Towards ITS Map Database Interoperability — Database Error and Rectification

Val Noronha
Vehicle Intelligence and Transportation Analysis Laboratory (VITAL)
NCGIA, University of California, Santa Barbara CA 93106-4060, USA
Phone +1.805.893.8992 E-mail noronha@ncgia.ucsb.edu URL http://www.ncgia.ucsb.edu/vital

Abstract. One of the impediments to the implementation of ITS is the lack of map database interoperability. Centreline databases are available from a number of sources, but few were designed specifically for ITS. Consequently there are a variety of problems — coordinate inaccuracy, errors of omission and commission, missing or wrong street names, incorrect topology — that are compounded when communicating parties use databases from different vendors. Many ITS applications (e.g. emergency response, ATIS) rely on the exchange of messages in which location is a component. Map error and interoperability problems can result in a variety of practical difficulties from inappropriate vehicle routing to delays in delivery of critical services. These problems can be addressed by (a) standards for map databases, (b) intelligent messaging, (c) national integration efforts to improve database quality in the long term. VITAL is a laboratory funded by the U.S. federal government to develop and test these solutions. This presentation outlines the problem and describes the solutions currently under development.

Introduction

The vision of Intelligent Transportation Systems (ITS) is reminiscent of "Star Wars" in terms of the leap in the implicit level of human organization, and its ambitiousness. A few isolated components of the technology have already been demonstrated: automated toll collection (Highway 407 in Ontario, Canada), platoons of driverless vehicles cruising at highway speeds (Demo97, San Diego 1997), and systems that call emergency services when the airbags deploy (e.g. Cadillac® On-Star®). ITS calls for the real-time gathering and communication of a large amount of information, e.g. vehicle identity, vehicle and incident location, passenger counts, payload characteristics. It would not be an exaggeration to state that location is one of the pivotal issues in the technology. As we know, all maps contain some inaccuracies or lack of currency; moreover, maps from different sources differ in their level of detail, and interpretation of features. Inaccurate or inconsistent maps could result in emergencies being mis-located, and vehicles being routed down already congested arteries. As ITS will undoubtedly involve multiple system vendors, each with a slightly different version of the road network, the interoperability challenge is to ensure that users, whether motorists or dispatchers, are not negatively affected by these differences, and that a driver could cross a continent, using maps from different vendors as (s)he crosses jurisdictional boundaries, or using one vendor for national highways and local vendors for minor roads, without being unduly impacted by seams in coverage.

Origins of street databases

The map databases currently in use for ITS were developed from a variety of sources. In the U.S. and Canada, the national census bodies produced the first extensive digital street coverages in the 1970s, originally as reference aids for enumerators. The U.S. Census developed the DIME files, and Statistics Canada the Area Master File (AMF). The Canadian AMF was limited to urban areas and was heavily protected by Crown copyrights, stifling private sector development; whereas in the U.S., DIME and its successor, TIGER, were available at low cost, and spawned a number of commercial products and services, from customer “spotting” (geocoding) to commercial vehicle tracking and routing. But the periodicity of update (10 years) was too long for many applications; it soon became clear that the data were not suited to mission critical tasks such as emergency service routing; and due to database errors in street names and address ranges, success rates in address matching were in the range of 60–75% at best. These problems are not unique to the U.S.; practically every country that creates digital street maps still faces these problems.

A second stream of digital street files evolved from such sources as Thomas Brothers® in the western U.S., who had established reputations as vendors of reliable hardcopy street gazetteers and maps. These firms had typically taken some liberties in cartographic representation for the sake of esthetics and ease of interpretation, e.g. the separation of service roads from nearby freeways was exaggerated to facilitate visual distinction, and sharp elbows were smoothed. In the world of digital georeferencing, these “features” of the previous era came to be considered as liabilities.

To support the specific demands of ITS, a handful of firms, notably Etak® and Navigation Technologies®, embarked on expensive surveys of the national street network, using GPS, video-logging, photogrammetry, and coordination of local government sources. The quality specifications of these products vary widely; while some are designed for in-vehicle use, others are better suited for the somewhat less exacting needs of traffic management centres (TMCs).

ITS requirements

In ITS, the location of an object or event (moving vehicle, accident, available hotel, highway closure) must be communicated digitally, unambiguously and in real time, to suitably equipped recipients. A car with an unconscious driver makes a call to emergency services; a route closure is broadcast inaudibly to in-vehicle navigation systems that recompute routes. There may be no opportunity for human intervention, and no room for error. There are competing map and information service vendors, offering different levels and qualities of service. Yet a motorist should expect to drive across a continent using just one in-vehicle navigation system, crossing political boundaries and interacting with several information service providers (ISPs), seamlessly. Major highways from a national vendor should dovetail with city streets from a local vendor.

Fletcher (1999) identifies four levels of interplay: (a) *independent*: systems do not communicate with one another; (b) at the other end of the scale, within a single organization such as a state department of transportation, databases are *integrated* by means of enterprise planning and internally recognizable identifiers; between these two extremes, (c) *interfaces* or translators enable two disparate systems to communicate by transforming the language, and thereby the semantics, of one system so that it is acceptable to the other system; and (d) *interoperability* is an attempt to harmonize further, at the level of semantics, so that systems differ only in the details of implementation.

Standards organizations in several countries (e.g. ERTICO in Europe, VICS in Japan, Society of Automotive Engineers (SAE) in the U.S.) are working to develop standards to facilitate interoperation. The standards efforts are in many cases in advanced stages, but placeholders still exist for the representation of location.

VITAL was contracted by the U.S. Federal Highways Administration (FHWA) to test map database interoperability in the context of SAE standards, and we have done so with respect to six commercial databases for the County of Santa Barbara, California. One of these is an engineering scale product, another is the national TIGER database; the others are popular commercial databases that are widely employed for ITS applications.

Types of error

Disagreement between maps arises due to (a) differences in standards or interpretation, and (b) error. One has to be careful not to confuse error with uncertainty, scale and resolution. At a small cartographic scale, say 1:50,000, a winding mountain road is generalized, and if a product built at that scale is compared with a 1:10,000 scale product, disagreements are inevitable. It is also unfair to compare a $50,000 product built to 1:1000 engineering specifications, with a $1,000 product designed for use in market demographics. The following discussion is not intended to be critical of vendors, but merely to point out that problems exist, that pose interoperability challenges.

Scale, existence, classification and topology

Any digital representation of geographic reality entails some degree of interpretation. A divided urban road may be usefully represented as a single centerline for those funding maintenance and planning services, whereas those *delivering* emergency service needs to know the barriers to accessibility, and a detailed dual-line representation with turning points is necessary. Similarly traffic circles and channelized turn lanes are not required for all applications, and are represented in some maps and not in others. This is really a matter of scale, resolution and cost — at small mapping scales, even large freeway interchanges are reduced to points.

Since there is no single standard for road classification, vendors differ considerably on how they classify roads. One vendor uses a handful of classes: freeway, other highway, urban arterial, residential; another has more than 40 classes, distinguishing for example between roads that share a bed with railway tracks, and those that do not. Consequently it is impossible to find agreement even on a map of "major" streets between two vendors. Since vendors have different interpretations of importance of a road, there are substantial differences in inclusion and exclusion. Driveways to major buildings appear in one map and not in the other. Finally, vendors are on different time-tables; some use the national census product (TIGER in the U.S.) as a starting point, giving them immediate national coverage; others build from the ground up, and their coverage may be sparse in some areas. New neighbourhoods and changes to the network are not immediately reflected in all products.

Disagreement in both existence and classification result in topological discrepancies: a single link in one map may be fragmented in another map as it is interrupted by minor roads and driveways.

Geometric alignment

Due to the varied lineage of digital maps, some being surveyed by GPS and photogrammetry, others evolving from hardcopy products at various scales, there are significant differences in their geometric representation of features. While it is true that "ground truth" itself is evasive — centrelines and other road features wander with widening and periodic striping — one could hardly characterize a positional discrepancy of 200m on a 10m-wide road as being due to uncertainty (Figure 1).

Geometric accuracy has gained importance recently due to the low-cost availability of GPS. For ITS purposes, GPS is currently the only practical, universal and cost-effective way to determine the location of a vehicle. To express a location in any terms, say to determine a street address where a vehicle is parked, the GPS coordinates must be snapped to the nearest segment of the street network, and a location expression is derived by processing the network data.

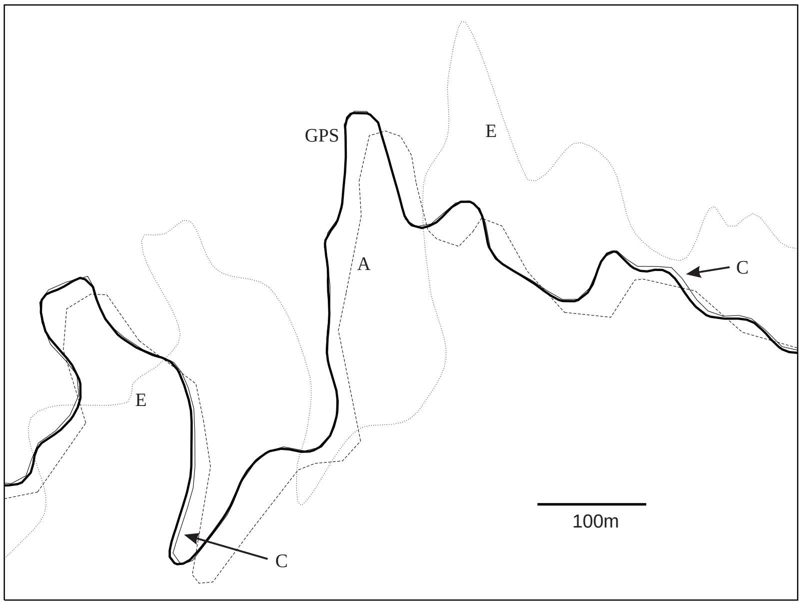

Figure 1. Overlay of three maps, A, C and E, with our GPS-survey. The engineering database C is almost obscured by our GPS trace, while databases A and E exhibit considerable positional error.

When the Selective Availability (S/A) error of GPS is compounded with map database error, locations reported by GPS can be grossly inaccurate. VITAL (1998, 1999) has shown, based on current maps of the County of Santa Barbara, that there is a 50% chance that a GPS-derived coordinate will snap to the wrong road (with 100m S/A error). This can easily result in an emergency being reported on the wrong ramp of a freeway interchange, or the wrong block of an urban development. But with an accurate reference map and differential GPS, the likelihood of accurate identification is virtually 100%.

Recently there have been efforts to characterize and to model error (Goodchild and Hunter 1997; Church et al 1998, Funk et al 1998), so as to develop methods to correct it. From the ITS point of view, a correction algorithm has to run in real-time, perhaps in a relatively small vehicle-based processor. This problem is revisited later in this paper, in the discussion of datums.

Street names and addresses

When ITS was faced with the problems of map and GPS inaccuracy, it was thought that street names may be an appropriate way to express and to communicate a location. However, there are considerable problems with street names too. In the County of Santa Barbara, between 20% and 46% of database records have blank name fields — this accounts for 50–60% of the length of the road network. Many of the blank roads are private ranch roads, and driveways to gated communities, and it could be argued that they are of little consequence for ITS. On the other hand, freeway ramps are prominent among unnamed roads in urban areas, and they are critically important for ITS incident reporting.

A second problem is that not all vendors have adopted standards on street type abbreviation (“Ave” vs “Av”; although the U.S. Postal Service has a national standard) or parsing of a name. While the fields for “Main Avenue” are obvious, treatment rules for “Avenue of the Americas,” “Avenida Redondo” and “State Route 45” are not so clear. Only one vendor uses a comprehensive and logical breakdown of components: direction prefix, type prefix, proper name, type suffix, direction suffix. Another vendor is alone in storing an alias field (Highway 401, also Macdonald Cartier Freeway), hence disagreements arise when one vendor quotes the name of a road, and another vendor uses the highway number. Table 1 highlights a few of the problems.

Table 1. Five stretches of road, as named in 5 databases. A road may have several names over the sampled distance; databases may not agree on where changes occur. Only A provides for an alias.

A	B	C	D	E
E CAMINO CIELO	CAMINO CIELO		no record	E CM CIELO
SAN MARCOS PASS HWY 154	HIGHWAY 154		SAN MARCOS PASS CA-154	SAN MARCOS PASS blank
MOUNTAIN E MOUNTAIN	MOUNTAIN PARK BELLA VISTA		MOUNTAIN	W MOUNTAIN E MOUNTAIN PARK BELLA VISTA
FOOTHILL CATHEDRAL OAKS	CATHEDRAL OAKS	HWY 192	CATHEDRAL OAKS FOOTHILL	CATHEDRAL OAKS FOOTHILL

The challenges of name matching are not trivial. Interest in the matter dates back to the turn of the century — the popular Soundex method of phonetic matching was patented in 1904 — and the general problem of lexical matching is addressed in the computer science literature (e.g. Knuth 1973). Software solutions for street name matching have been developed by Statistics Canada (Fellegi & Sunter 1969) and the U.S. Bureau of the Census (Jaro 1989), and there are now commercial products (e.g. Matchware) that perform fuzzy matches on addresses. The software must be geography-specific to deal with local variations. Currently such software is not widely used; it will have to be deployed extensively to make ITS messaging successful.

We have not tested accuracy of street address data — this would involve intensive field work. Two points deserve particular mention. First, civic addresses do not exist along highways, hence ITS has limited interest in them, although the more general problem of emergency service delivery, which relies on ITS, is heavily dependent on accurate processing of addresses. Second, because most databases store address ranges rather than individual address points, there is potential for error on long roads where the allocation of addresses is non-linear with respect to distance. One sometimes encounters situations where the address assignment method is non-standard, or two sides of a street lie in different municipalities and follow different addressing systems.

Length from coordinates

Another method to express location is by means of a linear offset from an agreed starting point. GIS-T is heavily reliant on linear referencing, and locations of sign posts and other highway furniture are stored and communicated in linear terms. VITAL (1999) conducted extensive tests of linear messaging, with mixed results. On the one hand we found excellent agreement between linear measures derived from a vehicle-based distance measuring instrument (DMI; precision 1m), cumulative measurements on the engineering database, and our own D-GPS survey. On a 7.5 km sinuous mountain road (Figure 1), the DMI and engineering measurements were within 5m of each other, and the D-GPS trace within 30m of the others. But on the other hand, even the most reputable ITS databases under-reported distance by 100–200m because of coordinate generalization, and positional errors were apparent. In general, DMI and D-GPS measurements were within 0.1–0.5% of the engineering measurement, whereas the other databases varied by 5–15%.

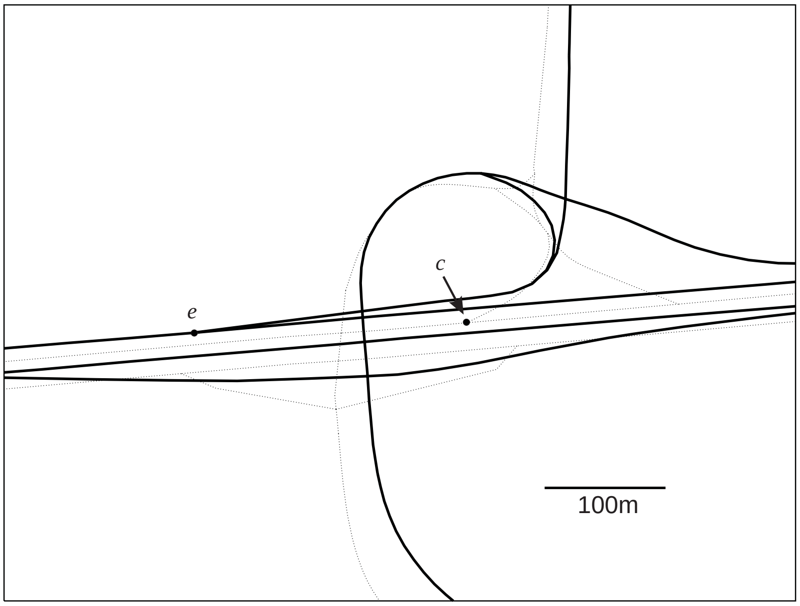

Figure 2. Overlay of maps C and E, showing different positions (c and e) where ramp meets freeway. The impact on freeway segment length is more than 200m

Due to alignment error and generalization in databases, there is on average a 50m error (800m in the worst cases) in translating a 2-dimensional coordinate into a linear reference. Due to disagreement in linear measures, there is a further 50m error incurred in communicating that location to a linear reference in a different database — in other words, the offset of a fixed point on the ground could be 800m with respect to one database, and 850m with respect to another. Paradoxically, both these errors are worse on freeways (100–200m) than on city streets, because of inaccurate representation of ramps (Figure 2). Fortunately tolerances on freeways also tend to be greater.

Other errors

Reference was made above to topological inconsistencies resulting from differences in inclusion. In addition there are factual errors: roads are shown to intersect when they do not, and vice versa, and connectivity may be generalized so that two intersections are blended into one. Attribute errors (e.g. number of lanes, speed limit) are known to exist in all databases. Vendors may attach attributes based on general notions of street characteristics, e.g. a street of moderate sinuosity passing through a residential neighbourhood would normally have a speed limit of 50 km/h, hence the value is assigned without field verification. There are always instances of human error: a bike path described as a major artery, several streets with the same name, spelling error, etc. These will always exist, although appropriate data structures can minimize the likelihood of some of these types of errors.

In general, database quality is geography-dependent. Vendors in the U.S. are preparing for a burgeoning ITS market, and have focussed their attention on major cities and metropolitan regions. While positional errors in Santa Barbara are in the 30–100m range, they are in the 1–5m range in the San Francisco area — more in the league of uncertainty than error. Similarly there are differences between rural and urban areas, winding and straight roads.

Remediation strategies

Awareness of map database error has been growing over the last decade, and both the ITS and GIS communities are now working towards solutions.

Messaging

In the U.S., the Society of Automotive Engineers (SAE) is the standards development organization that has taken the lead in ITS messaging standardization. A number of message specifications exist, to poll a vehicle for its location, and to return location information data to an enquirer. Until recently, the authors of these specifications had left placeholders to represent location, uncertain about the volume and type of data required to communicate location efficiently and unambiguously. The Location Referencing Message Specification (LRMS) was developed by Oak Ridge National Laboratory, and recently abstracted into SAE standard J2374. The LRMS provides profiles to communicate location in different ways, by coordinates, cross street names, linear references, street addresses, grids, etc. Each profile is designed with a particular purpose or user group in mind, e.g. the geometry profile communicates generic points, lines or polygons; cross streets are used for most ISP-to-vehicle messaging; and linear references are designed to be used primarily between agencies that have a standard identification scheme for network nodes and links. VITAL was contracted to test two LRMS profiles, the Cross Streets profile and the Linear Referencing profile. In the case of the Cross Streets profile, success rates were extremely low, in the 5–25% range, largely due to the number of blank data fields. Moreover, we pointed out that a matching triad of street names might be found anywhere in the world, therefore the profile should be extended by including a coordinate pair, to avoid gross error. This is a general principle for all forms of messaging, and in particular for mission-critical situations. Redundancy, feedback and iterative convergence assure accuracy of communication — this is the basis of check digits and parity bits in electrical signal processing and digital packet communications. VITAL is currently developing this area of research, towards “LX-100,” a near-100% reliable location expression and exchange profile for mission critical applications.

Map database standards

An elaborate messaging technique such as LX-100 can certainly *improve* upon the quality of messaging, but cannot guarantee near-100% reliability unless it can make assumptions about the quality of inputs. Databases must meet some basic quality standards in geometric alignment, street naming (including parsing and abbreviation), and preferably classification, inclusion and topology. It should then be possible to certify whether a given database complies with the standard, and if it does, the burden of error compensation by intelligent messaging and other methods is diminished, and near-100% success can be assured.

Datums

A datum is a frame of reference, in relation to which measurements are taken, and with which we relate measurements to reality. The global graticule of latitude and longitude is a datum, and there are standard horizontal and vertical datums such as NAD83, WGS84 and NAVD88, by means of which coordinate measurements are tied to the earth.

Currently there are two major transportation datum initiatives in the U.S., the 3-dimensional ITS Datum from the ITS georeferencing community, and a purely linear datum that arose from GIS-T requirements. A third associated initiative is from the National Spatial Data Infrastructure (NSDI). All these efforts are ultimately aimed at smoothing the process of data exchange in GIS, GIS-T and ITS.

Figure 3. VITAL's rubber-streeting algorithm applies a geometric adjustment to an erroneous map to make it agree with a more accurately surveyed map or datum. Useful attributes associated with the geometrically poor map are retained, or easily transferred to the other map.

NSDI

The NSDI objective (FGDC 1998) is to facilitate broad data sharing between public sector agencies, by establishing identifiers on unique field-recoverable locations, such as street intersections and the roads that link them. The identifiers act as external keys, and every participating data custodian is responsible for tying its internal identification scheme to this national standard. This can be laborious work, but the NSDI argument is that it need be performed only once by each agency.

Linear Datum for GIS-T

The linear datum initiative (Vonderohe et al 1995) is to create a system of anchor points, with anchor sections linking them. Accurate linear measurements are made along anchor sections, and these are the basis for correcting other linear measurements taken along the roads covered by that section. Anchor point locations are determined by an optimization procedure that seeks to minimize linear error (i.e. error in distance measurement) only, hence anchor sections cover roads in linear sequences, there is no recognition of topological relationships, and intersections are no more significant than any other point on the network.

ITS Datum

The ITS Datum proposal (Siegel et al 1996) calls for a network of datum nodes and links, forming a topological network that is a subset of the road network. A datum prototype was constructed in 1996, consisting of 50,000 points across the U.S., essentially limited to major highways. Because of the small scale source of highway maps used, the prototype datum was specified with a single point to represent a complex interchange, at ±80m accuracy — this is clearly insufficient for ITS purposes. To solve the location referencing problems described earlier in this paper (e.g. to distinguish between ramps in an interchange), the datum would have to be densified, with a datum point probably at each exit, usually 8 per interchange, with accuracy of about ±5m. It would also have to be extended to cover at least urban arteries.

Clearly the greater the number of datum points, the greater the cost of creation and maintenance. The current plan is to optimize density with regard to 2-dimensional error (contrast this with the linear datum which minimizes 1-dimensional error), taking into account the costs. The benefits of the datum are accuracy improvements. Clearly accuracy at the millimetre level is excessive and probably spurious, and not worth the cost. The challenge is to find the density and accuracy that reduces error to the point of best justifying the costs. VITAL has developed a rubber-streeting™ algorithm (Figure 3) to adjust maps geometrically in real-time, using a selection of datum points; this method will probably be used as the basis for determining the appropriate density of the ITS datum, and for datum-based corrections.

Towards a single multipurpose datum?

Proponents of the three efforts discussed above have been meeting regularly in recent months, to assist in developing each others' plans, and to compare and contrast the objectives and methods. It is clear that there are fundamental differences in technical goals, but to the extent that all three involve (a) field survey, (b) identification rules for points and lines, and (c) administrative arrangements for custodianship and maintenance, there is hope that the cross-fertilization could lead to coordination of expensive work components and thereby reduced costs, if not conceptual unification.

Lessons for the future

How do the above developments impact transportation agencies planning for the future? Our tests in the County of Santa Barbara indicate that (a) with good centreline databases and differential GPS, location referencing and communication can be virtually error-free, and (b) serious errors are possible with uncorrected GPS (±100m) and practically all commercial databases: in general one can expect 50–75% success rates in positioning and communication, but in some areas commercial data vendors have expended the effort to improve their databases.

These findings clearly indicate that local agencies faced with poor commercial maps of their area can and should move immediately to create databases based on differential GPS or photogrammetry to meet immediate needs, but providing hooks in data models so as to exchange data with other agencies that may have better data, or attributes of interest. A good quality base map can be the basis for value-added cooperative agreements with other agencies, and as datums and other formal data exchange mechanisms become commonplace, quality can evolve to the more demanding standards of the future.

Acknowledgement

Statements in this paper are based partly on research funded by the U.S. Federal Highways Administration and the California Department of Transportation.

References

Church R, Curtin K, Fohl P, Funk C, Goodchild M, Kyriakidis P, Noronha V 1998. Positional distortion in geographic data sets as a barrier to interoperation. Proceedings, ACSM Baltimore [document available at www.ncgia.ucsb.edu/vital]

Fellegi IP, Sunter A 1969. A Theory for Record Linkage. Journal of the American Statistical Association, 64 (328):1183-1210

FGDC 1998. Content Standard for Digital Geospatial Metadata. United States Geological Survey, Reston VA. Federal Geographic Data Committee, Metadata Ad Hoc Working Group. FGDC-STD-001-1998.

Fletcher D 1999. Road data model workshop, Washington DC. Meeting Notes.

Funk C, Curtin K, Goodchild M, Montello D, Noronha V 1998. Formulation and test of a model of positional distortion fields. Third International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Quebec City [URL www.ncgia.ucsb.edu/vital]

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

Jaro M 1989. Advances in Records Linkage Methodology as Applied to Matching the 1985 Census of Tampa. Journal of the American Statistical Association 84 (406): 414-420

Knuth DE 1973. The Art of Computer Programming, Volume 3: Sorting and Searching. Addition Wesley.

Noronha V, Goodchild M, Church R, Fohl P 1999. Location Expression Standards for ITS — Testing the LRMS Cross Streets Profile. Annals of Regional Science, Special Issue on GIS Sharing and Standardization, in press.

Siegel D, Goodwin C, Gordon SR 1996. ITS Datum Final Design Report. United States Department of Transportation, FHWA Contract 61-94-Y-00001, Review Draft, June 28, 1996

VITAL 1998. The Cross Streets LRMS Profile with Coordinates — Technical Evaluation. United States Department of Transportation, FHWA Contract DTFH61-91-Y-30066, Final Report.

VITAL 1999. The Linear Referencing Profile — Technical Evaluation. United States Department of Transportation, FHWA Contract DTFH61-91-Y-30066, Final Report.

Vonderohe AP, Chou CL, Sun F, Adams T 1995. Results of a Workshop on a Generic Data Model for Linear Referencing Systems. Proceedings, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks NV

Many points made in this paper are elaborated in reports which are summarized on our web site, http://www.ncgia.ucsb.edu/vital, and additional illustrations may be found there.