OBTAINING INFORMATION ON QUALITY OF DIGITAL DATA

Nicholas R. Chrisman
Department of Landscape Architecture
University of Wisconsin - Madison
Madison WI 53706 USA

ABSTRACT

Quality information has been recognized as an important component of geographic information systems, at least in theory. This paper expands on the definitions of quality information with particular reference to sources. Many systems developers believe that quality information will make enormous demands, and that the user may be unwilling to pay for such extravagance. This paper provides a counterargument, maintaining that quality information is currently processed and maintained. Digital cartography usually proceeds on the presumption that the manual system is hopelessly outdated and encrusted with peculiar rites of dubious importance. This paper argues that quality information is a continual concern in the traditional practice of cartography. Examination of a few diverse cases reveals a broad range of quality information maintained by producers which can not be fully integrated into current spatial data bases. Better exploitation of quality information can improve communication with users and it can also cut costs inside the producing agency.

BACKGROUND: DEVELOPMENT OF QUALITY CONCERNS

The intellectual history of AUTO-CARTOs and geographic information in general has been dominated by various aspects of geometry. Quality, although it is a common concern in the traditional manual practice of cartography, has not been converted fully into the digital arena. One example, the DIME development, shows how the field of geographic information systems development has focused on geometric concerns. DIME was the first full-fledged development of a topological data structure in the United States (Cooke and Maxfield, 1967). DIME is the intellectual precursor to many other systems, such as the POLYVRT structure described by the author (Peucker and Chrisman, 1975). In that description, the focus falls almost exclusively on the use of the topological structure to construct a clearer, more useful model of geometric relationships. It is interesting to note that the original purpose of DIME, and one of its great strengths (see White, 1978) was the verification of logical consistency. For the mathematical purist, logical consistency may be a property of a set of geometric axioms, but it also forms a part of quality.

Whatever the reasons, perhaps the field finally got tired of the never-ending debate between rasterologists and vector purists, the issues of data quality are now perceived as important. It is difficult for some of the contributing disciplines to attach great importance to issues outside positional accuracy, but even that situation is changing. The US National Committee for Digital Cartographic Data Standards [NCDCDS] (1985) Interim Proposed Standard recognizes five components to a complete quality report. The first component is lineage which records all the source materials and transformations which produced the final product. Lineage is a major topic on its own. The other four categories, positional accuracy, attribute accuracy, logical consistency and completeness, relate to testable concerns. These categories have been accepted for the standard to be used in the exchange of digital data; the categories are logical but they do not appear with these labels in the traditional cartographic textbooks.

One of the most artificial parts of the practice of digital cartography has been that data bases are seen as static. The standards are developed to accomodate exchange of data, but not how to incorporate part of it into another data base. The NCDCDS Working Group has recognized this problem, at least as far as requiring lineage to record the process. However, the long-term fate of digital data will require substantial development towards a dynamic system that creates its own audit trail. Some of these issues have been presented at a previous AUTO-CARTO (Chrisman, 1983), and some will be presented below.

QUALITY INFORMATION: A DEFINITION

Quality information is not a new component of the cartographic production process. Over the past century, the typical steps of inspection and review have endured without significant change, whereas the steps of collection and representation have been dramatically altered by material technology. For example, photography affected the production process from collection (photogrammetry) to printing (photo-offset). Such material technologies may change the quality (or cost) of the results, but they rarely affect the procedures used to ascertain and control quality. These quality procedures are most strongly affected by institutional structures. Information systems are only the current wave of change in the collection and representation of spatial data. Due to the abstract nature of this technology, the quality control component will be affected more deeply. This paper explores the sources and organization of quality information.

The fundamental changes in the role of quality information in cartography can illuminated by the alternative definitions of quality advanced in the field of industrial quality control.

> "Quality has been variously defined as 'fitness for use', 'meeting an expectation', 'degree of excellence', and 'conformance to a standard', along with other phrases. These all have merit depending on one's point of view." (Hayes and Romig, 1977, p. 9)

In the days of graphic reproduction, rather than digital transmission, the physical medium - the map product - set many limits on the use of cartographic information. A production agency could internalize the needs of a preconceived user and set a fixed standard of accuracy. Typically, these standards were dictated by the available technology, not the user's desires. The producing agency was responsible for all the steps required to meet the expectation. The nature of the steps was not of particular interest to the consuming public, as long as the aura of accuracy could be supported. Under this arrangement, cartographic production could be easily fit into the quality control models used for industrial production of objects like light bulbs. The definitions of quality such as *meeting an expectation* , or *conformance to a standard* fit the situation. The user need not be told much more than the simple notice of the standard applied [for example the US National Map Accuracy Standard (Bureau of the Budget, 1947) requires the statement **This map complies with National Map Accuracy Standards**]. As a statement about quality this sentence is useful only if the NMAS incorporates all relevant factors and correctly distinguishes useful maps from useless ones. To avoid unacceptable false positive cases, NMAS must be conservative, but that may lead to false negatives - maps which have certain uses, but do not fit all of the criteria. One common procedure to handle this problem creates a hierarchy of standards with successively less strict thresholds. Such an approach may work for information such as geodetic surveying where the criteria are essentially unidimensional.

Digital data can adapt to a broader range of uses with a broader range of special demands, some of them beyond the imagination of the producing agency. This trend is both good and bad. The expansion of uses will enhance the value of the product, broaden the user community, and sustain the producer's bureaucratic viability. On the negative side, the flexibility of digital processing increases the possibilities for inappropriate handling. One reaction to misuse is to tighten standards and produce the best possible information. However, given a fixed budget, the tighter standards will probably lead to less complete coverage and thence to duplication of effort by other agencies working at less strict standards. The result is not stricter standards, but two incompatible standards. This situation should be familiar to Americans (General Accounting Office, 1982), but it can happen almost anywhere.

In the definitions advanced by Hayes and Romig above, *fitness for use* stands out. This definition can imply a different relationship between producer and consumer. The root of data abuse is not in the quality of the data, but in the awareness and understanding of the quality of the data. By converting to the fitness for use approach, the problem of data abuse is moved from producer to consumer. The evaluation and judgement of fitness for use must be the responsibility of the user, not the producer. To carry out this responsibility, the user must be presented with much more information to permit an informed decision. Conversion from the old kind of standard to the fitness for use approach will not be simple.

In the deliberations of the NCDCDS Working Group on Data Set Quality, the fitness for use concept has formed the base of the Interim Proposed Standard (Chrisman, 1984; NCDCDS, 1985). By the time this paper is presented, this concept will probably be a part of the proposed US national standard. Fitness for use creates distinct responsibilities for producers and consumers. The producer is not responsible for achieving any particular threshold of quality, only for informing the user about what was done and the actual level achieved. The user has the responsibility to examine this level of performance in order to judge if the particular product is fit for the given use. This standard applies the fitness for use concept in a method characterized as *truth in labelling*. Truth in labelling has its origins in the reformist consumer movements of one hundred years ago that replaced the ancient standard of commercial behavior - *caveat emptor* - with the registered ingredients list certified by the Pennsylvania Department of Agriculture and other aspects of the Pure Food and Drug Act. The NCDCDS proposal is more complex than an ingredients list, but it comes out of that consumer-oriented tradition.

Truth in labelling promises to alter the operation of producers and consumers of geographic information. Consumers will have to become aware of the issues of data quality and informed about technical processes not currently publicized. Producers particularly will have to make adjustments. To some extent, truth in labelling is a permissive standard. It permits a producer to distribute whatever product is obtained; there is no fixed threshold of performance required. The lack of fixed thresholds is a necessity in a polyglot world of varied applications. The difference between the duplicative effort mentioned above and the possibility of divergent thresholds here may seem to be slight. One difference is that under a fitness for use concept, there is no particular need for defensive arguments about abstract standards or professional bias; the debate can focus on the ability to meet user needs.

Most importantly, the truth in labelling approach may be able to drive out lax standards by exposing inner workings to scrutiny. Producers may be more likely to upgrade the quality of their products when they have to report all of the details to the professional community. Under the old kind of standard, there was much less possibility of embarrassment. In a more concrete way, consumers could vote with their feet and simply not acquire data which does not fulfill their needs. Thus the truth in labelling standard fits into its consumer reform, but capitalist, origin.

One major objection to the truth in labelling approach might be that it will cost too much to implement. Producers would have to develop quality reports for public consumption, instead of the current internal procedures of quality control and quality assurance. This paper will attempt to demolish this objection to truth in labelling. To demonstrate that a full quality report can be developed at a relatively minor cost, it will be shown that the components of that quality report are already known to most agencies. The difference is in organization and availability of the quality information.

METHODS OF ENSURING QUALITY

While the NCDCDS Interim Proposed Standard is constructed to be understood by a consumer, there is a need to convert its five components into the information as it is organized by a producer. From the perspective of a producer, quality information derives from quality control and quality assurance (QC/QA) procedures. The work of the NCDCDS Working Group recognizes four levels of testing that can be applied to obtain information for the quality report. The four are ranked more or less in increasing rigor: deductive estimate, internal evidence, comparison to source, and test against an independent source of higher accuracy. The list of possible tests is broader than normally considered in the industrial QC/QA literature. The sheer size of the earth imposes special restrictions, along with issues of expense.

Quality tests for spatial data are also in a formative state. The NCDCDS Working Group felt hesitant to push a particular research idea into the standard before it got out of the development stage. Even the test procedures for the positional accuracy of well-defined points are still in the process of adoption by the American Society of Photogrammetry (1985). Considering the mass of digital data in geographic information systems, there is remarkably little work done on the testing of positional accuracy for features that are not well-defined (ill-defined?). Similarly, the procedures used for attribute accuracy assessment are open to debate. If the Working Group had created a standard tied to the most rigorous tests, there would be no way for producers to comply without massive research efforts to create the methods. It would be easy to draw the wrong conclusion from this discussion. The state of quality assessment for spatial data is not completely hopeless. There are many procedures used to check quality. These procedures have been in place for many years, without being integrated into the digital processing system. A digital system cannot take over as the primary operation until it can incorporate these concerns.

Deductive estimate

Deduction, though the weakest level of test, forms the basis for a large bulk of information systems. It is the way to take the scarce resources (such as testing) and apply them to the greatest number. The most common example is found in procedure manuals and specifications for tasks. Manuals and specfications may make reference to testing, in which case the section on that level of testing applies. In many other parts, the logic is clearly deductive. Testing for a particular element is not required if that element is obtained and processed through a thoroughly known procedure.

Another facet of deduction is professionalism. If the actual procedure cannot be written down in detail in a manual, the idea is that a trained professional will take the correct action. Considering the ability of mapping professionals to disagree, this approach does not have a high degree of reliability. For instance, there are numerous section corners which have been remonumented many times each time by well-esteemed members of the profession. In a sour view of the world, the call for professionalism can sound like a guild system of oligopoly. Still, there are elements in quality which are not fully reduced to numbers. The training and experience of a professional may permit as valid an estimate of quality as a much more expensive and laborious test.

Repeated measurement (internal evidence)

The fact that internal evidence rates just above deduction should not make it seem weak. A system of repeated measurement can provide strong test results. In fact, as the NCDCDS reports point out, internal evidence may be the highest level of test possible for some topics. Logical consistency of a data structure can only be judged internally. Also, geodetic surveying involves a global (literally) network of measurements with only the most exotic procedures to test it from external sources. Internal evidence works by some form of redundancy designed into the data collection system. In the topological model, this is the dual encoding of the graph. In surveying, it is repetition of measurement through closure of traverse, and adjustment. Many other sciences have similar experimental structures. Biological sciences have used rats in cages and other repeated experiments to weed out the signal from the noise. Social sciences also depend on statistical notions where there is no absolute standard outside to permit testing.

For all the similarities to other sciences, spatial information has some complexities which require special treatment. Much of the work in social sciences, for example, is to discover the dimensionality of the objects described. Short of a very few obscurantists (for example, Atkin, 1974), dimensionality of geographic information is not an issue that merits much debate. By contrast, the statistical models applicable in biological and social sciences rely on strictly independent individuals which cannot apply to spatial phenomena. Surveyors understand that adjustment of repeated measurements can only be performed with the correct model of error; the differences between triangulation and trilateration must be reflected in the model. Similarly any new device must be modelled, such as the accelerations in inertial surveying. Error models have only been developed for a limited set of cases. There may be need for more distinct spatial error models than needed for the rest of known science.

Comparison to source: an incremental strategy

Comparison to source is practiced in a rudimentary way in the typical digitizing operation by checkplots. This procedure is much less sophisticated than the repeated measurement systems of surveying. However, it contains a germ of a system which uses something more than strictly internal evidence. A checkplot can detect a systematic bias in the digitized product under certain conditions. The use of this strategy operates incrementally; each step in the production cycle is checked against the last. It is still possible for the errors to propagate, but each addition is checked and noted. This strategy is also the scheme used when instruments are calibrated. Calibration tests are not performed on the data, but on the instruments. The result on quality analysis is similar. Many agencies operate on the basis of independent calibration of each machine. The specifications for error in each step may not be assembled at the end to determine overall levels of error. It is too easy to presume that the whole error tolerance can be spent at each step, leaving later steps to accumulate error well beyond specifications. A major advantage of the quality report is that each agency will be forced to examine its whole operation to find this kind of problem.

Tests against independent sources of higher accuracy

The best determination of total error in a system is not a complex model of propagated error between calibrated instruments. At some point, it is necessary to take a separate path from the surface of the earth, using more reliable procedures. Testing is mentioned in the NMAS, but in rather the manner used in nuclear arms agreements (national technical means of verification...).

> **The accuracy of any map may be tested** by comparing the positions of points whose locations or elevations are shown upon it with corresponding positions as determined by surveys of a higher accuracy. Tests shall be made by the producing agency, which shall also determine which of its maps are to be tested and the extent of such testing. (Bureau of the Budget, 1947)

By placing testing at the option of the producer, and providing for no direct report of test results, there was no incentive to perform tests. Some of the major map series of the US have not been tested in decades, despite many alterations in the production system. This situation is not particular to the US, it is a consequence of the centralized approach.

Tests are complex to perform, and may not seem earthshattering when they are done. Considerable education is needed so that users will be able to assimilate test results. However, these problems are no reason to abandon a testing program.

AVAILABILITY OF QUALITY INFORMATION

The theory of the NCDCDS quality report may be sound, but adoption is not certain. One approach is obtain compliance to the new standard is to find a large amount of new funds to add to the budgets of all agencies that subscribe to the standard. This is most unlikely in the current fiscal climate nearly everywhere. Another approach is bureaucratic strong-arm tactics. With sufficient political force, agencies can be made to comply. Without an incentive, this approach is very likely to fall apart over time. The most stable solution is to make the new standard work to the best interest of all parties. To ensure that quality reports get written, they must be in the interest of the producing agency. They also cannot introduce large additional costs.

NCDCDS Testing cycle

There is limited evidence about the costs of quality reports, however, the NCDCDS has just finished a testing cycle which offers a preliminary suggestion. In the external test conducted for Working Group II, Bell South (a service company owned by the telephone utility holding company for the southern US) prepared a quality report for a facility base file. This work was performed by their mapping contractor, Donohue Intelligraphics and AeroMetric Engineering, a photogrammetric subcontractor. These contractors were asked to follow the specifications of the Interim Proposed Standard, using information that they had kept concerning a project that was complete. An adequate quality report resulted without substantial additional cost. The main difficulty was the writing and word processing. Donohue thought that the standard could help in reducing certain costs by forcing reports into a more uniform format. It is particularly interesting that even for this spaghetti data base, produced on an Intergraph, there were items to be reported under all the topics. The positional accuracy of the product was assigned by deduction from a set of calibration tests performed by AeroMetric. The attribute tagging was quite crucial to Bell South and had been verified closely. Completeness was checked by reference to a master address list. Even the closure of certain features had been checked (laboriously) to provide information to report on logical consistency.

National Ocean Service

A more comprehensive case study on quality information has been conducted over the past year at the University of Wisconsin. Robert Gurda, Kate Beard and myself have conducted a study of the procedures used in the Nautical Charting Branch of the National Ocean Service. This agency is responsible for the hydrographic charting of the United States. The overall goal of the project is to revise the accuracy standards associated with manual and digital products. The first phase, a survey of the steps used to produce chart information, is of particular relevance here.

The goal in the first step was not to review all the steps of production. NOS was currently engaged in a process with the Office of Personnel Management to produce such a manual. When we saw the draft, it ran to dozens of volumes and it was increasing. Our interest was in the steps where some aspect of positional information was treated and transformed. (The similar inquiry for attribute information would have been much more difficult since there are more decisions made on attributes.) The information we gathered was condensed onto six pages of complex diagrams. One of the simpler pages, covering the hydrographic survey component, appears on the next page (Figure 1). The lesson about nautical charting is that it is complex, but so is almost any mapping operation. The NOS procedures included many steps of review, and each decision is backed by signatures or initials of those responsible. This procedure may seem overly military or bureaucratic, but it tends to emphasize the importance of quality information.

At NOS, there was a very complete audit trail available on every product of the agency. The nautical chart itself has a chart history which shows each alteration and addition. An automated version of the chart history could be created that would log each transaction as it occurs. Operations inside the information system could reformulate this raw data into summaries of changes for geographic coverages, or for operators or whatever. The hydrographic surveys covered in Figure 1 come with a report that runs to forty or fifty pages. This Descriptive Report includes the following sections in its table of contents:

- Hydrographic title sheet (describing location, personnel, dates, etc.)
- Project (description of project instructions with ammendments)
- Area surveyed (general description, including reference to nautical traffic)
- Sounding vessels
- Control stations (by name) ; Position control (range/range, etc.) ; Crosslines ; Shoreline
- Junctions and connections to previous surveys
- Aids to navigation ; MIscellaneous (history of failures of equipment, etc.)
- Letters of transmittal ; Approvals ; Signatures
- Calibration results for equipment used ; Field tide notes ; Verifier's Report ; Inspection Report

This exhaustive list should provide ample proof that some agencies will have little problem complying with the quality standards. In the quality information maintained by the NOS, there are calibration tests, comparisons with independent sources of the same accuracy (junctions to adjacent surveys), and many cases of internal evidence. The whole process of review and crosschecking is another attempt to drive out the flukes of personal bias and to stick to the facts. Some of the process is controlled by deduction and the rule of the Instructions and the Manual.

Not all cartographic production is performed by agencies with the kind of bureaucratic background of NOS (dating back almost two hundred years). However, the approach to quality information may be worth examination by others. One important facet of the nautical chart operation is the sense of the long-term responsibility. In typical geographic information systems applications, the analysis chases the current hot topics, rather than building up the institution to handle the data base over the long haul.

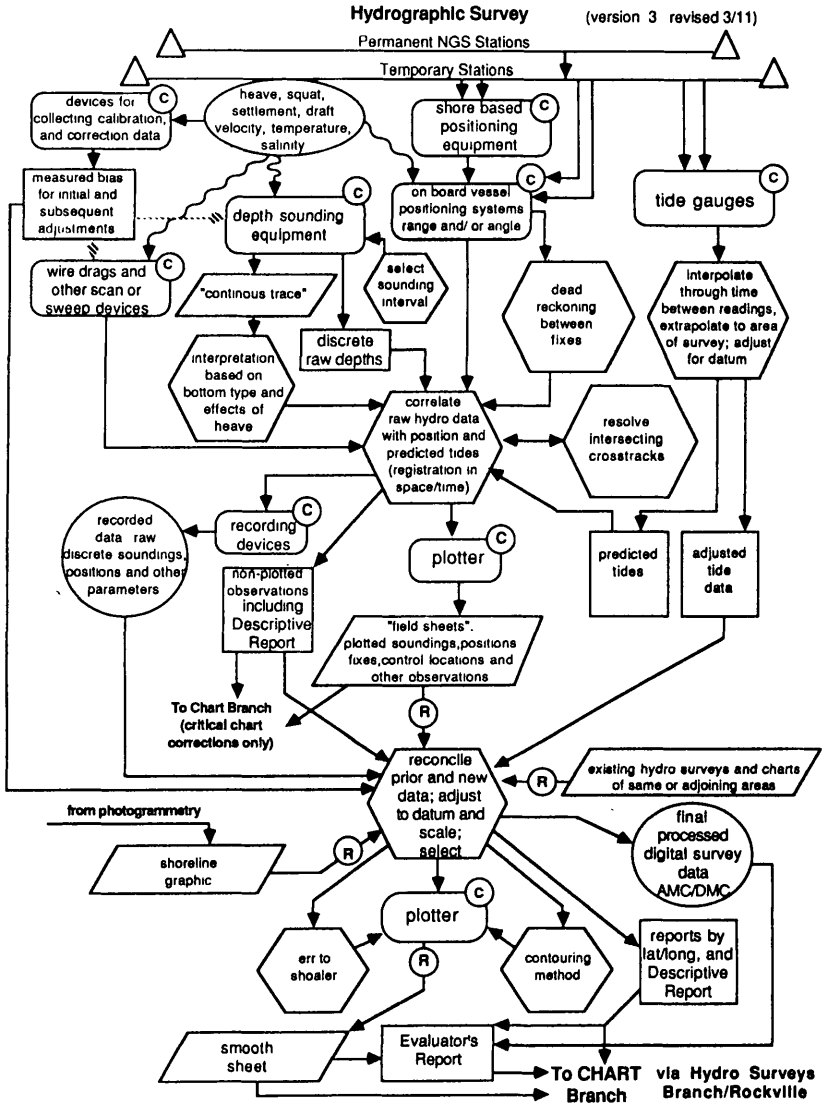
Hydrographic Survey
(version 3 revised 3/11)
Permanent NGS Stations
Temporary Stations
devices for collecting calibration, and correction data
C
heave, squat, settlement, draft velocity, temperature, salinity
shore based positioning equipment
on board vessel positioning systems range and/ or angle
tide gauges
measured bias for initial and subsequent adjustments
depth sounding equipment
wire drags and other scan or sweep devices
"continous trace"
select sounding interval
dead reckoning between fixes
Interpolate through time between readings, extrapolate to area of survey; adjust for datum
discrete raw depths
interpretation based on bottom type and effects of heave
correlate raw hydro data with position and predicted tides (registration in space/time)
resolve intersecting crosstracks
recorded data raw discrete soundings, positions and other parameters
recording devices
plotter
predicted tides
adjusted tide data
non-plotted observations including Descriptive Report
"field sheets". plotted soundings,positions fixes,control locations and other observations
R
To Chart Branch (critical chart corrections only)
reconcile prior and new data; adjust to datum and scale; select
existing hydro surveys and charts of same or adjoining areas
from photogrammetry
shoreline graphic
final processed digital survey data AMC/DMC
err to shoaler
contouring method
reports by lat/long, and Descriptive Report
smooth sheet
Evaluator's Report
To CHART Branch via Hydro Surveys Branch/Rockville

Figure 1

CONCLUSION

The general categories of quality information are a starting point for the study of data collection and organization. Useful quality information can be obtained without massive testing programs, although the tests will resolve difficult issues more easily. Agencies which will have an easy time are those with specific mandates to provide the information. This kind of relationship, with its long term implications, makes for higher quality information.

ACKNOWLEDGEMENTS

The research reflected in this paper was carried out in a number of projects, including the Dane County Land Records Project. Much of the specific work on the Nautical Chart Branch was supported by a contract with the National Ocean Service, which provided funds for this paper. The author acknowledges the interactions with other members of the National Committee for Digital Cartographic Data Standards, but the views expressed in this paper are strictly personal.

REFERENCES

American Society of Photogrammetry, 1985, Accuracy specifications for large-scale line maps, ***Photogrammetric Engineering and Remote Sensing***, 51, 195-199.

Atkin, R.H., 1974, ***Mathematical Structure in Human Affairs***, New York, Crane Russak.

Bureau of the Budget, 1947, ***United States National Map Accuracy Standards:*** Washington, Government Printing Office.

Cooke, D.F. and Maxfield, W.F., 1967, The development of a geographic base file and its uses for mapping, ***Proceedings URISA***, 207-218.

Chrisman, N.R., 1983, The role of quality information in the long-term functioning of a geographic information system: ***Proceedings AUTO-CARTO 6,*** 1, 303-317.

Chrisman, N.R., 1984, Alternatives for specifying quality standards for digital cartographic data: p. 43-71 in Moellering, H. (ed.) ***Report 4***, Columbus OH, National Committee for Digital Cartographic Data Standards.

General Accounting Office, 1982, Duplicative federal mapping: a growing problem: ***Report GAO/RCED-83-19***, Gaithersburg MD, General Accounting Office.

Hayes, G.E. and Romig, H.G., 1977, ***Modern Quality Control:*** Encino CA, Bruce.

National Committee for Digital Cartographic Data Standards, 1985, An interim proposed standard for digital cartographic data quality, p. 17-22 plus Supporting Documentation p. 113-132 in ***Report 6***, Columbus OH, NCDCDS.

Peucker, T.K. and Chrisman, N.R., 1975, Cartographic data structures: ***American Cartographer***, 2, 55-69.

White, M.S., 1978, A geometric model for error detection and correction: ***Proceedings AUTO-CARTO III***, 439-456.