INT. J. GEOGRAPHICAL INFORMATION SCIENCE, 1997, VOL. 11, NO. 3, 299–306

Technical Communication

A simple positional accuracy measure for linear features

MICHAEL F. GOODCHILD
National Center for Geographic Information & Analysis, 3510 Phelps Hall, University of California, Santa Barbara, CA 93106, U.S.A.
e-mail: good@ncgia.ucsb.edu

and GARY J. HUNTER
Center for GIS & Modeling, Department of Geomatics, The University of Melbourne, Parkville, VIC 3052, Australia
e-mail: gary_hunter@mac.unimelb.edu.au

(*Received 16 June 1995; accepted 29 February 1996*)

Abstract. In this paper we propose a simple technique for assessing the positional accuracy of digitized linear features. The approach relies on a comparison with a representation of higher accuracy, and estimates the percentage of the total length of the low accuracy representation that is within a specified distance of the high accuracy representation. The approach deals successfully with three deficiencies of other methods: it is statistically based; is relatively insensitive to extreme outliers; and does not require matching of points between representations. It can be implemented using standard functions and a standard scripting language in any raster or vector GIS. We present the results of a test using data from the Digital Chart of the World.

1. Introduction

The positional accuracy of a spatial object, or a digital representation of a feature, can be defined through measures of the difference between the apparent location of the feature as recorded in the database, and the true location. Unfortunately, abundant difficulties exist in identifying the true location of a feature because of problems with measuring instruments, frames of geodetic reference, and feature definitions. Instead, positional accuracy of a feature's digital representation in a database may need to be defined in practice through measures of the difference between the location recorded in the database, and a location determined with higher accuracy. We term the first representation the *tested source* and the second representation the *reference source*. If the accuracy of the reference source is sufficiently high, one can ignore the (unmeasured) difference between it and the truth, and treat the measure of accuracy as a property of the tested source only.

In the case of point features, the distance between the tested and reference source provides a convenient basis for measures of accuracy. Suitable measures include the root mean square distance, and percentiles of the distance distribution. The 90th and 95th percentiles are often used as the basis of map accuracy standards and as measures of accuracy for measurement instruments such as the Global Positioning System (GPS), and are readily interpreted without substantial understanding of statistics. If the distribution of observed positions around the true position is normal

1365-8816/97 $12·00

or Gaussian, simple mathematical relationships exist between the percentiles of the distance distribution and other parameters such as the root mean square distance and the standard error.

When features are lines, areas, or volumes the comparison of tested and reference sources presents substantial problems. The epsilon band (Perkal 1966) has been discussed frequently in the literature (Blakemore 1984, Aspinall and Pearson 1995), and is implemented in various GIS algorithms in the form of a tolerance. Epsilon is often interpreted in a deterministic sense as the minimum buffer width around the reference object that contains all of the tested object, or *vice versa*. However, this definition is very sensitive to outliers. Since outliers are likely to be rare, we can expect epsilon to rise as the sampled length of the feature rises, and thus to be dependent on decisions that may have little to do with the process of data production. For example, epsilon estimated in this way for the coastline of the entire U.S. is likely to be larger than epsilon estimated for any sample part of the U.S. coastline, even though the coastline may have been produced using a uniform method. This dependence of the metric on outliers and sample size makes it less than ideally robust, and less than ideal as a quality control measure.

Another possible approach is to compare representations through a series of point comparisons. If points on the tested source can be matched to points on the reference source, then measurement of accuracy reduces to the well-defined point case. But while certain types of features such as rectilinear road networks may contain sufficiently well-defined points, such as intersections, the approach is much more difficult to carry out with other types of features such as coastlines, especially since comparisons must be made across large differences in scale, and therefore in feature generalization. Moreover, it is possible that the processes of data production lead to different distributions of error for well-defined features, in which case such measures may be biased. For example, road intersection locations may have been obtained from a more accurate source than the links between them, or may have been subject to different processes of generalization. If matching is possible and there is no reason to expect bias, then statistical parameters of the resulting distance distribution will provide robust metrics of accuracy.

Where no matching of points is possible, it is necessary to resort to some other metric of the separation between the tested and reference sources. One might sample points along the tested source and in each case measure the distance to the closest point on the reference source. Similarly one might sample along the reference source and measure to the closest point on the tested source. Both of these approaches give a sample of distances from which distributions can be obtained, and statistics measured such as the mean or the percentiles. One can also combine the two approaches to obtain a single average. But these methods are all subject to the same basic objection—we do not know how the shortest distance is related to the distortion of individual points, which can be measured only if points can be matched. Certainly the true distortion is always at least as large as the shortest distance.

The method we propose requires that a reference source be available, and that it contain a complete representation of the feature. Skidmore and Turner (1992) discuss a somewhat different situation, common in forestry, where reference information is available only in the field, and where the cost of field observation is high. In such cases it is feasible to collect only a sample of the reference source, in their case using field transects; the separation between the tested location of the feature and the reference location is determined at each intersection between a transect and the tested source. The number of transects determines the confidence one has in the estimate of separation—the more transects, the greater the confidence.

The objectives of Skidmore and Turner (1992) are different from ours in a second respect. The mapping of forest boundaries is normally subject to some positional uncertainty imposed by the mapping technology, for example, one might know that a positional uncertainty of 10 m was introduced by the act of digitizing from a paper map. In addition, mapped forest boundaries will vary from their true positions in the field because of observer error, and this effect will frequently be greater than 10 m. It is desirable to know what proportion of mapped boundaries are within the limits of uncertainty imposed by the technology, and what proportion contain additional uncertainty. Skidmore and Turner assume that the first form of uncertainty is known, and focus on measuring the second. We see our proposed method as directed at the more fundamental problem of measuring positional uncertainty in the most general circumstance where nothing is known.

The assumptions of standard error theory lead to the expectation that the distribution of differences between observed measurements and their true values should be Gaussian. Although this principle may extend to points, and is assumed in many approaches to the measurement of positional uncertainty in points, no comparable theory exists for the case of measures of separation between complex geographical features. The method we propose is non-parametric, which seems appropriate until adequate theory is developed, or perhaps until extensive empirical results demonstrate that the Gaussian distribution is endemic in these circumstances, although this seems unlikely.

2. Method

Figure 1 shows an example of a reference source (the 'true' coastline) and a tested source. Consider a buffer of width x around the reference source object. We compute the proportion p of the tested source length that lies within the buffer. The function $p(x)$ is expected to be robust, in the sense that it is relatively insensitive to outliers

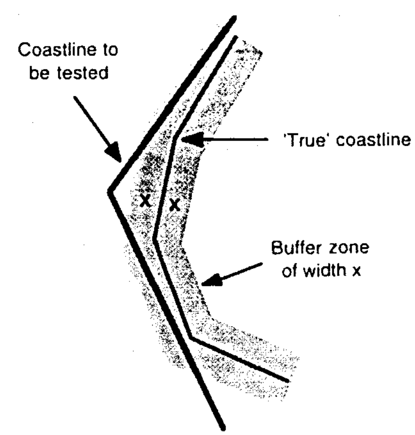

Figure 1. A buffer of width x around the 'true' coastline boundary is intersected with the boundary to be tested, to determine the percentage of coastline lying within the buffer polygon.

and rare events, unlikely to depend systematically on the length of feature sampled, and invariant under rotation, translation, and rescaling of the coordinate system. If we think of $p(x)$ as a cumulative probability distribution, with $p(0)=0$ and $p(\infty)=1$, then values of x can be thought of as corresponding to percentiles of the distribution. For example, we could identify the 95th percentile as the distance within which 95 per cent of the length of the tested source lies. This should be readily interpretable.

The measure depends on measurement of the length of the tested object. In vector representations, length is a function of the digitizing process, and this might be raised as an objection as it makes the measure sensitive to factors other than positional accuracy. But since both the numerator and denominator of p depend on this same factor, we expect their effects to tend to cancel out. Similarly, in raster systems where length is measured by counting cells, both numerator and denominator will be approximately equally sensitive to cell size. Thus, more careful digitizing of the same test source, or change of cell size, should not affect the measure directly.

By analogy to accepted measures of positional accuracy for points, we expect that most applications will require the determination of x for a given value of p. Let y represent the given value, perhaps 90 or 95 per cent. Then we require a procedure for determining x such that $p(x)=y$. If we could assume a model for p, such as the Gaussian distribution, then we could use its inverse function to determine x. However, it seems unnecessary to make this assumption when a simple iterative procedure will suffice. We propose the following procedure:

1. Set $x_o=0$ and $p_o=0$, and determine the target percentile y. Select an initial value x_1, based on available indicators of positional accuracy such as test source scale, or map accuracy standards. Set $i=1$.
2. Construct a buffer of width equal to x_i around the reference object; intersect this buffer with the tested object; and compute the proportion of the tested object enclosed within the buffer as p_i. If $|p_i - y| < 0{\cdot}001$ then stop.
3. Compute a new estimate of x based on a linear approximation to the function p:

$$x_{i+1} = \frac{(y - p_{i-1})(x_i - x_{i-1})}{(p_i - p_{i-1})} + x_{i-1} \tag{1}$$

4. Set $i=i+1$, and go to (2).

Since the function $p(x)$ must be monotonically increasing, there will be exactly one solution to $p(x)=y$, $0<y<1$, and the iterative search should be well-behaved, although any algorithm to search for a crossing of a function could be used. Step (2) requires successive estimates to be within 0·1 per cent of each other for the process to terminate.

3. Case study and discussion of results

To test the method in practice, a segment of coastline from the Digital Chart of the World data set was selected (figure 2 (*a*)). The segment chosen depicts the southern coastline of Victoria, Australia, around Port Phillip Bay and the city of Melbourne. The data quality statement accompanying the DCW data (ESRI 1993, pp. 2–10) notes that '... horizontal accuracy, at a 90 per cent confidence level for circular error, ranges from 1600 feet (488 m) to 7300 feet (2225 m)'. For some tiles in the data set, DCW also contains local positional accuracy statements as supplied by the source mapping agencies, but these are absent for this coastline segment.

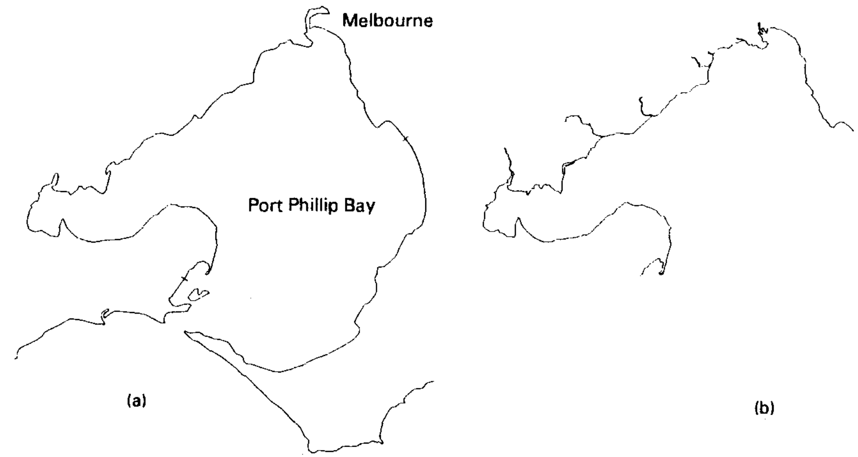

Figure 2. In (*a*) the coastline to be tested comes from the Digital Chart of the World, and represents part of the southern coast of Victoria, Australia. The segment of coastline extracted for comparison lies between the two short line markers. In (*b*) the 'true' coastline has been digitized from local 1:25 000 topographic mapping.

For purposes of comparison, a reference source was created by careful digitizing of local topographic maps compiled at 1:25 000 scale (figure 2 (*b*)). A coordinate transformation was applied to the DCW data to bring both sources into the same projection and datum. The endpoints of the reference source were matched to the DCW, which was then clipped at the positions marked in figure 2 (*a*). The reference source had a total length of 247 km, compared to a test source length of 179 km, the difference being due largely to the omission of several large rivers and estuaries from the DCW.

The results are shown in table 1 and figure 3. In both cases we have evaluated $p(x)$ over a range of discrete values; in practice, we suggest the automated search described above be used to determine x given y. The 90th percentile is approximately 330 m, which is consistent with the data quality statement quoted earlier, although the latter was presumably verified by comparison of well-defined points. Table 1 and figure 3 also show a computed Gaussian distribution for comparison, as estimated from the 68th percentile ($z=1$ for a one-tailed Gaussian distribution at a cumulative probability of 0·6826). The standard deviation is 189·2, and the fit is very close with the exception of the extreme tail.

The test case clearly illustrates the problems associated with estimating the 100th percentile noted earlier. For example, an increase to $x=1000$ leaves 1·9 per cent of the tested source still outside the buffer, due to the prominent inlet shown in figure 2 (*a*) near Melbourne at the top of the data set appearing on the generalized tested source but not on the detailed reference source. The reason for this omission is unknown, and indeed it would have been expected to have been shown in greater detail on the reference source than the tested source. While it is considered by the authors to be an anomaly in the reference source, it nevertheless exists in practice and cannot be ignored.

Clearly, the approach described here focuses on information loss due to positional error, rather than to other cartographic processes. For example, figure 2 illustrates

Table 1. Case study results showing the length and percentage of coastline inside the buffer in terms of its width. The Gaussian distribution is also shown for comparison.

Buffer width (m)	Length of coastline within buffer (km)	% of coastline within buffer	Gaussian distribution
20	17·2	9·6	8·4
40	32·0	17·9	16·7
60	49·0	27·4	24·9
80	62·5	34·9	32·8
100	75·7	42·3	40·3
125	91·5	51·1	49·1
150	105·7	59·0	57·2
200	126·7	70·8	70·9
250	145·4	81·2	81·4
300	156·8	87·6	88·7
350	164·8	92·1	93·6
400	169·8	94·9	96·6
450	172·2	96·2	98·3
500	173·0	96·6	99·2
550	173·3	96·8	99·6
600	173·5	96·9	99·8
700	173·9	97·1	99·98
800	174·3	97·4	
900	174·7	97·6	
1000	175·6	98·1	

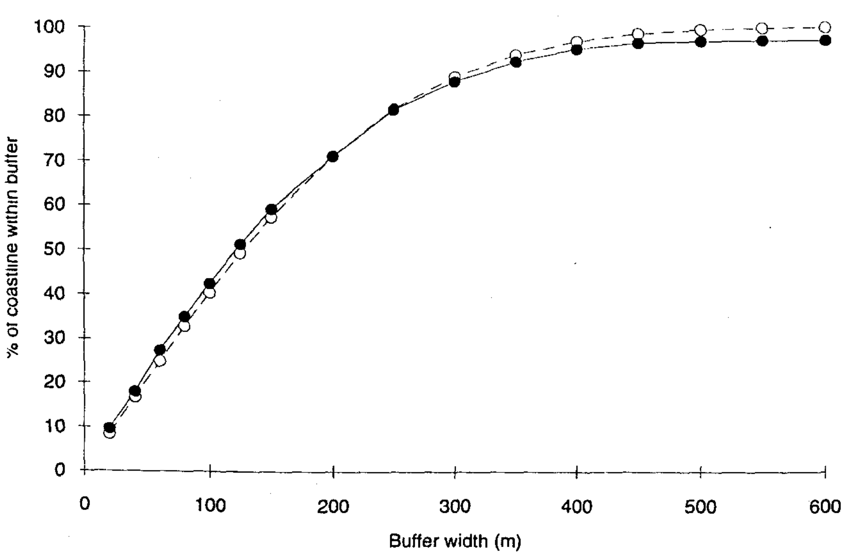

Figure 3. Plot of the percentage of coastline lying within the buffer versus the buffer width (solid line connecting black circles). For comparison the Gaussian distribution is also shown (dashed line connecting white circles).

the effects of cartographic generalization, in the omission of the large river and estuary features of figure 2(*b*) from the relatively coarse DCW representation of figure 2(*a*). These features do not affect the length of the tested source, nor the proposed measures of positional accuracy, because their omission is a reflection of cartographic generalization rather than error. On the other hand, the large inlet which appears in the coarse figure 2(*a*) but not in the detailed figure 2(*b*), will affect the proposed measure, as its absence in the detailed representation is more consistent with the notion of error than generalization. We suggest that a similar approach based on the percentage of the reference source length rather than the tested source might be more related to generalization, although we acknowledge that the two concepts are not as simple to separate as this comment might suggest.

The approach also lends itself well to various forms of visualization. Figure 4 shows those parts of the DCW feature lying inside buffers of widths 40 m (17·9 per cent), 100 m (42·3 per cent), and 500 m (96·9 per cent); displays such as this might be useful in quality control and in educating spatial data users in the meaning of map accuracy statements (Paradise and Beard 1994). Indeed, it is suggested that the written statement and its visual counterpart should be shown together so that the

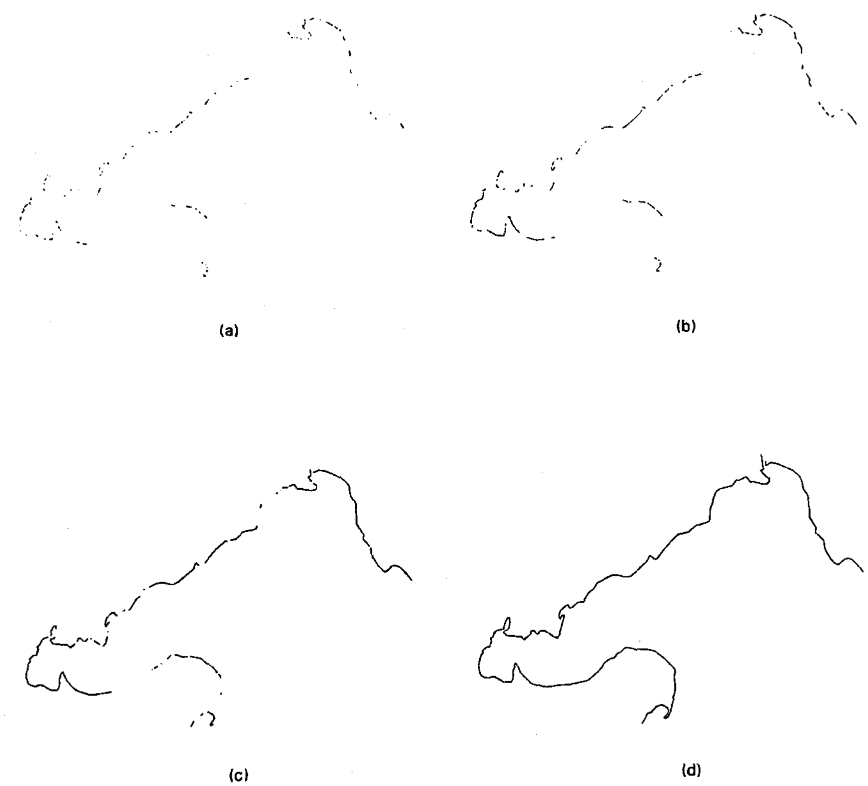

Figure 4. Illustrations of the portions of coastline lying inside the buffer for widths of (*a*) 40 m (17·9 per cent), (*b*) 100 m (42·3 per cent), (*c*) 200 m (70·8 per cent), and (*d*) 500 m (96·6 per cent). Line breaks depict sections of the coastline lying outside the buffer.

consequences of the statement may be more clearly understood. For instance, by being able to visualize what it means to have only 17·9 per cent of the map feature lying within 40 m of the true position, a decision-maker may well decide that the dataset is not suitable for the intended application—hence map accuracy statements can assist users to make decisions about data quality and fitness for use, which is their ultimate goal.

4. Conclusion

A simple method has been presented for obtaining useful descriptions of positional accuracy for linear features such as coastlines which are usually of extreme length and difficult or uneconomical to field check. It is considered that the method can be applied to other forms of linear features and generalized to area and volume features, however this has not been confirmed in practice. The procedure compares a tested source to a reference source by computing the percentage of the length of the tested source lying within given distances of the reference source. The approach is robust, being comparatively insensitive to the nature of the digital representation of the feature, and is easy to implement in a GIS.

Acknowledgments

The National Center for Geographic Information and Analysis is supported by the U.S. National Science Foundation under Cooperative Agreement SBR 88-10917. The authors wish to acknowledge the assistance of Colin Arrowsmith in providing comparative coastline boundary data which was collected in conjunction with the Port Phillip Environmental Study, conducted by the CSIRO Institute of Natural Resources and Environment, Melbourne, Australia.

References

Aspinall, R. J. and Pearson, D. M., 1995, Describing and managing uncertainty of categorial maps in GIS. In *Innovations in GIS 2*, edited by P. Fisher (London: Taylor & Francis), pp. 71–84.

Blakemore, M., 1984, Generalization and error in spatial databases. *Cartographica*, **21**, 131–139.

ESRI, 1993, *The Digital Chart of the World for Use with ARC/INFO: Data Dictionary* (Redlands, CA: Environmental Systems Research Institute).

Paradise, J. and Beard, M. K., 1994, Visualization of spatial data quality for the decision-maker: a data quality filter. *Journal of the Urban and Regional Information Systems Association*, **6**, 25–34.

Perkal, J., 1966, *On the length of empirical curves*. Discussion Paper No. 10 (Ann Arbor: Michigan Inter-University Community of Mathematical Geographers).

Skidmore, A. L. and Turner, B. J., 1992, Map accuracy assessment using line intersect sampling. *Photogrammetric Engineering and Remote Sensing*, **58**, 1453–1457.