POSITIONAL DISTORTION IN GEOGRAPHIC DATA SETS AS A BARRIER TO INTEROPERATION

Richard Church
Kevin Curtin
Peter Fohl
Chris Funk
Michael Goodchild
University of California, Santa Barbara, USA
Phaedon Kyriakidis
Stanford University, USA
and
Val Noronha
Digital Geographic Research Corporation, Canada.

ABSTRACT

A simple model of positional distortion in vector databases is presented. Positions are distorted by the addition of a vector field. The model is compared to the familiar epsilon band, which is shown to seriously underestimate positional displacement. The model is calibrated against a set of street centerline databases for the Goleta, CA area, using matched points. Properties of the distortion field are determined, including the variograms of the x and y components. Using conditional simulation it is possible to generate a population of possible distorted positions, and thus to study the properties of the distortion under various forms of analysis, including propagation through GIS operations. The paper reviews several general strategies for interoperation of positionally distorted geographic data sets in the context of Intelligent Transportation Systems. One proposed strategy, the ITS Datum, is reviewed in detail. The paper concludes with a general discussion of interoperability of geographic data sets in the context of positional distortion.

1. Introduction

As spatial databases proliferate, it is becoming increasingly common to encounter situations where multiple sources of the same information exist. National mapping agencies may offer digital topographic databases of the same area at different scales, and if the spread of scales is small, the two databases may offer different versions of the same features. Databases produced by local government agencies may offer different versions of the features identified in the databases produced by national mapping agencies. And in some domains, two or more companies may offer competing databases of the same area in the open market. To a user of digitized street centerlines, the TIGER files produced by the U.S. Bureau of the Census and the U.S. Geological Survey provide one potential source in the public sector, while for most areas of the U.S. coverage is also provided by varying numbers of companies in the private sector.

Many options are open to the user in such situations. One might reasonably ask for measures of the accuracy of each database, and select the most accurate; or evaluate the options by weighing the relative costs against the relative accuracies. One might use one database as a more accurate source, as a basis for correcting other databases or for determining their accuracies. In comparing two databases it might be that one has more accurate attributes, but the other has more accurate positions. In this case one might choose to conflate the two databases, to produce a single database that has the best characteristics of both. If positions are known to be equally inaccurate in both databases, one might want to average them, on the grounds that an average of two observations is a better estimate of the truth.

Street centerline databases play major roles in the set of advanced services known as Intelligent Transportation Systems (ITS; see, for example, recent publications of ITS America, or of the Transportation Research Board of the National Research Council). Consider a scenario in which a driver has a navigation system installed in a vehicle, and is receiving information by wireless communication from a central server, consisting of dynamic updates to driving information. For example, the database in the vehicle might know about one-way streets and other semi-permanent driving conditions, but the server might provide more recent information on road construction, congestion, or accidents. Since the market for street centerline databases has many suppliers, it is possible that the database being used by the server is different from the one being used by the driver's client. The two databases will likely differ in the set of streets represented, the precise naming and other attributes of streets, the positions of streets, and the topology of their connections. All of these cause problems when the server must send unambiguous information about the locations of events and road conditions to the client.

The purpose of this paper is to explore one specific aspect of these problems, that of positional accuracy. Section Two presents a basic theory of positional accuracy in spatial databases. The theory is then applied to a sample of commercial street centerline databases from an area of Goleta, CA, in Section Three, and the positional distortions present in the data are analyzed. Section Four discusses alternative models, and their implications for interoperability between databases in the ITS context. Section Five presents other aspects of the accuracy problem for these databases related to attributes and topology. Finally, Section Six places the work in the context of proposed standards for ITS.

2. Positional Accuracy

Consider a well-defined point, such as a survey monument, whose position can be measured unambiguously on the ground. Let the true location of this point in some coordinate system be denoted by **x**, and let the position recorded in some database be **x'**. The difference between these two positions is the positional error of the point's measurement, a property both of the point and of the method of measurement. If the method of measurement is constant over a number of points or a number of measurements of the same point, we can characterize its accuracy in terms of descriptive statistics of the difference between **x** and **x'**.

Suppose (Hunter and Goodchild, 1996) that the displacement between the true location of a point and its recorded position can be represented by a vector field **e**(**x**). In other words, the apparent position (x',y') is related to the true position (x,y) by the equations $x' = x + e_x(x,y)$; $y' = y + e_y(x,y)$; where e_x, e_y are the x and y components of the vector field **e**. For example, each point might be displaced systematically in one direction by a constant amount, implying a constant **e**. The precise form of **e** defines the 'rubber-sheet' distortions of position in the plane—if we think of the map as drawn on a rubber sheet, then after distortion each point **x** in the plane will be shifted by an amount and in a direction determined by **e**(**x**). To ensure that the sheet is not 'folded'—one of the necessary conditions for topology to be preserved—it is necessary that the partial derivatives of **e** be greater than -1, that is:

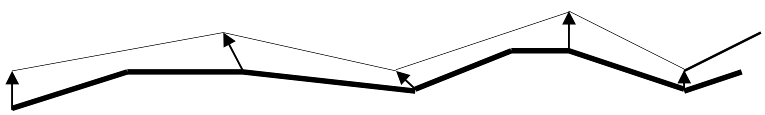
$$\partial e_x/\partial x < -1 \quad \text{and} \quad \partial e_y/\partial y < -1 \qquad (1)$$

This theory of positional accuracy clearly breaks down if the point is not well-defined; that is, if it is not possible to associate a position recorded in a database with a more accurate measurement unambiguously. For example, problems arise in assessing the accuracy of the recorded position of 'Main and First' if Main intersects more than once with First, or if it is not clear exactly where position was measured in the area of the intersection. Survey monuments resolve this problem by establishing unique identifiers for each monument, and by fixing physical monuments in the field.

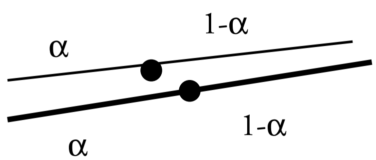
This problem of matching unambiguously occurs frequently in attempts to define the positional accuracy of lines or the boundaries of areas. The epsilon band provides a convenient and widely used basis for describing positional uncertainty of such objects. Perkal (1966, and see Chrisman, 1982) defined the band as the set of points S(ε) such that the shortest distance d(**x**) from any point **x** in the set to the line or boundary was less than some amount epsilon: $\{\mathbf{x}|\mathbf{x}\in S \text{ if } d(\mathbf{x})<\varepsilon\}$. Suppose the line or boundary is the result of some process of interpretation from source information, and is therefore subject to various forms of uncertainty, including uncertainties due to variations between interpreters, misregistration, or digitizing error. Suppose further that some minimum epsilon ε_{min} can be defined such that $S(\varepsilon_{min})$ just contains all possible versions of the line. Then ε_{min} is a useful description of the positional uncertainty of the line or boundary. Many applications of the concept have been described in the literature. Goodchild and Hunter (1997) have recently extended the concept to a probabilistic version, in order to deal with certain difficulties in the original concept, including the problems of estimating it from limited samples.

In the epsilon band formulation S is defined by distances measured perpendicular to the line or boundary. In reality, the positions of lines or boundaries can be displaced in any direction, but the direction of displacement is often unknown, and it is precisely in these situations that ε_{min} is the most effective measure.

In order to estimate **e** at any point, we need to be able to 'match points', in other words we need knowledge of the apparent and true position of a point. As Figure 1a shows, this is sometimes possible if the line or boundary contains identifiable, well-defined points, such as sharp bends, intersections, crossings of other linear features, monuments, etc. In general, however, such point matching is not possible except at a few locations. It is also possible to interpolate (Figure 1b), by assuming that in between matched points, other points can be matched according to a simple rule. Consider a point on the true line intermediate between two matched points. Let the distance from one point be a proportion α of the distance between the two matched points along the true line. Locate a point on the apparent line, the same proportion of the distance along the apparent line between the matched points. Then it is reasonable to assume that these two intermediate points also match (see, for example, Edwards, 1994a,b; Goodchild, Cova, and Ehlschlaeger, 1995). Goodchild and Hunter (1997) show, however, that such simple interpolation can often be erroneous, particularly if the specifications of the two lines are substantially different (e.g., the true line is at a more detailed scale than the apparent line).



(a)



(b)

Figure 1: Matching points to estimate displacement: (a) matching well-defined points; (b) matching interpolated points based on distance along the line.

When displacements are measured between matched points the displacement vector can be at any orientation to the line, including parallel to it. Displacement parallel to the line would occur, for example, if an east-west street were subject to a registration error that resulted in displacement to the east. Mean absolute positional error will be measured as the mean length of the displacement vector at the matched points, and RMS error will be measured as the root mean square. Let $\mathbf{e}_i$ denote the displacement vector at the *i*th matched point. Then RMS error is:

$$\bar{e} = \sqrt{\sum_i \mathbf{e}_i . \mathbf{e}_i \Big/ n} \qquad (2)$$

When no matching is possible, it is necessary to make additional assumptions in order to measure the displacement between two lines. Implicit in the epsilon band and its derivatives is the assumption that the shortest distance between a point **x** and a line is a useful measure of the point's displacement from the line. If the point could be matched, then its true displacement, or the magnitude of the vector **e**, would be at least as large as the shortest distance d(**x**). While **e** may be at any orientation to the line, the shortest distance between the point and the line will always be measured in a direction that is locally perpendicular to the line.

It follows that estimates of positional accuracy implied by the epsilon band model and its derivatives are underestimates of the true positional displacement in geographic coordinates. To evaluate the bias, we now consider four cases, in decreasing order of generality.

Case 1. Random orientation of lines and displacements

Consider first the case of a database containing lines at random orientations. For example, the lines in a database of digitized contours might be expected to occur at all orientations, such that a histogram of line orientations would show a uniform distribution. With a few notable exceptions, this seems a reasonable assumption to make about databases of naturally occurring phenomena (the oriented lakes on the North Slope of Alaska are a notable exception, as are oriented dune systems).

We assume also that displacement occurs in all directions with equal likelihood. Although there must be strong spatial autocorrelation in the displacement field e, nevertheless over a large enough area the assumption of a uniform distribution of displacement directions seems reasonable.

Without loss of generality, consider the orientation of the displacement vector with respect to a direction that is locally perpendicular to the line, as shown in Figure 2. Writing e for the magnitude of e, the projection of e onto the perpendicular is of length $e\cos\alpha$. Assuming that all displacement directions are equally likely, and limiting β to the range $\{0,\pi/2\}$ (the other three quadrants are symmetrical), the probability that a displacement lies between β and $\beta+d\beta$ of the perpendicular is given by $2/\pi\ d\beta$. The mean value of $e\cos\beta$ is obtained by integrating over the quadrant, that is:

$$\bar{e} = \frac{2}{\pi}\int_0^{\pi/2} e\cos\beta d\beta = \frac{2}{\pi}e \qquad (3)$$

If we treat e as a random variable as well as β, and assume that there is no correlation between the two variables, then e on the right hand side of the equation can be interpreted as the mean magnitude of the displacement vector. Thus the measurement of displacement perpendicular to the line, rather than in the true direction of displacement, results in an underestimation of displacement by a factor of $2/\pi$, or 0.6366. A similar analysis, integrating $\cos 2\beta$ instead of $\cos\beta$, gives a factor of $1/\sqrt{2}$ or 0.707 for the underestimation of RMSE.

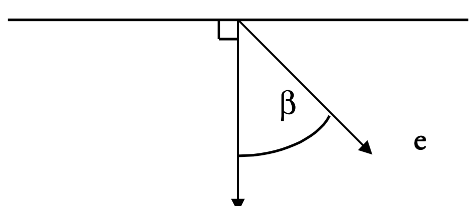


Figure 2: Relationship between the true displacement vector, line direction, and a local perpendicular.

Case 2. Random displacement direction, lines parallel to axes

Assume now that the database contains only lines parallel to the axes, in equal proportions. If displacements are oriented randomly, then all angles of orientation of displacements with respect to perpendiculars are still equally likely, and the results obtained in (1) above are still valid.

Case 3. One displacement direction, lines in random directions

The results obtained in (1) above are also valid if displacement occurs in one direction, but lines are oriented in all directions with equal likelihood.

Case 4. One displacement direction, lines parallel to axes

Finally, assume that displacement occurs in one direction, but equal proportions of lines are oriented parallel to each of the axes. The expected distance of displacement perpendicular to the line is now $1/2 \cos\beta + 1/2 \sin\beta$, which reduces to its minimum value of 1/2 when $\beta = 0$ or $\pi/2$, that is, when the displacement is parallel to one of the axes, and has its maximum value of $1/\sqrt{2}$ or 0.707 when displacement is at 45 degrees to the axes. Thus when displacement is parallel to one of the travel directions in a rectilinear street pattern, the apparent value of epsilon will be only one half the true mean displacement distance.

3. Positional Accuracy Case Study

As part of a program to test various ITS protocols and standards, the VITAL (Vehicle Intelligence Testing and Analysis Laboratory) group of the National Center for Geographic Information and Analysis at the University of California, Santa Barbara, acquired copies of six different street centerline databases for part of Goleta, CA. The databases were transformed to a common coordinate system. Each pair of databases was searched for common, unique intersections between named streets. Finally, every instance of a matched intersection was used to obtain a vector $\mathbf{e}_{AB}(\mathbf{x})$, where $\mathbf{x}$ is the location of the intersection recorded in Database A, and $\mathbf{x}+\mathbf{e}_{AB}(\mathbf{x})$ is the location recorded in Database B.

The magnitude of the displacement vector varied from less than a meter to more than one hundred meters. Strong spatial autocorrelation was observed in the distribution of displacement, as entire streets or even neighborhoods were misaligned in x or y or both (Figure 3). Straight sections of major streets appeared least subject to error, whereas sinuous roads in hilly suburban neighborhoods produced large displacements. The results appeared logical when considered in the context of how the databases are generated by map vendors, the emphasis being on major streets. Moreover, as subdivisions are developed, vendors append them piecemeal to existing databases, each section with its own displacement characteristics.

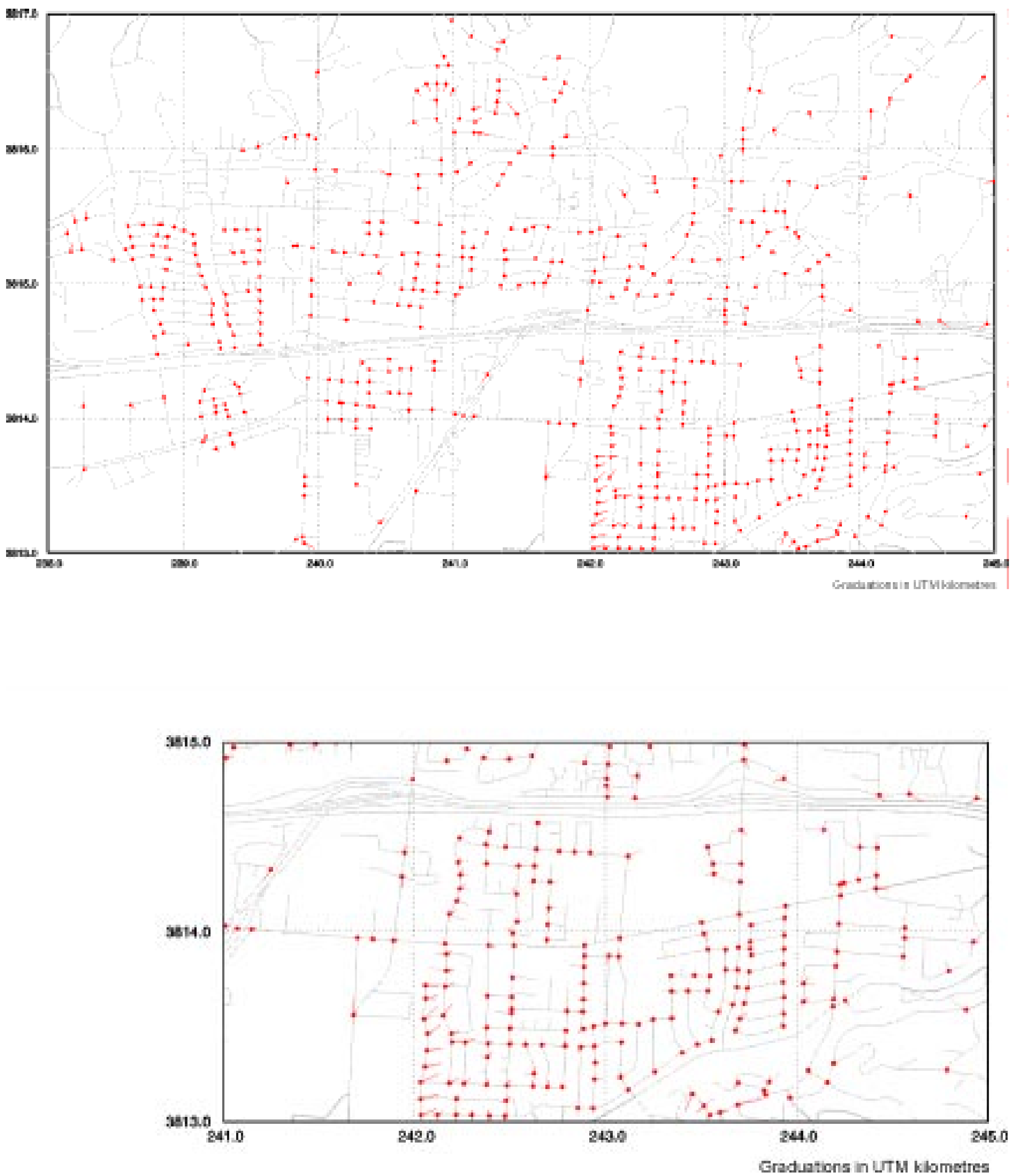


Figure 3: Distortion vectors computed by matching nodes between two commercial databases for part of Goleta, CA. The lower illustration is an enlargement of part of the upper.

4. Models of Distortion Fields

If a complete model of $\mathbf{e}_{AB}$ were available, it would be possible to correct A to match B exactly, by applying a correction vector to every coordinate pair in A. Alternatively B could be corrected to match A. In the context of ITS, any position derived from A would be unambiguously meaningful to another system using B, and interoperability of the two databases would have been achieved, in the narrow sense of unambiguous specification of location. In the case study described in Section 3 a comparatively dense sampling of $\mathbf{e}_{AB}$ was obtained, but not a complete model. In this section we discuss properties of such a model.

Kiiveri (1997, p. 34) notes that 'smoothness' is a necessary property of distortion fields. If zero-order discontinuities or 'cliffs' were allowed, then breaks would occur in any linear features intersected by such discontinuities. In general, we require $\mathbf{e}(\mathbf{x}+\boldsymbol{\delta}\mathbf{x})-\mathbf{e}(\mathbf{x})$ to tend to zero as the magnitude of $\boldsymbol{\delta}\mathbf{x}$ tends to zero. In addition to an absence of cliffs, note the conditions given earlier on the magnitude of the signed gradient of $\mathbf{e}$, to exclude the possibility of 'folding' the rubber sheet.

Kiiveri (1997) models $\mathbf{e}$ with a truncated series defined by trigonometric functions of the coordinates x and y. This results in a convenient set of boundary conditions, since $\mathbf{e}$ is required to go to zero at the edges of the (rectangular) study region. In our case, however, there seems no reason to impose this boundary condition, since the boundaries of our study region were arbitrarily defined by us, and bear no relationship to the boundaries used by the database creators. Thus we have no reason to expect distortion to go to zero everywhere on our boundary.

The literature on spatial interpolation and the modeling of surfaces is clearly divided into two schools. The first school models surfaces as functions of the spatial variables, using two convenient approaches. Any function $f(x,y)$ can be described by a polynomial in x and y, and this principle lies behind *trend surface analysis*, which estimates such a polynomial from point observations, arbitrarily limiting the number of terms in the polynomial to be much less than the number of observations, and using only the low-order terms. Similarly, any such function can be described by an appropriate *Fourier series* in trigonometric functions of x and y, and again the series is arbitrarily limited by the number of observations to the low-frequency terms. Kiiveri's approach is a variant of this second method.

Two strong arguments can be made against this first school. The use of the spatial variables as the basis of modeling often leads to spurious artifacts at the edge of the study area (Unwin, 1975), unless arbitrary boundary conditions are applied, such as those used by Kiiveri (1997). Second, the need to truncate series because of limited numbers of observations ensures that although these surfaces are in principle completely general, in practice they have the form of the low-order terms in the series, such as simple waves in the Fourier case, or simple linear or quadratic surfaces in the trend surface case. Yet it is very unusual to find theoretical justification for such very specific forms. For example, we can see no reason why the errors introduced by map-making or surveying would lead to a distortion surface that has the form of a simple linear or quadratic function of position, or a simple trigonometric function.

The second school makes no such assumptions about the form of the surface, but instead invokes other principles for surface construction. For example, the procedure known widely as *inverse-distance weighting* determines the function f at any point as a weighted average of the nearest observed values, with weights that decrease with distance. *Kriging* also uses a linear combination of nearest observed values, but with weights that are computed from the theory of regionalized variables. For a general review of interpolation methods see Lam (1983); and for reviews of geostatistics and Kriging see Isaaks and Srivastava (1989) or Deutsch and Journel (1992). We choose Kriging in this analysis because we know very little about the form of the distortion surface that results from the errors inherent in map-making methods, and because of its strong theoretical basis.

Figure 4 shows a typical distortion field semivariogram, computed from the observations resulting from node matching two of the Goleta databases. We computed variograms for both components of **e** and for all pairs of databases.

With a complete interpolated surface, it is now possible to correct all of the coordinate pairs in any one database to match any other. Since Kriging is an 'exact' interpolator, producing a surface whose values at points of observation match the observations exactly, the corrected databases will fit perfectly at all matched nodes. There will still be differences in general, however, between features in the databases that were not matched. For example, although the nodes at both ends of a curved street may have been matched, we cannot guarantee that the representations of the intermediate curves will also match. But because distortion fields must be smooth, we expect that fits will generally be good.

It was argued earlier that cliffs could not exist in a distortion field because any such cliff would introduce a sharp break in any linear feature that crossed it. Let **s** and **s**+**δs** denote two points on a linear feature a small displacement **δs** apart. As noted earlier, a cliff exists if **e**(**s**+**δs**) - **e**(**s**) tends to some finite vector **a** as the magnitude of **δs** tends to 0. Where this condition exists, a 'fold' or a 'rip' will result depending on whether the scalar product **a.δs** is less than or greater than zero, respectively. Thus it was argued earlier that such cliffs cannot exist in distortion fields.

But this argument has a major flaw, specifically, that such linear features exist everywhere, and therefore expose the existence of cliffs. In practice, even a dense street network like that shown in Figure 3 has substantial areas where no linear features exist, and where cliffs would go undetected. The set of point observations of **e** that results from matching nodes is very far from a randomly-located sample.

Suppose, for example, that a Goleta street centerline database had been assembled from a patchwork of varying quality, using various methods of registration to a uniform, geodetically-controlled base. Within each patch, distortion might be characterized by a constant **e**, or by the result of a simple affine transformation, described by a distortion field whose components e_x and e_y are linear functions of x and y. But sharp discontinuities might exist at the edges of patches, where one patch's data were edgematched to another's. This model of **e** might be expected to apply to a database constructed by georegistering subdivision plats, or to one built from a mosaic of aerial photographs. Hunter and Goodchild (1995) show that the tiling used by the Gestalt photomapper introduces similar patterns of discontinuity in digital elevation data. Figure 5 illustrates the model.

One might expect that such a pattern of distortions would be unacceptable, because of the extreme distortions and breaks it would introduce in any features crossing patch boundaries. But in reality the degree of connectivity between subdivisions in a suburban area like Goleta can be relatively low, as illustrated in Figure 3. Areas of this figure do indeed show the kinds of distortion being suggested here, of persistent systematic distortion over substantial areas, and sharp breaks elsewhere. We plan to continue this work by experimenting with patch-building algorithms, to see if they provide a better-fitting estimate of the distortion field than standard Kriging, with its assumption of smoothness.

5. Other Aspects of Inaccuracy

Positional accuracy is just one of several aspects of database disagreement that present problems for interoperability. Disagreement also arises in street names, topology, and in attributes such as address ranges. Some differences are simply variations in vendor practice, while others are due to operator interpretation, and there are errors of commission and omission. In the County of Santa Barbara, 20–40% of all centerlines have blank street names—this is true of all the databases examined in this study. Vendors differ in their treatment of freeways: some store a single centerline; others represent the two sets of lanes separately. Similarly, traffic circles and divided streets may be interpreted differently. There are numerous instances of alleys and non-navigable pathways being coded as streets. This leads to topological disagreements with other databases.

The exercise of sampling the error vector field (Figure 3) illustrates the problems that arise due to these disagreements. We needed to find street intersections in one database that could be unambiguously identified in the other database. This was difficult because of database differences: Avenue was abbreviated as 'Ave' or 'Av'; there were problems of operator interpretation (e.g. Ward Memorial Boulevard vs Clarence Ward Memorial Boulevard) and differences in spelling the same

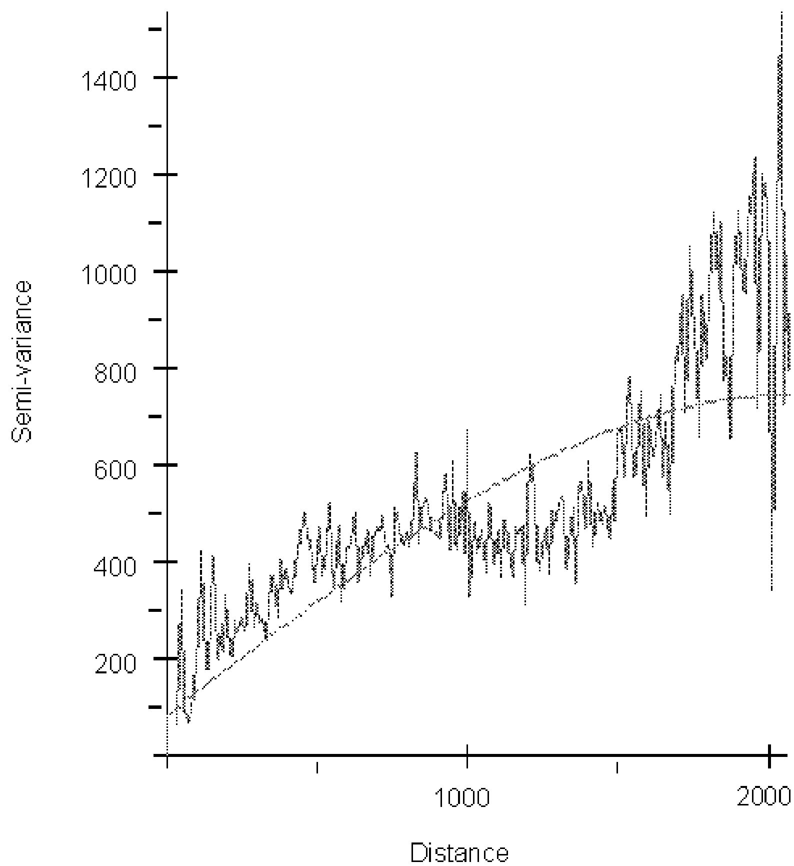


Figure 4. Example of an empirical variogram for the magnitude of the distortion field between two street centerline databases. The smooth curve shows the fitted variogram model.

name (Brien vs Brian). A routine was devised to search for containment of the street name root ("Main") of database A in the combined name fields ("Main St") of database B. This strategy produced a relatively good match rate of about 90%, although some spurious matches had to be eliminated manually. There were systematic errors in that freeways and ramps were almost never matched, due to blank data fields or differences in coding conventions (e.g., US-101 vs Hwy 101). Further work on name-based matching and other messaging methods is ongoing at VITAL.

6. Implications for ITS

In these data sets we have been able to obtain comparatively dense point estimates of distortion fields, by matching nodes. In practice, however, a user of Database B in a vehicle receiving instructions based on Database A installed at the server would know virtually nothing about their compatibility. With positional errors commonly as high as 50m, it is possible for a given coordinate pair transmitted from the server to be assigned to the wrong street by the client. Without a model of $\mathbf{e}_{AB}$, it is impossible for the client to correct the coordinate pair. Moreover, in a competitive, commercial environment it is very unlikely that all vendors of street centerline databases would agree to produce only one, 'true' database.

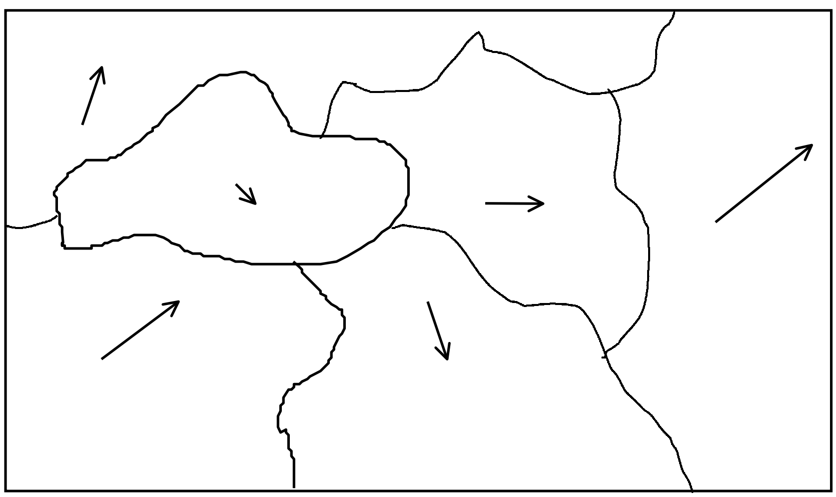


Figure 5: A 'patch' model of a distortion field, with discontinuities at patch boundaries (arrow lengths not to scale)

The ITS Datum proposal (Siegel et al., 1997) would identify a small number of common control points at public expense, on the understanding that each vendor would establish the coordinates of each control point in its own coordinate system. Each control point would be established with a physical monument, or by a set of rules for selecting reference points at interchanges, and could be surveyed as accurately as possible by any database vendor or agency. Suppose control point i is determined to be at (x_{iA},y_{iA}) by the vendor of Database A, and at (x_{iB},y_{iB}) by the vendor of Database B. In effect, one observation of the field $\mathbf{e}_{AB}$ is now established. From a series of such observations, any system would be able to estimate all of $\mathbf{e}_{AB}$, and determine the corrections to be applied to any coordinate pair from A to make it meaningful in the context of B.

One simple operational procedure would be as follows. Any coordinate pair to be communicated from A to B would be associated with an ITS Datum point, by expressing location as *x* and *y* offsets. B would then apply the same offsets to its own version of the datum point's location. In effect, this procedure forces $\mathbf{e}_{AB}$ to be modeled as a set of domains, each domain being associated with one ITS Datum point, with a constant value of $\mathbf{e}$ in each domain, while the definition of domains is left up to the server.

Alternatively, B might anticipate this situation by acquiring in advance a list of all ITS Datum points with their coordinates in both A and B databases. Then a model of $\mathbf{e}_{AB}$ could be built using some convenient method of spatial interpolation. Kriging would likely not be appropriate because of its high computational complexity, which might overwhelm a 'thin' client, and because the number of available observations is likely to be small. Instead, a simple triangulation with linear interpolation within triangles might be appropriate, although this method would inherently assume that the ITS Datum points had been carefully located at the locations of extreme high and low distortion. We intend to experiment with various procedures in the coming months, as part of VITAL's project to test ITS standards and protocols.

References

Chrisman, N.R. (1982). Methods of Spatial Analysis Based on Error in Categorical Maps. PhD thesis, University of Bristol.

Deutsch, C.V., and A.G. Journel (1992). GSLIB: Geostatistical Software Library and User’s Guide. Oxford University Press, New York.

Edwards, G. (1994a). Characterising spatial uncertainty and variability in forestry data bases. In: Proceedings, International Symposium on Spatial Accuracy of Natural Resource Data Bases, pp. 88-97.

Edwards, G. (1994b) Characterising and maintaining polygons with fuzzy boundaries in geographic information systems. In: Proceedings, 6th International Symposium on Spatial Data Handling, pp. 223-238.

Goodchild, M.F., T.J. Cova, and C.R. Ehlschlaeger (1995). Mean objects: extending the concept of central tendency to complex spatial objects in GIS. In Proceedings, GIS/LIS ’95, Nashville, TN, pp. 354-364.

Goodchild, M.F., and G.J. Hunter (1997). A simple positional accuracy measure for linear features. International Journal of Geographical Information Science, 11(3):299-306.

Hunter, G.J., and M.F. Goodchild (1995). Dealing with error in spatial databases: a simple case study. Photogrammetric Engineering and Remote Sensing 61(5):529-537.

Hunter, G.J, and M.F. Goodchild (1996). A new model for handling vector data uncertainty in geographic information systems. Journal of the Urban and Regional Information Systems Association, 8(1):51-57.

Isaaks, E.H., and R.M. Srivastava (1989) Applied Geostatistics. Oxford University Press, New York.

Kiiveri, H.T. (1997). Assessing, representing, and transmitting positional uncertainty in maps. International Journal of Geographical Information Science, 11(1):33-52.

Lam, N.S.-N. (1983). Spatial interpolation methods: a review. The American Cartographer 10(2):129-149.

Perkal, J. (1966). On the Length of Empirical Curves. Discussion Paper No. 10, Michigan Inter-University Community of Mathematical Geographers, Ann Arbor.

Siegel, D., C.W.H. Goodwin, and S.R. Gordon (1997). ITS Datum Specification, Draft-A. Federal Highway Administration, McLean, VA-USA.

Unwin, D.J. (1975). An Introduction to Trend Surface Analysis. Concepts and Techniques in Modern Geography No. 5. GeoAbstracts, Norwich, UK.