New Tools For Handling Spatial Data Quality: Moving from Academic Concepts to Practical Reality

Dr. Gary J. Hunter

Abstract: *While the growth in spatial data quality research has increased dramatically during the past decade in response to user needs, many of the concepts developed have yet to come to fruition in the form of usable tools that can be readily applied in practice using commercial GIS packages. As such, the purpose of this paper is to report on some of the tools becoming available for implementation by users of spatial data that have been developed by the author and his colleagues over the past few years. The examples presented cover: (1) the tracking of feature coordinate edits and their reporting in visual data quality statements; (2) testing and reporting the positional accuracy of linear features of unknown lineage; (3) simulating uncertainty in products derived from Digital Elevation Models; (4) incorporating uncertainty modeling in vector point, line and polygon files, and (5) reporting data quality information at different levels of database structure. In several of these cases, public domain software already exists on the Internet for application by GIS users, which serves to demonstrate that data quality research is indeed moving from academic concepts to practical reality.*

Introduction

For several years now the author and many other researchers worldwide have been engaged in investigating the topic of data quality and uncertainty in spatial databases. Indeed, the subject was listed as a key research initiative by both the National Center for Geographic Information and Analysis (NCGIA) in the late-1980s and the University Consortium for Geographic Information Science (UCGIS) in the mid-1990s. To date, the research has spanned a wide variety of sub-topics ranging from error modeling and communication (or portrayal) through to liability issues and responses, and using risk management procedures to help deal with spatial data uncertainty in decision-making.

However, while data quality has long been considered a 'hot' topic at GIS conferences and good attendance is always assured at such sessions, the user community continues to ask when the results of our academic research will finally 'bear fruit' in the form of usable tools that can be applied in everyday practice. In response to that question there are two things that need to happen, viz.: (1) current research findings need to be transformed into practical applications, and (2) those tools and techniques need to be further developed so that they can be used in conjunction with current commercial GIS software.

Dr. Gary J. Hunter *is a Senior Lecturer and Deputy-Head of the Department of Geomatics, and his research specialties are modeling and communicating uncertainty in spatial data and decision-making under conditions of uncertainty. He has presented papers on this and other topics at URISA conferences for the past 11 years and has had his work previously published in the URISA Journal. In conjunction with Professor Michael Goodchild, he has twice previously been awarded URISA's Horwood Critique Prize.*

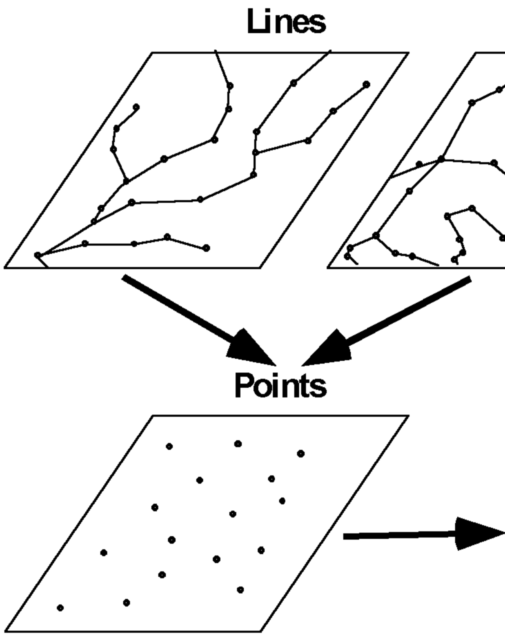
While the latter problem will require closer liaison between researchers and system developers over the longer term, in the shorter term the research community must show that its concepts, algorithms and methodologies are in fact capable of being converted into useful tools, and that they are not simply esoteric exercises in higher mathematics and statistics. Accordingly, the purpose of this paper is to report on some of the developments in spatial data quality that have been made by the author and his research team over the past few years. The examples discussed are:

- the tracking of feature coordinate edits and their reporting in the form of visual data quality statements;
- testing and reporting the positional accuracy of linear features of unknown lineage;
- simulating output uncertainty in products derived from Digital Elevation Models;
- uncertainty modeling in vector point, line and polygon databases; and
- reporting data quality information at different levels of database structure.

In several of these cases, public domain software already exists on the Internet for application by GIS users, which serves to demonstrate that the data quality research agenda is indeed moving from academic concepts to practical reality.

Example #1: Visual Reporting of Feature Coordinate Edits

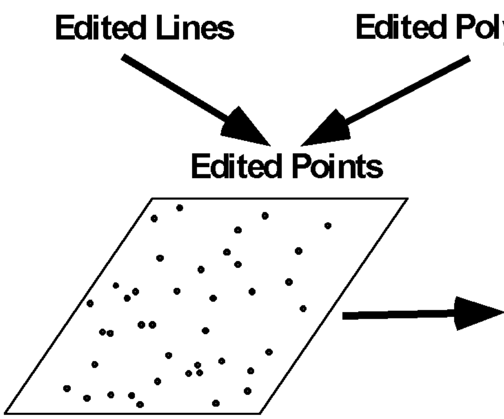
With the widespread introduction of spatial data transfer standards in the 1990s, a key feature of many of them has been the requirement for data quality reporting of components such as lineage, positional accuracy, attribute accuracy, logical consistency,



STAGE 1

Line and polygon feature datasets are decomposed to point files and new attributes are added.

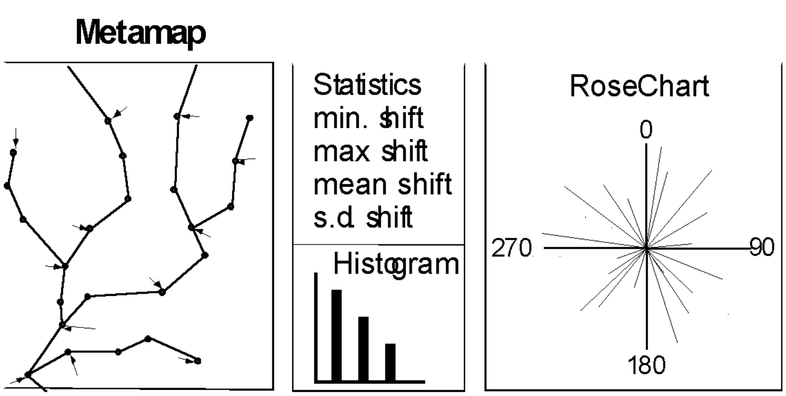
ID	START-X	START-Y	X-COORD	Y-COORD	SHIFT	BRG	STATUS
1	100.6	113.1	100.6	113.1	0.0	0.0	1
2	200.3	310.1	200.3	310.1	0.0	0.0	1
3	156.8	188.0	156.8	188.0	0.0	0.0	1
4	177.8	167.1	177.8	167.1	0.0	0.0	1
-	-	-	-	-	-	-	-



STAGE 2

Edits are made to the point, line and polygon datasets, and the coordinate attributes are updated.

ID	START-X	START-Y	X-COORD	Y-COORD	SHIFT	BRG	STATUS
1	100.6	113.1	130.7	127.1	33.2	65.1	1
2	200.3	310.1	200.3	310.1	0.0	0.0	1
3	177.8	167.1	177.8	167.1	0.0	0.0	0
4	156.8	188.0	171.8	224.2	39.2	22.6	1
-	-	-	-	-	-	-	-



STAGE 3

A 'metamap' is constructed showing all features and their shifts, summary statistics and graphs. Stages 2 and 3 are repeated whenever required to show an updated version of the metamap to document the positional changes that have occured to the dataset.

Figure 1. A method for visually reporting feature coordinate edits.

completeness and currency. Focussing specifically on positional accuracy, although the lineage component is intended to faithfully document the history of a data set and the processes that have led to its creation, there are often few details supplied regarding the various feature coordinate edits that may have been applied along the way—other than written notations, for instance, that digitized map sheets have been edge matched or certain data themes generalized. While such global reporting remains useful, it is argued that lineage reporting has yet to be widely undertaken at the individual feature level and it is suggested that the absence of more detailed information, particularly in a digital form capable of being quantified and visualized, deprives users of valuable metadata that could help them to better assess the potential usefulness of a data set.

Dealing solely with vector data in this example, the typical forms of coordinate edits that add, delete or shift features include: correction of line undershoots and overshoots; reshaping; realignment; rotation; simplification; smoothing; displacement; spurious polygon removal; rubber sheeting and edge matching. Because several of these coordinate adjustment methods are (1) unpredictable in their outcome, (2) highly localized and (3) sometimes irreversible, it is suggested here that the only option for effectively reporting them is to record and store the changes at the feature level as they occur so that they can later be recalled and communicated to users.

In current research being conducted by the author, a method of documenting, measuring and displaying the results of feature coordinate edits is being developed which could provide a useful supplement to data lineage documentation. Importantly, visualization of the spatial variation arising from these adjustments has the potential to enable users to gain detailed local, rather than global, information on the coordinate shifts that have occurred in a data set. This digital visual aid to interpreting data quality is termed a 'metamap'.

In developing a methodology that reports coordinate edits, there are obvious parallels with the challenges faced in designing and implementing a temporal component in spatial databases. However while there is a need to record and display feature changes, it is not considered to be necessary to be able to chronologically progress from one edit version to the next. Instead, the key data to be supplied to users are the cumulative coordinate shifts that have occurred to features since the time when they were first added to the data set. This means that only the original plus current coordinates need be stored—not the intermediate ones—which simplifies the problem. The process developed to achieve this task is shown in Figure 1.

The method is based around the fact that line and polygon files are essentially formed from point features which possess x and y coordinates. By storing the original coordinates for each point and then comparing them with the current set of coordinates at any time thereafter, the magnitude and direction of the cumulative positional shift that has been induced in each point through editing processes can be computed and displayed. The visualization can be portrayed in several ways such as in the form of a 'metamap', histograms showing the distribution of shift magnitudes throughout the data set, and shift statistics and rose charts.

At this stage of the research a simple macro language script has been developed for use in conjunction with Arc/Info point files such that visual reports, as in Stage 3 of Figure 1, are automatically generated on demand by the user. Further development of the method is still required to cater for line and polygon files, since the underlying point coordinates are usually not made explicitly available to users of these files in many commercial GIS. However, one option is to store coordinate shift information against line and polygon nodes which in some packages are permitted to be assigned separate attributes. Another proposed enhancement to the method is the retention of deleted points so that users can see which points were originally held in the file but have since been removed, usually after generalization operations.

Example #2: measuring Positional Accuracy of Linear Features

In the second example, a simple positional accuracy measure for linear features is presented as described in Goodchild and Hunter (1997). 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 a database and its true location. Unfortunately, several difficulties exist in identifying the true location of a feature because of problems with measuring instruments and feature definitions. Instead, the positional accuracy of a feature's digital representation 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. If the accuracy of the reference data set (accepted as being the 'truth') is sufficiently high, then we can ignore the (unmeasured) difference between it and the truth, and treat the measure of accuracy as a property of the tested data set only.

The method estimates the percentage of the total length of the low accuracy representation that lies within a specified distance of the higher accuracy representation. The technique deals successfully with three deficiencies of other approaches, viz.: (1) it is statistically based; (2) it is relatively insensitive to outlying values; and (3) it does not require matching of points between the two representations. Furthermore, it can be implemented using basic operations and scripting languages present in most vector GIS. Figure 2 shows an example of a reference data set (the 'true' coastline) and a data set to be tested. A buffer of width x is placed around the reference data set and the proportion of the length of tested data set that lies within the buffer is computed by a simple overlay process. In this way a statement to the effect that "x% of the tested feature lies within y meters of its true position" can be derived.

To test the method in practice, a segment of coastline from the Digital Chart of the World (DCW) data set was selected for which there was no local positional accuracy information. The scale of the source material from which the digital coastline was derived was believed to be 1:1,000,000 or smaller, and for the purpose of comparison a reference data set was created by digitizing local topographic maps compiled at a scale of 1:25,000. The reference data set had a total length of 247km, compared to the test data set length of 179km—the difference being due primarily to the omission of several large rivers and estuaries from the smaller-scale DCW file.

A macro language script was prepared which operates on Arc/Info linear feature files and, before the calculations begin, prompts the user for the names of the source and test data sets, and either the buffer widths to be applied or the percentage lengths to be computed. The interim output is a set of linear feature files that contain the individual portions of the linear feature to be tested which lie within the specified buffer widths or represent the required percentiles of the tested line. The approach lends itself to various forms of visualization, and Figure 3 shows the results of a simple Visual Basic software application designed to communicate the results of the positional accuracy testing. Users can select a buffer width and receive a statement of the percentage of the tested line that lies within that distance from its true position (shown in the upper left window). In addition, the large screen on the right hand side provides a graphic of how the overlaid segments of the tested data set within the buffer (in red) compare with the reference source (yellow). The graph on the left hand side of the screen shows the variation in buffer width versus the percentage of tested line lying within the buffer zone.

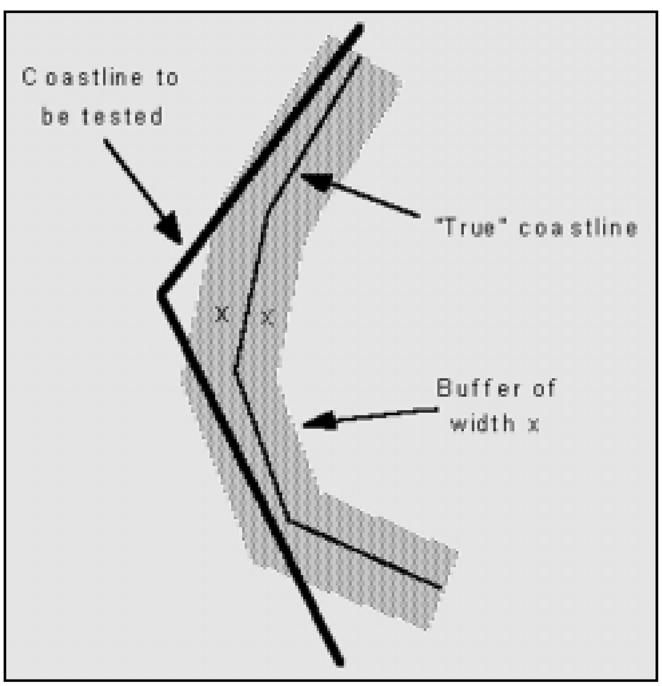


Figure 2. A buffer of width x is placed around the 'true' coastline boundary, and intersected with the boundary to be tested to determine the percentage of coastline length lying within the buffer.

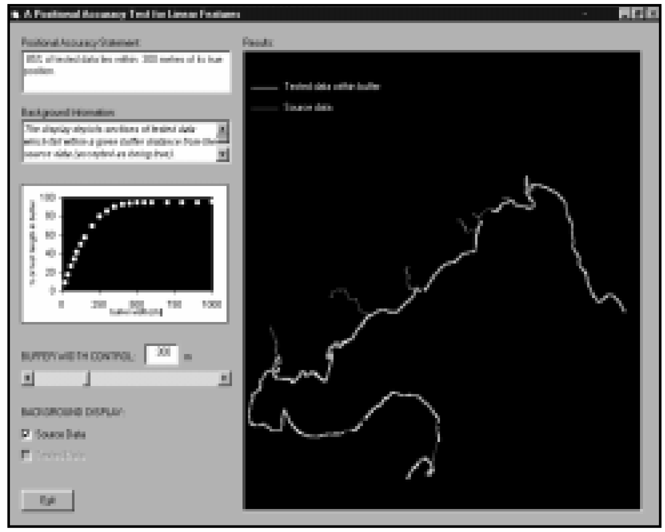


Figure 3. An example of the output of a Visual Basic application program designed to communicate to users the results of the positional accuracy testing of linear features.

Example #3: Simulating Uncertainty in DEM Outputs

In this section, three case studies are presented in representing uncertainty in the outputs derived from Digital Elevation Models. The examples relate to (1) representation of elevation uncertainty in a DEM, (2) the uncertainty in a LANDSAT image following normalization procedures to account for terrain effects, and (3) the uncertainty associated with calculating landslide susceptibility.

Representing Elevation Uncertainty in a DEM

In the paper by Hunter and Goodchild (1995a), a simple method of representing the elevation uncertainty in a DEM was illustrated. The method involved taking the Root Mean Square Error (RMSE) provided by the DEM producer (in this case the USGS), and applying it to derive a probability overlay for the DEM for a given elevation value. An example of the research is presented in Figure 4 below, in which Figure 4(a) shows the test site measuring 10km by 13 km with contours at a 20m interval. The elevation of interest is the 350m value and an initial attempt at showing where this contour would lie in the DEM is given in Figure 4(b), where reclassification of the DEM into values above and below 350m yields a portrayal of the contour but says nothing about its uncertainty. On the other hand, the RMSE statistic can be used to derive grid-cell probabilities given our knowledge of the Normal distribution of elevation error in the DEM. For example, in Figure 4(c) suppose the test site was to be flooded by a new reservoir to a level of 350m. Cells at the brown end of the color scale are those with an elevation value lying 3 x RMSE above 350m and thus have about a 99% chance of possessing a true value in

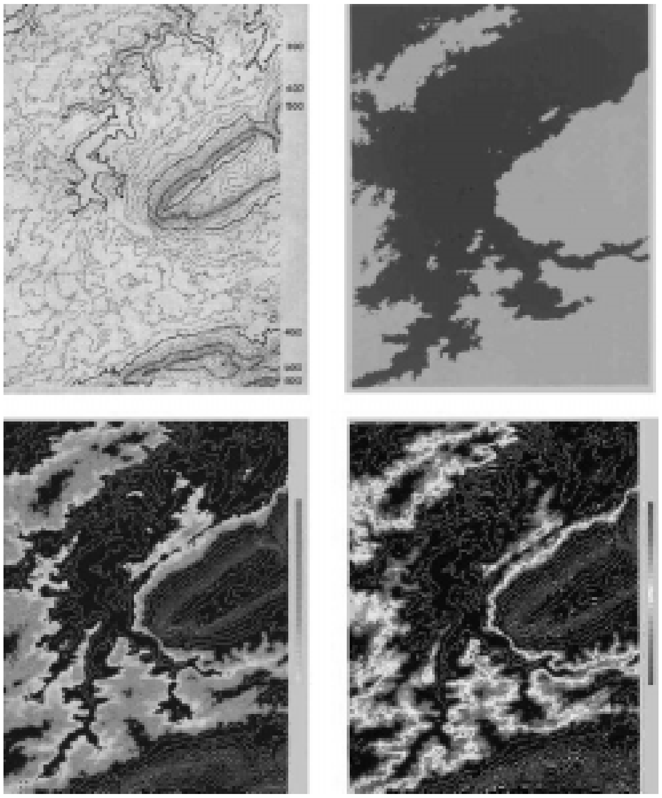


Figure 4. (a) top left, the test site contours at 20m interval, (b) top right, raster definition of the 350m ‘contour’, (c) bottom left, a probability map for cells exceeding 350m, and (d) bottom right, a probability map for cells equaling 350m.

Figure 5a

Figure 5b

Figure 5c

Figure 5. In (a) the LANDSAT image of the test site shows shadowed areas and those burnt by fires colored white and gray in the middle part of the image. In (b) a hill-shaded image of the DEM illustrates the rugged nature of the terrain. The image in (c) represents the uncertainty in the adjusted LANDSAT image as a result of the error in both the DEM and that which has been propagated during the spatial processes. Cells colored mid-gray have remained relatively stable after a series of DEM error simulations, while those colored white denote cells which have large variation in their computed values—in this case due to shadowing.

excess of 350m and not being inundated, while those at the opposite (blue) end of the scale have a value 3 x RSME less than 350m and thus only a 1% likelihood of exceeding 350m or a 99% chance of being inundated. Alternatively, in Figure 4(d), cells with the highest probability (50%) of equaling 350m are shown in white, while those with only a 1% chance of equaling 350m are at the (Red) ends of the color scale.

Uncertainty Due to Normalization for Terrain Effects in a Landsat Image

In the second example (described in Hunter and Goodchild, 1995b and 1997), LANDSAT imagery taken of Central Portugal (Figure 5a) was used by NCGIA researchers to detect the presence of land damaged by forest fires, after having first separated out those cells which are in deep shadow and may cause error in the analysis. This terrain normalization involves taking a DEM (Figure 5b) of the test area and calculating the slope gradient and aspect for each cell, which were in turn used to compute a normalized intensity value for each corresponding pixel in the image. Usually the process would be performed only once, but the task in this instance was to estimate and visualize the resultant uncertainty in the adjusted pixel values as a result of elevation error in the DEM used, and also of any error that may have been propagated during the computational processes applied.

A series of simulated error surfaces were computed and added one at a time to the original DEM to produce a family of different, but equally probable, DEMs which were then applied to create a series of adjusted image files. By studying the set of output files a mean pixel intensity value and its corresponding standard deviation were computed for each pixel in the image. Figure 5(c) shows the resulting uncertainty map which portrays pixels with small standard deviations (and are therefore relatively stable and unsusceptible to error) in gray color, while the white pixels are those that exhibited large variation in their values during error simulation (and are therefore considered unstable). In this case the unstable pixels are those on the western slopes of north-south ridgelines lying in heavy shadow as a result of the low solar elevation at the time the image was recorded. From a user’s perspective, if the area of interest is heavily affected by such unstable pixels then it is not appropriate to work with such data and better imagery should be obtained.

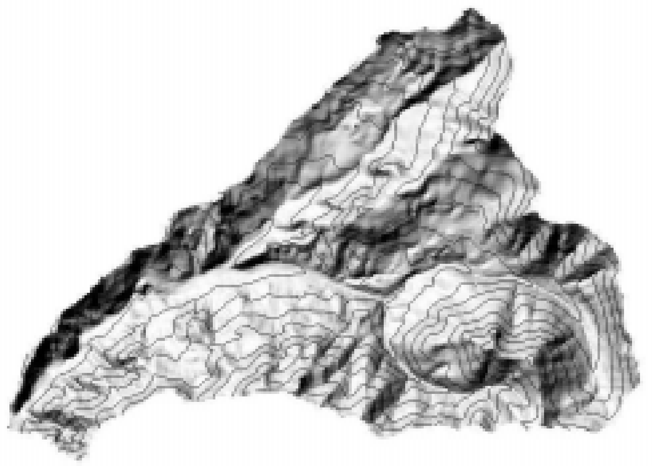


Figure 6. Showing a hill-shaded DEM of the landslide susceptibility test site in Oregon.

Uncertainty in Calculating Landslide Hazard Areas

The final example in this section deals with assessing the uncertainty associated with using DEMs when calculating landslide hazard areas in Oregon (Murillo and Hunter, 1997). The usual approach would be to calculate the slope gradient of the cells in the DEM and overlay this information with geological data to determine the steepest slopes / weakest soils combinations that would indicate high susceptibility to landslides. In the case study undertaken, the simulation approach previously described was again adopted to derive a family of perturbed DEMs which were used to generate a set of output grids in which cells were classified as either having a low susceptibility (a value of zero) or high susceptibility (a value of one).

The study area in Oregon (Figure 6) is a watershed in the H.J. Andrews Long-Term Ecological Research site in which experiments were underway to determine the effects of logging and road construction in areas that had a long history of being susceptible to landslide hazards. After 50 simulations, the output grids were added together and the values for each cell were accumulated to give a 'score' out of 50. Some cells proved to be extremely susceptible to landslides each time and scored the maximum 50, while others had zero score at the end of the set of simulations. The remaining cells scored between these limits. Figure 7 shows the outcome of the uncertainty assessment and cells that scored a high susceptibility classification at least 41 times out of 50 (that is, greater than 80% of the time) are shown in red color. From this visualization of the uncertainty in the landslide assessment model due to error in the DEM employed, users are able to adopt a more probabilistic approach to the contentious issue of whether or not to permit road construction and logging to occur in a region. For instance, rather than just employing the GIS to create a single map of landslide hazards, a set of simulations such as these could be applied so that some measure of data uncertainty is introduced to the rating model and a consensus might be reached by the parties concerned, for example, such that grid-cells which have at least an x% probability of being prone to landslides should be excluded from logging or road making.

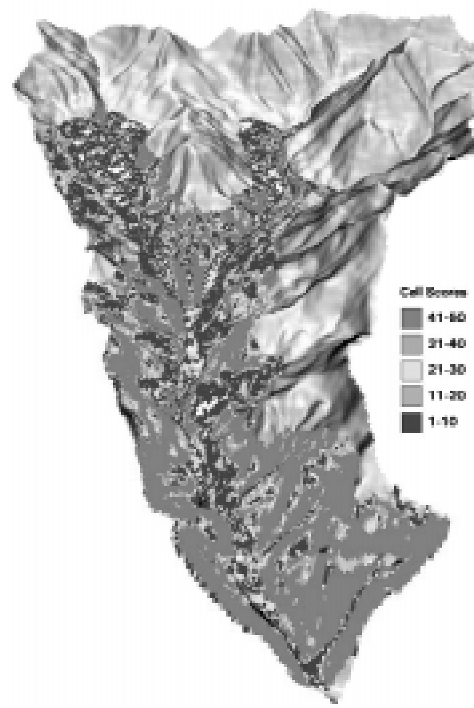


Figure 7. Looking north-east along the watershed, a landslide susceptibility uncertainty map is overlaid on a hill-shaded DEM. Cells colored red scored a 'high' landslide rating at least of 80% of the time during simulation trials.

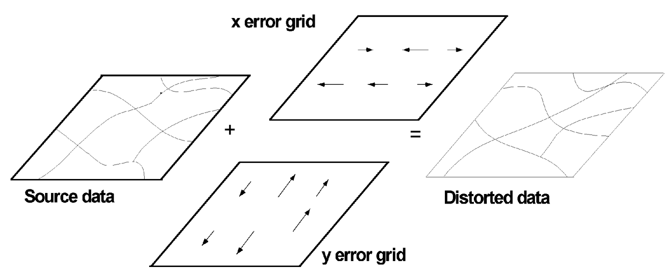


Figure 8. The vector uncertainty model uses random error grids in the x and y directions to produce a distorted, but equally probable, version of the original data set.

Example #4: Uncertainty Modeling in Vector Databases

In addition to grid-cell data, the author's research has also investigated uncertainty modeling in vector databases (Hunter and

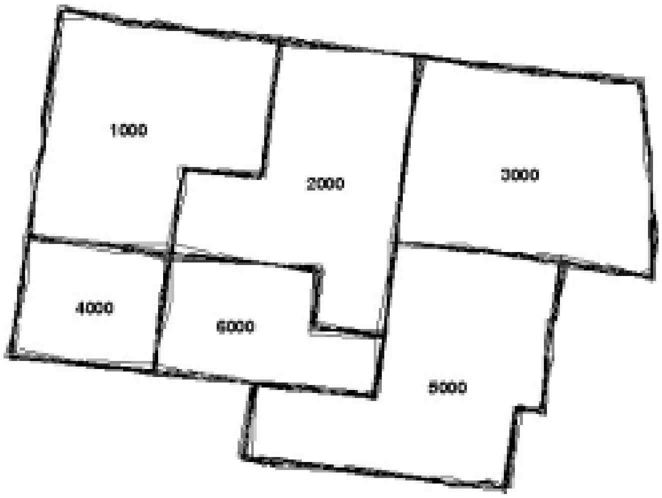


Figure 9. Application of the vector uncertainty model, showing the overlaid results of perturbing a set of six polygons 20 times.

Polygon ID	Mean Area (sq. m)	Standard Deviation (sq. m)
1000	891858.3	5419.6
2000	890108.5	9920.3
3000	945221.7	3889.6
4000	358774.9	5407.7
5000	980114.9	6748.4
6000	459806.7	7175.6

Table 1. With reference to Figure 9, showing each polygon's mean and standard deviation values after 20 simulations - information which has traditionally not been provided in commercial GIS packages. These statistics are important when polygon areas are required for computing density values, and allows further statistical analysis of the final results to be performed.

Goodchild, 1996). The model developed involves the creation of two independent, normally distributed, random error grids in the x and y directions. The grids are combined to provide the two components of a set of simulated horizontal positional error vectors regularly distributed throughout the region of the data set to be perturbed (Figure 8). The assumptions made are (1) that the error for each node or vertex has a circular normal distribution, and (2) that the x and y error components are independent of each other. The grids are generated with a mean and standard deviation equal to the estimate for positional error of the data set to be perturbed (a necessary condition for use of the model). These error estimates, for example, might come from the residuals at control points reported during digitizer setup or from an accompanying data quality statement.

By overlaying the two grids with the data to be perturbed, x and y positional shifts can be applied to the coordinates of each node and vertex in the data set to create a new, but equally likely, version of it. With the distorted version of the data, the user then applies the same set of procedures as required previously to create the final GIS product, and by repeating the procedure a number of times the variability residing in the end product may be assessed. Alternatively, several different data sets may be perturbed (each with their own error estimate) before being combined to assess final output uncertainty. While the model does require a positional error estimate for creation of the two distortion grids, it is the resultant uncertainty arising from the use of perturbed data due to simulation which is under investigation—hence its label as an 'uncertainty' model.

The first application of the model involved estimating the uncertainty in the areas of a set of six polygons that had been digitized from a source map at a scale of 1:50,000. It was determined that the digitizing had been performed with a standard deviation of 25m. The set of polygons was perturbed 20 times and the results of overlaying the 20 realizations can be seen in Figure 9. Then, by appending the 20 sets of polygon and statistically summarizing the areas of each of the six polygons, a table of mean polygon areas and their standard deviations was constructed (Table 1).

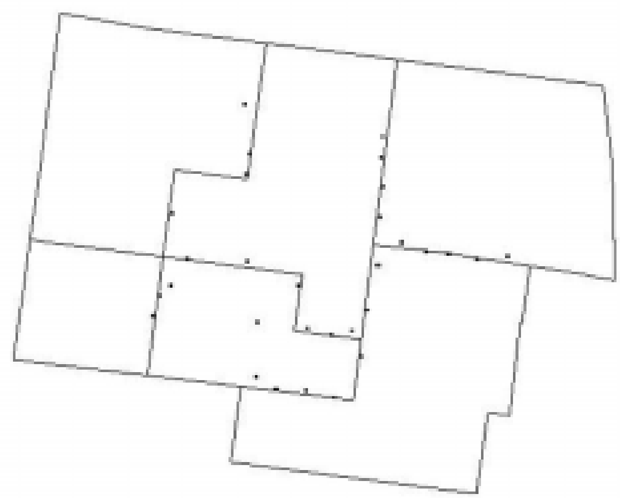


Figure 10. As shown above, 30 points were deliberately placed near the boundaries of a set of 6 polygons prior to perturbation to test the vector uncertainty model during point-in-polygon overlay operations.

In the next application a set of 30 point features, deliberately placed near polygon boundaries and junctions, was overlaid with the set of six polygons used previously. Both the points and polygons were perturbed 20 times each and then overlaid a total of 400 (20 x 20) times—a process that was easily automated with a short macro language script. As each perturbed point set was overlaid with the fixed polygon boundary file, the identifier of the polygon in which each point was deemed to lie was recorded. When the 400 overlays were completed a frequency count was taken and the results were summarized in Table 2.

Point ID	Poly 1000	Poly 2000	Poly 3000	Poly 4000	Poly 5000	Poly 6000
2	290 (0.73)	110 (0.27)	–	–	–	–
9	–	260 (0.65)	–	–	–	140 (0.35)
13	–	27 (0.07)	373 (0.93)	–	–	–
17	–	–	156 (0.39)	–	244 (0.61)	–
18	–	–	280 (0.70)	–	120 (0.30)	–
19	–	–	129 (0.32)	–	271 (0.68)	–
23	–	–	–	–	108 (0.27)	292 (0.73)
27	–	–	–	–	140 (0.35)	260 (0.65)

Table 2. With reference to Figure 10, the 30 points and 6 polygons were each perturbed 20 times and overlaid 400 (20x20) times in a point-in-polygon operation. The values shown above represent the number of times (out of 400) that each point was computed to lie in each polygon (with probabilities in brackets). Thus, useful likelihood estimates can be derived for points that lie particularly close to polygon boundaries—taking into account the different positional accuracies of the point and polygon data sets involved.

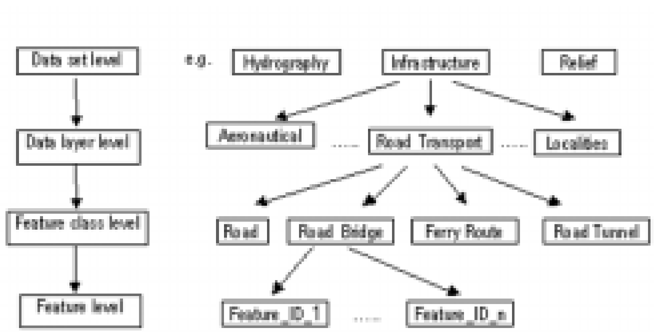


Figure 11. The hierarchical structure of objects in a GIS.

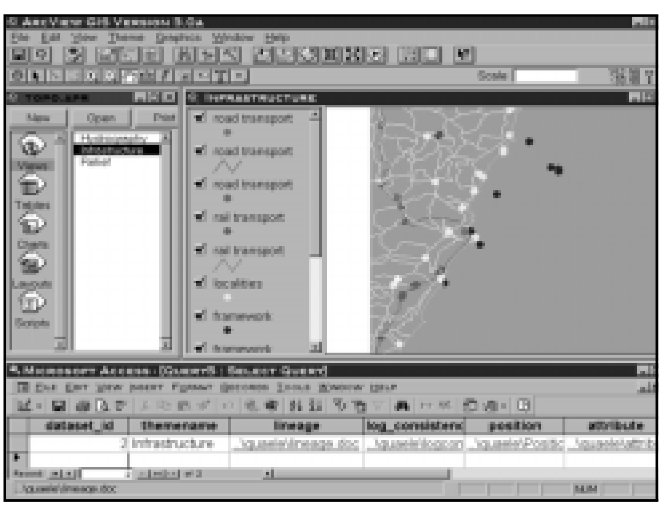


Figure 12. An example of accessing dataset level quality information via the customized "Q" (Quality) button (shown at the top left of the ArcView screen)

Example #5: Handling Data Quality Information at Different Levels

In this final example (see Qiu and Hunter, 1999), the hierarchical structure of objects in a GIS (as defined in Figure 11) was used as a means of storing and displaying data quality information at several different levels in the database structure. The concept has been implemented using Microsoft Access and ESRI's ArcView and applied to the AUSLIG TOPO-250 sample data set (which is a digital version of the 1:250,000 national topographic map series prepared by the Australian Surveying and Land Information Group). Three data sets (relief, infrastructure and hydrography) are included, and each data set further contains a set of data layers, for example, the Hydrography data set has four layers: drainage, framework, offshore and waterbodies. There is one quality statement file for each data set which contains information about each tile. Apart from the data quality information stored in the files, there is also a data quality table within which each record is linked to a feature occurrence in the TOPO-250K data. This table holds data quality information for individual features with respect to feature reliability, attribute reliability, planimetric accuracy and elevation accuracy. The data quality information prepared by AUSLIG for the TOPO 250K product is an excellent example of detailed quality reporting and can be readily accessed online at: http://www.auslig.gov.au. By linking both the data and data quality information, it was possible to store quality information at the data set, feature class and individual feature levels for query and display (as shown in Figures 12, 13, 14 and 15).

Conclusions

In this paper a variety of practical examples of spatial data quality reporting and assessment have been presented to demonstrate that the research agenda undertaken in this field over the past

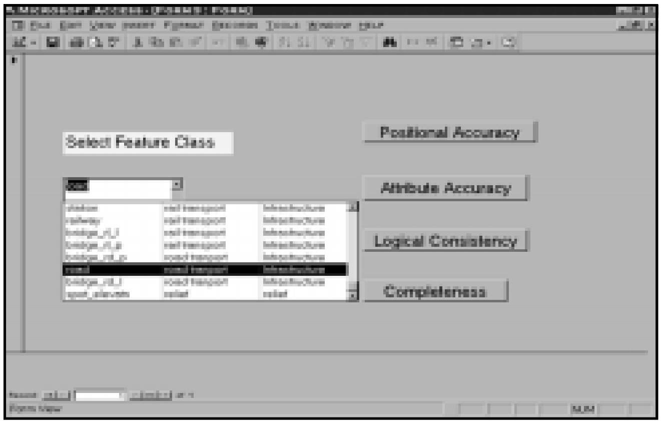


Figure 13. At the feature class level, the type of data quality information for the chosen class (road) is selected interactively once the "Q" button is selected.

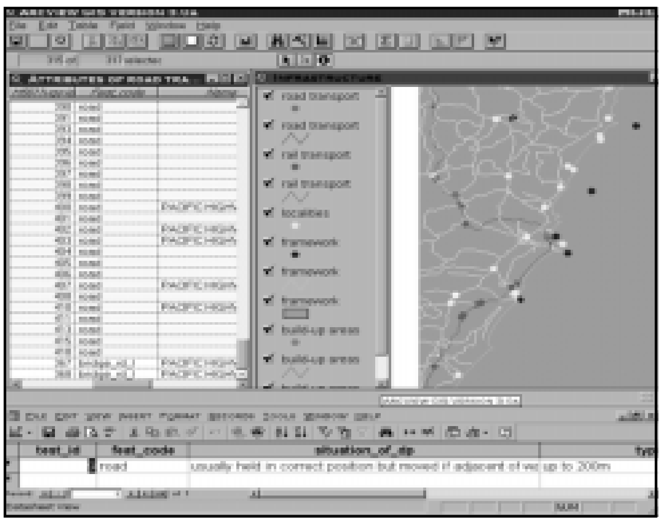


Figure 14. The result of the query is displayed on the screen with the selected feature class 'roads' highlighted in yellow, and (in this case) the positional accuracy information is shown in a table below.

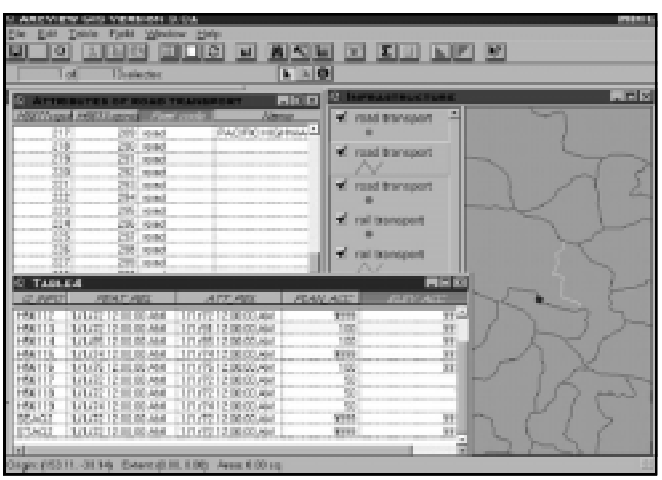


Figure 15. Data quality information can also be displayed for individual features.

decade is becoming more productive. The examples given here cover the tracking of feature coordinate edits and their reporting in visual data quality statements; testing and reporting the positional accuracy of linear features of unknown lineage; simulating uncertainty in products derived from Digital Elevation Models; incorporating uncertainty modeling in vector point, line and polygon files; and reporting data quality information at different levels of GIS database structure. Software code and tutorials for several of the cases illustrated are available via the Internet at the author's website (see below), and it is argued that the examples help demonstrate that data quality research is indeed moving from academic concepts to practical reality. While the next step required is to implement these tools and techniques in commonly available commercial GIS packages, which in turn will need closer liaison between researchers, users and system developers, it can at least be said that the theoretical research undertaken in the past is beginning to lead to practical applications which have the potential to become part of everyday GIS packages.

Acknowledgments

The author wishes to acknowledge the support received for this research through Australian Research Council (ARC) Large Grant No. A49601183 titled "Managing Uncertainty in Spatial Databases", and Australian Research Council Small Grant No. S49813054 titled "Quantification and Application of Data Quality Statements". The author also acknowledges the valuable research contributions made by Ms Karin Reinke in developing the Visual Basic software application for testing the positional accuracy of linear features, and by Ms Jibo Qiu in developing the software for modeling vector data uncertainty and in reporting data quality at different levels.

References

Hunter, G.J. and Goodchild, M.F., 1995a, "Dealing with Error in Spatial Databases: A Simple Case Study". Photogrammetric Engineering & Remote Sensing, 61, 5, pp. 529-37.

Hunter, G.J and Goodchild, M.F., 1995b, "A Methodology for Reporting Uncertainty in Spatial Database Products". Journal of the Urban and Regional Information Systems Association, 7, 2, pp. 11-21.

Hunter, G.J and Goodchild, M.F., 1996, "A New Model for Handling Vector Data Uncertainty in GIS". Journal of the Urban and Regional Information Systems Association, 8, 1, pp. 51-57.

Goodchild, M.F. and Hunter, G.J., 1997, "A Simple Positional Accuracy Measure for Linear Features". International Journal of Geographical Information Science, 11, 3, pp. 299-306.

Hunter, G.J. and Goodchild, M.F., 1997, "Modeling the Uncertainty of Slope Gradient and Aspect Estimates in Spatial Databases". Geographical Analysis, 29, 1, pp. 35-49.

Murillo, M. and Hunter. G.J., 1997, "Evaluating Uncertainty in a Landslide Susceptibility Model". Transactions in Geographic Information Systems, 2, 4, pp. 289-299.

Qiu, J. and Hunter, G.J., 1999, "Managing Data Quality Information". To be presented at the International Symposium on Spatial Data Quality, Hong Kong, July 1999, 10 pp.

URLs for the Applications Described in this Paper

Linear Feature Accuracy Testing Code and Tutorial
http://www.geom.unimelb.edu.au/people/gjh_notes/linetest.htm

Vector Uncertainty Model Code and Tutorial
http://www.geom.unimelb.edu.au/people/gjh_notes/vector.htm

Grid-Cell Uncertainty Model Code and Tutorial
http://www.geom.unimelb.edu.au/people/gjh_notes/grid.htm