Speaking Truth to Power: An Agenda for Change

1999: ***Spatial Accuracy Assessment Land Information Uncertainty in Natural Resources*** K. Lowell and A. Jaton, editors. Ann Arbor Press, Chelsea MI. pp. 27-31.

Nicholas R. Chrisman
Associate Professor
Department of Geography Box 353550
University of Washington
Seattle WA 98195-3550 USA
chrisman@u.washington.edu
(+1) 206 543 5870
FAX (+1) 206 543 3313

ABSTRACT

Despite substantial attention by the research community, data quality perspectives still do not inform the developments of new tools. This paper introduces a different perspective on "truth" in science based on studies of science and technology. Then it develops an example of the quality dimensions hidden in the black box of coordinate transformations. This paper calls for actions to bring greater capacities (like improved estimation algorithms) to users.

DATA QUALITY CONCERNS IN GIS

Fifteen years ago, I travelled to Hull, Qué bec to deliver a talk titled "The role of quality information in the long-term functioning of a GIS" (Chrisman 1984). The issue of accuracy and data quality was not considered the most important concern at Auto-Carto 6. My session was at the very end of the day, when most attendees were on their way to more relaxed pursuits. Many of the well-established experts on GIS were highly sceptical of my claim that data quality information would grow to become a major proportion of a total system. I do not need to detail the extent of the changes over the past fifteen years, but data quality is certainly much more prominent. The term metadata has created new kinds of information resources, as clearinghouses and digital libraries become a part of the information landscape. So it is no longer possible to claim that data quality issues have been ignored.

This third conference devoted to studies of accuracy of spatial data demonstrates the expanding interest. The research community has recognized the need and has vastly expanded the knowledge of many aspects of accuracy assessment and data quality more generally. Though it took many years to come to fruition, the concept of "fitness for use" has become an element of most data standards efforts worldwide (National Institute of Standards and Technology 1992, for example) See also: (Moellering 1991; Guptill and Morrison 1995). This signals a remarkable change from the era of technological paternalism when agencies promulgated their own criteria for accuracy assessment, and simply informed a user whether the map sheet "conforms to National Map Accuracy Standards".

Before we give ourselves too much credit, these events are by no means unusual. The overall trends during the past fifteen years have seen enormous alterations. The world economy has continued to globalize and has continued to place great reliance on information to reduce all forms of wasted effort. This trend has favored "flexible production" in the place of the old "Fordist" model of uniform products (Scott and Storper 1986). The emphasis on fitness for use can be interpreted as the cartographic community's reaction to these larger trends (Chrisman 1991). Since 1983, the trend in computer capacities has increased at ever expanding rates. Yet, the nature of our GIS software has not evolved as rapidly. Data quality, while it has become recognized in the research arena, has not become a driving force in the software packages available to most users. Most of the software is still emulates the procedures of the predigital era. There are many challenges that have not been adequately addressed. This paper will present one detailed example of a procedure that would have direct impact on virtually all GIS users. Before beginning the example, it is important to establish some revisions in the underlying reasons for studying data quality and map accuracy.

SPEAKING TRUTH TO POWER

The main title of this paper comes from the Quaker community. It encapsulates a spirit of resistance to established power structures, and the belief that persuasion will defeat prejudice. The connection to the study of map accuracy might seem immediate. Accuracy makes a claim to compare a measurement to the "truth", thus giving a strength in any argument. While this interpretation of accuracy is quite common, and may characterize the majority of those attending this conference, I have come to doubt this simple interpretation. I have found a different way of interpret the work of accuracy assessment that leads me to a more indirect reading of the Quaker maxim.

The "scientific method", at least since the time of Francis Bacon, has placed great emphasis on direct observations of nature. Truth is supposed to lie ready for discovery by the properly prepared scientist. The story of "objective discovery" implies that the researcher is not really responsible for their discovery; the truth was implicit in nature, simply waiting to be revealed. By contrast, untrue scientific statements are entirely due to human errors by the researcher. Desire to conform to social expectations, pressure to support political structures, economic weakness can be marshalled to explain away the mistakes of a scientist influenced by social distractions. As Bruno Latour (1993, p. 92) characterizes the common interpretation of science:

"Errors, beliefs could be explained socially, but truth remained self-

> explanatory. It was certainly possible to analyze a belief in flying saucers, but not the knowledge of black holes".

It is very hard to avoid the hero-worship of the great explorers of science whose efforts revealed such a succession of grand new truths. Like a history written after the fact to justify the actions of the victors, these heroic accounts do not stand up to careful scrutiny. Boyle's experiments on air pressure, to pick one well-documented account (Shapin and Schaffer 1985), did not seem at the time to have such pure objective truth due to the problems with the nature of his mechanical equipment – his air pump leaked and his observations flew in the face of long-established explanatory logic (dating back to Aristotle). Now we may think that Boyle's critics (such as Hobbes) were influenced by their context, their scholastic training, their reliance on authority, but so was Boyle who adopted a jury-like form of "witnessing" for his demonstrations. Boyle's scientific method was as strongly rooted in the image of the gentleman as it was in technical details of his pump. As a result of such studies, the sociologists of science have come to understand that error and truth do not come from two different kinds of work (Bloor 1976). Latour (1993) goes further in demanding a strict symmetry in seeing the constant interaction between nature and society. Truth is not directly or necessarily located in "nature", and error is not simply the result of human failings.

What do such academic arguments matter to practitioners of GIS? Any project manager knows that error is inevitable. No database can hope to represent every object in nature without any flaw. No one expects a Nobel prize for their efforts in operating a GIS, the work is closer to routine utility maintenance than high science. Yet, in the accuracy assessment sphere, we often set up standards of truth without thinking them through sufficiently. The standards and procedures of our measurements do not come from a pristine nature, but are deeply social in their origins and in their operations. As one who has sat through ten years of standards committee meetings, I can attest to the multiple levels of political, cultural, psychological interactions that produce the agreements that make science (or GIS data transfers) possible. Yet, the social component does not wall itself off from the rigor of observing the world, and trying to confront it as it is, not as we would like it to be. The critical element of data quality assessment is the willingness to confront all the array of powerful human institutions with the truth. It is in this act of persuasion, itself intensely social, that the Quaker maxim takes on its application to data quality assessment. All the social forces in the world become unconvincing when they suggest some phenomenon that simply does not exist. Persuasion, and the assembly of allies who will agree with your statements, is the real core of the scientific method (Latour 1987).

AN AGENDA FOR CHANGE

One of the most pervasive forces in technology is the tyranny of the way things have always been done. The research community can produce great advances, but they do not influence the practice of GIS until the research concepts change the routine ways of doing things. This paper cannot recount all the possible topics, but it will focus on one problem that highlights missed opportunities.

Coordinate transformations

If the central commonality of GIS is the integration of different sources, almost every application requires some form of geometric registration. To connect the measurements obtained on a digitizer, the device units are transformed into map coordinates by solving the correspondence for a set of known points. To transform a remotely sensed image to other layers, a set of known features are located on the image. Perhaps the greatest quantity of research on this process is in the field of photogrammetry. This procedure is critical to the whole GIS enterprise, yet the tools presented to users are considered so simple that they are rarely called into question. Like many elements of technology, once the decisions are made, it becomes a "black box" whose internal structure is not worth considering (Latour 1987).

To pry open the black box, some set of points are used to estimate the parameters for some mathematical formulation that will translate other points, not just those whose values are known. The decisions embedded in the software concern the nature of the mathematical formulae, for instance the use of linear equations as in the affine, the projective, or piecewise equations (White and Griffin 1985). The estimation method typically uses least squares, an decision that embeds certain assumptions. In addition, the number of points used for registration strongly influences the result.

The first issue is the number of points used for the transformation. This may seem like the least controversial from the research perspective. Considering that the user is fitting a linear equation with six unknowns (in the case of an affine), having three coordinate pairs provides no redundancy, hence no check for errors. How many points are "enough"? One might debate whether 12 or 20 would begin to be enough, but four (4) would not pass muster as a statistical procedure. Yet, the market leader in GIS software suggests four points (ESRI 1991, p. 5-13): "Select 4 widely spaced points common to maps A and B to be used as tics for A." ESRI software will handle many more points, but four is considered adequate. Some other systems only allow four, no more (Planet One Corp. 1997a). In the translation from the research community to the world of practice, important messages are not being communicated. In the case of Planet One, it seems to be a drive for a simple, minimal interface that led to a fixed number of control points. Also, they are simply following the lead of their Business Partner, ESRI.

Another part of the process involves the mathematics of the coordinate transformation. The affine, the overwhelming favorite in GIS software, is a rather odd choice. If the sources are in the same projection, and simply need to be subject to the rotation, translation and scaling described in most textbooks, then the similarity transformation should be adequate. This assumes that the scaling on the input axes should be preserved. The affine relaxes the assumption about a common scale and permits some degree of distortion (varying scales from one side to another). In many cases, sources run through these transformations are not in the same projection, In he USA, this is common when topographic quadrangles (compiled on state plane projections or the antiquated polyconic) are merged with UTM sources such as digital elevation matrices. The affine might paper over some of the local differences between two projections, but it is not the appropriate method to apply. Users are unlikely to be able to decipher the report of the transformation parameters to be able to see if the differences between the X and Y scales are reasonable, since the parameters reported include the rotation component in the same numbers. Here the software designers are serving their own needs to carry out routine mathematical calculations, but they are not reporting them back to the user in such a way to warn them of possible problems with data quality. It is most amusing to see the dialog box (Planet One Corp. 1997b) provided in one ArcView tool that calculates an affine transformation although the software provides no way to actually rotate a background image. The user is told to go rotate by some many degrees in some other package. Some software packages also provide Helmert's projective transformation. This is particularly appropriate for unrectified air photographs and any source where the scale would vary outwards from a center. Of course, this requires more parameters, and four points do not provide any redundancy. To solve for a center and the scaling away from it, the points should also be distributed more densely than the four-corner approach that seems to be common in GIS documentation. The four corner points are all equidistant from the center of a rectangle, so they provide no information on change in scale from the center. Though the option to use the projective is there, most users have very little reason to deviate from the default.

The deepest dark corner of the black box is the method of estimation. Software packages will use least squares solutions to generate the parameters that fit the observed points. Least squares is a common fixture of statistical methods and other applied mathematics. The time-tested technique does indeed produce the most efficient use of the observed information, under a set of assumptions that need to be recalled and more clearly advertised to users. The conditions for least squares are simple, and quite strict. And they are fundamentally social at their core, they create a division of labor between the estimation technique and the user. They also distinguish "good" (random) error from "bad" (human-created) error. These divisions are counterproductive.

Least squares provides a Best Linear Unbiased Estimate (BLUE) of the parameters under the condition that the errors in the points come from a common normal distribution. The short hand is *iid* – independently and identically distributed normal variates. Academic treatment of error and adjustments (Mikhail and Ackermann 1976, for example) divides error into three basic categories: systematic, random and blunders. Each of the coordinate transformation methods will remove various kinds of systematic effects. The least squares method is designed to provide the best estimate of the parameters, given the random error in the points. Blunders are meant to be removed prior to invoking the procedure. Here is a grand chasm between the world of research and the world of practice. With only four points, the user has next to no information to alert to the existence of blunders. In some software, the text message about the fit of the parameters flashes by on a screen that is covered over by the graphics display. Here the software conspires to hide the numerical details from the user, thus creating a dysfunctional situation.

Blunders are a residual category, created by the mathematical model. Since "random" error is defined as having iid Normal distribution, anything else is simply not the responsibility of the statistical model. Blunders are blamed on the user, they are tied to human failings such as reversed digits, selecting the wrong object on the photo, and whatever else might go wrong. The logic here is air-tight; least squares is the most efficient procedure given random error, blunders do not behave according to random distributions, hence all blunders must be removed prior to using least squares. There is a substantial literature on removing blunders from photogrammetric and surveying adjustments (Kavouras 1982; Kubik et al. 1988, for example). Detecting blunders certainly requires more points than the bare minimum, so the standards of practice in GIS do not support this necessary step.

Perhaps in an earlier period such a two-phase procedure made sense. On a manual calculator, least squares can be done in a single pass, but more fancy iterative techniques would impose a serious time cost. The explosion of computer capacity has changed all the ratios of effort, but the software industry has not fully exploited these changes. In other numerical disciplines such as statistics, there have been a continual development of outlier detection methods, of exploratory procedures more generally, and of robust estimation procedures (Draper and J.A. 1981; Hoaglin et al. 1985; Hampel et al. 1986, for example). The research community in the mapping sciences has taken notice of these developments (Kubik et al. 1987, for example). One particular estimation procedure, least median squares (Rousseeuw 1984; Rousseeuw and Leroy 1987), has been applied for estimation of mapping transformations (Shyue 1989). Rather than separating the steps, the points can be contaminated with up to fifty percent blunders, and the estimation will ignore the blunders and fit the parameters to the rest of the data. Unlike some of the complex weighting techniques, this requires virtually no user intervention or tricky tuning parameters. In the spirit of bootstrap methods, the least median procedure (Rousseeuw and Leroy 1987) is obtained by sampling the possible combinations of data points. The regression is estimated for each combination, and the one with the best fit (least median squared) is selected. For large numbers of points, the number of possible combinations rises very rapidly, but, for the typical numbers of points used for map registration (10-20), a few thousand samples are bound to separate out the outliers. Of course, least medians do not wring all the value out of every point, but they remove the division of labor. This estimation technique is built to compensate for the possibility of blunders.

Least median squares and other robust techniques are a technology with clear advantages over the earlier techniques. The academic research sector has done its job in exploring these techniques and reporting them in academic outlets. None of this changes the tools that users see until the software vendors adopt these innovations. There is nothing inevitable about this next step. Good ideas do not move inexorably into the large software packages. As long as the documentation presented to the user is so slim, the users barely recognize how the process is done by their current software. It is unlikely that consumer demand will work. The academic sector tends to continue chasing off towards every better improvements, without ensuring that their innovations become converted into practice. And, to recognize the harried circumstances of the software team, no software manager wants to waste effort on some feature that no one will recognize. There are plenty of other priorities for each actor in the network. Consequentially, the story of coordinate transformations is the story of many other situations where the software remains stuck on the tried and true, not the most recent research results.

CONCLUSION

The example of coordinate transformation procedures points out that data quality issues cannot be described by setting up some external "truth" as the arbiter. The concept of best fit has developed from a given historical context when certain kinds of computing were more possible than others. These approaches, once adopted, do not simply shrivel up when the original rationale changes. Each group involved, users, programmers, sales persons and researchers share the responsibility for the current failures of communication. The power of the way things have always been is what must be resisted. New ideas must be demonstrated so that all can benefit. Too frequently the research community thinks their job is over when the publication has been accepted. There is a responsibility to convert important ideas into changes in practice that benefit a larger community.

REFERENCES CITED

Bloor, D. *Knowledge and Social Imagery*. London, Routledge & Kegan Paul, 1976.

Chrisman, N. R. The role of quality information in the long-term functioning of a geographic information system. *Cartographica.* 21(3&4), pp. 79-87, 1984.

Chrisman, N. R. Building a Geography of Cartography: Cartographic Institutions in Cultural Context. *Proceedings International Cartographic Association 15th Conference,* Bournemouth, UK, 1, 83-92, 1991.

Draper, N. R. and J.A., J. Influential observations and outliers in regression. *Technometrics.* 23, pp. 21-26, 1981.

ESRI. *Map Projections and Coordinate Management: Concepts and Procedures.* Redlands CA, Environmental Systems Research Institute, 1991.

Guptill, S. K. and Morrison, J. (Eds.). *Elements of Spatial Data Quality.* published on behalf of the International Cartographic Association Oxford, Elsevier, 1995.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. *Robust Statistics: the approach based on influence functions.* New York, John Wiley, 1986.

Hoaglin, D. C., Mosteller, F. and Tukey, J. W. *Exploring Data Tables, Trends, and Shapes.* New York, John Wiley, 1985.

Kavouras, M. On the detection of outliers and the determination of reliability in geodetic networks. *Technical Report 87.* Department of Surveying Engineering, University of New Brunswick, 1982.

Kubik, K., Lyons, K. and Merchant, D. Photogrammetric work without blunders. *Photogrammetric Engineering and Remote Sensing.* 54, pp. 167-169, 1988.

Kubik, K., Merchant, D. and Schenk, T. Robust estimation in photogrammetry. *Photogrammetric Engineering and Remote Sensing.* 53, pp. 167-169, 1987.

Latour, B. *Science in Action*. Cambridge MA, Harvard University Press, 1987.

Latour, B. *We Never Were Modern*. Cambridge MA, Harvard University Press, 1993.

Mikhail, E. M. and Ackermann, F. *Observations and least squares*. New York, IEP-Dunn-Donnely, 1976.

Moellering, H. (Ed.). *Spatial Database Transfer Standards: Current International Status.* New York, Elsevier, 1991.

National Institute of Standards and Technology. *Spatial Data Transfer Standard*. Washington DC, National Institute of Standards and Technology, Department of Commerce, 1992.

Planet One Corp. Image Registration and Reference Point Tools. http://www.planetonegis.com/pages/image/text_image13.htm, 1997a.

Planet One Corp. Recommending a Rotation. http://www.planetonegis.com/pages/image/text_image22.htm, 1997b.

Rousseeuw, P. J. Least median squares regression. *Journal of the American Statistical Association.* 79, pp. 871-880, 1984.

Rousseeuw, P. J. and Leroy, A. M. *Robust Regression and Outlier Detection*. New York, John Wiley, 1987.

Scott, A. J. and Storper, M. (Eds.). *Production, Work, Territory: The geographical anatomy of industrial capitalism*. London, Allen & Unwin, 1986.

Shapin, S. and Schaffer, S. *Leviathan and the Air Pump: Hobbes, Boyle, and the Experimental Life*. Princeton NJ, Princeton University Press, 1985.

Shyue, S. W. High breakdown point robust estimation for outlier detection in photogrammetry. unpublished PhD dissertation, University of Washington, 1989.

White, M. and Griffin, P. Piecewise linear rubber sheet map transformation. *The American Cartographer.* 12, pp. 123-131, 1985.