, (Milbert 1996), it will be possible in the USA to determine elevations from ellipsoid heights and geoid heights with a accuracy approaching or exceeding that of conventional leveling. For example, see the test documented in Appendix D of Burkholder (1997a). Similar progress in geoid modeling is being made in other parts of the world.
- NGS is responsible for the National Geodetic Reference System (NGRS) in the United States and has identified plans for an overall readjustment of the high accuracy reference networks (HARNs) to be completed by January 1, 2000 (NGS 1997). Should implementation of the GSDM should be considered as part of that datum update?

**ACCURACY CONSIDERATIONS**

The accuracy of spatial data is of critical concern to many spatial data users. The stochastic portion of the GSDM provides an efficient method for handling and addressing spatial data accuracy issues. Specific points about accuracy are:

- The GSDM is 3-dimensional, strictly geometrical, and equally applicable world-wide. Unlike the conformal map projection model which distorts horizontal distances, the functional model equations of the GSDM contain no inherent approximation. It is presumed any corrections needed for gravity-based observations, e.g. Laplace corrections, are applied to the spatial measurements prior to the data being made a part of a geospatial data set. The GSDM is used to describe where things are. Geodetic science explains why they are there.
• With the GSDM, the 3-dimensional accuracy of a point is described by its covariance matrix which is stored in the 3-D data base along with the geocentric X/Y/Z coordinates.

• The accuracy of spatial data is established during measurement of one or more fundamental physical quantities. Error propagation, according to some model (implicit or explicit) is employed to determine the accuracy of an indirectly measured quantity.

• The accuracy of spatial data is verified and the appropriateness of a model is confirmed by small residuals in a least squares adjustment of redundant observations.

• Spatial data users deserve statements of accuracy (standard deviations) which are both understandable and statistically reliable. The GSDM can be used to obtain the standard deviation of a point component-by-component in either the geocentric system or the local reference frame.

• The Draft Geospatial Positioning Accuracy Standards published by the Federal Geographic Data Committee (FGDC 1997) contain definitions for the concepts of network accuracy and local accuracy for geodetic control points. The GSDM includes specific mathematical procedures by which either or both can be computed. Equation (46) (Burkholder 1997a) utilizes the covariance matrix of each stored point, along with the correlations between them, to obtain local accuracy and equation (47) yields network accuracy by assuming no correlation between the points.

CONSEQUENCES OF USING THE GSDM

Although an incomplete list, consequences of using the GSDM include:

• Three-dimensional spatial data are stored in a single data base, eliminating the need to maintain separate horizontal and vertical data base files.

• The quality of spatial data is stored along with its position. Non-metric metadata are still important, but a provable statistical measure of spatial accuracy for all points can be readily available component-by-component.

• A 3-D control point would be expected to exhibit a small standard deviation in all directions. A bench mark would be expected to exhibit a small standard deviation in the local "up" direction. A horizontal control point would exhibit small standard deviations for local east and north. A GIS point might be allowed standard deviations of 1-2 meters in all directions. With no change of computational procedure, the GSDM accommodates any/all standard deviation combinations and discriminates according to stochastic data and tolerances provided by the user. Specifically:

1. When looking at the geocentric covariance matrix of a point, it is difficult, if not impossible, to ascertain whether a point is a horizontal control point, a vertical control point, or a GIS approximate point.

2. After a geocentric covariance matrix is converted to a local covariance matrix, one can readily determine the local component standard deviations by taking the square root of the diagonal elements.

3. When computing an inverse between points, the standard deviation of local coordinate differences is determined based upon the covariance data stored for each point.

• If the correlation between points, established during the measurement process and realized in a least squares network adjustment, is also saved and used during an inverse computation, the GSDM can provide reliable local accuracy estimates of one point relative to the other. Although the GSDM includes computational procedures which accommodate both network accuracy and local accuracy, the user is still responsible for knowing the answer "with respect to what?" The cliche "garbage in - garbage out" remains applicable, even with the GSDM.

Conceptually, many spatial data users are comfortable with interpolating between known or fixed values. For surveyors, a control traverse around a tract of land provides a level of confidence related to maximum errors expected for the location of points within and/or tied to the traverse. On a different scale, the "control" may be USPLSS section corners established and surveyed by the county surveyor. Or maybe a control survey for a highway project is tied into the national datum so that all points within the project are reliably related. Examples can also be described in which the issues are 3-dimensional and the "fixed" control is earth's center of mass. Can the concepts of network accuracy and local accuracy be applied in those cases? Without necessarily answering that question, the GSDM:
1. Accepts and uses any "fixed" control provided by the user in the same manner that "network" accuracy is defined by the FGDC (1997).

2. Automatically computes "local" accuracy between points if correlation data provided by the user are available and if correlations are not assumed to be zero, for example by a programmer. Given that "local" refers to directly connected stations having correlations and that distant stations are routinely connected using GPS, this is a non-trivial issue. Does that mean all participating stations are "local" if ground network stations and GPS satellite orbits are all included in a simultaneous "birdcage" adjustment? Yes, what is considered "local" in one case may be "network" at another level of application.

- The traditional practice of defining separate horizontal and vertical geodetic datums should be revisited. If attached to the ITRF, the GSDM could be viewed as a 3-D datum in which location is defined by geocentric X/Y/Z coordinates and the uncertainty of a point is given by its covariance matrix.

In that case, valid traditional horizontal datum coordinates and vertical datum coordinates can still be derived and their standard deviations can be determined using error propagation based upon:

1. The point's geocentric covariance matrix.
2. Appropriate datum transformation equations.
3. An appropriate geoid model.

Benefits and consequences of implementing the GSDM need further evaluation as applied to considerations of updating the NAD83. Should the CORS stations be held as "fixed" and how long will it be before better ITRF values for those stations justifies an update?

MOST IMPORTANT POINT:

If the GSDM is implemented on the basis of using local coordinate differences, the underlying datum values (NAD83, CORS, or otherwise) can be updated with little impact for most spatial data users. For precise applications in which an update does matter, standard deviations of local coordinate differences in each system respectively will provide evidence of whether the difference of the differences is significant. Some of those significant differences will be due to randomness in the original data. Others will be due to either gradual changes (continental drift and subsidence) or catastrophic changes (earthquakes).

REFERENCES

Abusali, P.; B. Schutz; B. Taply; and M. Bevis 1995; "Transformation between SLR/VLBI and WGS-84 reference frames," Bulletin Geodesique, No. 69.
Burkholder, E.F. 1997e; "Definition and Description of a Global Spatial Data Model (GSDM)," U.S. Copyright Office, Washington, D.C.
Milbert, D.G. 1996; "Readme File," GEOID96 program documentation, National Geodetic Survey (NGS), Silver Spring, Md.
Petersohn, F. 1997; Personal Communication, see also web page www.gov.state.nc.us/gsdf97.