Title
Accuracy of Spatial Databases: Annotated Bibliography (89-9)

Author
Veregin, Howard

Publication Date
1989-09-01

National Center for Geographic Information and Analysis

Accuracy of Spatial Databases:
Annotated Bibliography

Compiled by

Howard Veregin
National Center for Geographic Information & Analysis
University of California, Santa Barbara, CA 93106

Technical Report 89-9

September, 1989

Simonett Center for Spatial Analysis
University of California
35 10 Phelps Hall
Santa Barbara, CA 93106-4060
Office (805) 893-8224
Fax (805) 893-8617
ncgia@ncgia.ucsb.edu

State University of New York
301 Wilkeson Quad, Box 610023
Buffalo NY 14261-0001
Office (716) 645-2545
Fax (716) 645-5957
ncgia@ubvms.cc.buffalo.edu

University of Maine
348 Boardman Hall
Orono ME 04469-5711
Office (207) 581-2149
Fax (207) 581-2206
ncgia@spatial.maine.edu

PREFACE

This report was prepared for the National Center for Geographic Information and Analysis in con-junction with Research Initiative No. I on the accuracy of spatial data bases. It is being distributed as an aid to research.

The report consists of an annotated bibliography of articles related to the issue of spatial data base accuracy. Approximately one-half of these articles have been published in the professional literature. Conference papers account for the largest share of the remainder, which also includes technical reports, chapters of books and unpublished manuscripts.

Keywords are used to identify the primary topical focus of each article. Topics were selected to reflect a set of broad issues of interest to those involved in both pure and applied research. The latter pages of the report contain a keyword index, an author index and a brief description of each keyword.

Due to the usual constraints, it has been impossible to compile a thoroughly exhaustive listing of relevant articles. Readers will no doubt find that certain topics have not received the attention they deserve. Moreover, topics such as accuracy in surveying, for which a fairly mature and well-known professional literature exists, have not been dealt with in a comprehensive fashion. These shortcomings notwithstanding, the bibliography succeeds in bringing together a range of literature from a diversity of sources and should therefore prove to be a useful starting point for future research on the issue of spatial data base accuracy.

TABLE OF CONTENTS

ANNOTATED BIBLIOGRAPHY

1. Adejuwon, 0. (1975). A note on the comparison of chorochromatic surfaces. *Geographical Analysis*, 7(4), 435-440.
A method is described for comparing two thematic maps of the same area, based on the probability that different cover types from each map will occur at the same location. An index is derived that measures the degree of locational association for given cover types. This index may be used to compare alternate versions of the same map or compare one map against an accurate standard. [**Classification accuracy**]

2. Alonso, W. (1968). Predicting best with imperfect data. *American Institute of Planners Journal,* 34(4), 248-255.
A model is presented for assessing error propagation for simple arithmetic operations. Error propagation is shown to be a function of the arithmetic operator applied and the degree of correlation between parameters. As model complexity rises, error associated with model misspecification declines, but the effects of error propagation increase. Model complexity should reflect the reliability of the data. The optimal model is the one for which the sum of propagation and specification error is at a minimum. Strategies are proposed for reducing error propagation and techniques are presented for computing the marginal costs of obtaining more accurate parameter estimates. [**Geographic modeling**]

3. American Society of Civil Engineers (Committee on Cartographic Surveying, Surveying and Mapping Division). (1983). *Map uses, scales and accuracies for engineering and associated purposes*. New York, NY: American Society of Civil Engineers.
This manual details the American Society of Civil Engineers standards for selecting appropriate map types, scales and accuracies for engineering and planning purposes. The manual presents a means of determining the contour intervals and map scales best suited to the various stages of specific projects. Tables show the relationship between map scale and contour interval, and parameters such as map use, project stage, land use and topography. A method is presented for determining whether ground survey is required to meet accuracy requirements. An alternative to the US National Map Accuracy Standards (NMAS) is proposed for assessing positional accuracy, based on threshold levels of standard and absolute error. The relationship between this proposed standard and the NMAS is explained and examples of accuracy testing procedures are provided. Issues of data quality associated with map symbology and data availability are also addressed. [**Planimetric accuracy; Topographic maps**]

4. American Society of Photogrammetry (Committee for Specifications and Standards, Professional Practice Division). (1985). Accuracy specification for large-scale line maps. *Photogrammetric Engineering and Remote Sensing*, 51(2), 195-199.
An accuracy standard is proposed for large-scale (~! 1:20 000) maps. Accuracy is expressed in terms of the standard error, maximum error, circular map accuracy and vertical map accuracy. Accuracy testing is designed to indicate nominal accuracy equal to or better than the allowable error. Statistical tests are designed to assess both bias and precision. Map accuracy information is to be provided on the map sheet. (**Planimetric accuracy; Topographic maps**]

5. Amrhein, C. & Griffith, D. (1988). *GISs, spatial statistics, and statistical quality control*. Unpublished manuscript.
Error in digital representations of cartographic lines includes both digitization and generalization error. Digitization error results from inaccurate placement of the digitizer cursor over the true line. Generalization error refers to the deviations from the true line that result from the representation of this line as a set of points. Digitization error may be modeled using time-series methods since errors are likely to be serially correlated. Generalization error may be measured in terms of the deviations between the digitization-error-free digitized line and the true line described by a mathematical function. As the number of digitized points increases, digitization error rises while generalization error approaches zero. The point at which the two sources of error are equal determines the optimal number of points to digitize. [**Digitizing error; Generalization error]**

6. Anselin, L. (1988). *What is special about spatial errors?* Paper presented at First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis, Montecito, California.
This paper considers some of the distinctive aspects of spatial errors. Spatial aspects of specification error include location-specific and scale- and time-dependent phenomena. Spatial aspects of measurement error include spatial heterogeneity and dependence. Spatial statistics may be applied if error is viewed as information rather than a nuisance. Spatial statistics are relevant for estimation and prediction and for model validation. For model validation, important issues include summary versus partitive measures of accuracy, the trade-off between parametric and robust approaches, the nature of the decision making environment and the cost of accuracy assessment. [**Geographic modeling**]

7. Arbia, G. (1988). *Statistical effects of spatial data transformations*. Paper presented at First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis, Montecito, California.
The effects of aggregation of areal units may be addressed in terms of the spatial data configuration, which defines the relationship between pre-aggregated and post-aggregated processes. This approach gives an indirect measure of spatial autocorrelation and provides measures of the loss of accuracy with aggregation as a function of the aggregation criterion. [*Geographic modeling*]

8. Aronoff, S. (1982). Classification accuracy: A user approach. *Photogrammetric Engineering and Remote Sensing*, 48(8), 1299-1307.
Two types of risk are associated with hypothesis tests for thematic map accuracy. Consumer's risk refers to the probability of accepting of an inaccurate map, while producer's risk refers to the probability of rejecting a map of sufficient accuracy. Traditional hypothesis tests attempt to minimize consumer's risk at the expense of producer's risk, and thus acceptable maps often fail to pass the test. An alternate test is presented that incorporates both types of risk. [**Classification accuracy**]

9. Aronoff, S. (1982). The map accuracy report: A user's view. *Photogrammetric Engineering and Remote Sensing*, 48(8), 1309-1312.
This paper outlines a presentation format for thematic map accuracy tests that incorporates measures of overall map accuracy, consumer's risk, producer's risk, minimum map accuracy and individual cover type accuracies. The purpose of this format is to facilitate map accuracy interpretation for a variety of applications. [**Classification accuracy**]

10. Aronoff, S. (1984). An approach to optimized labeling of image classes. *Photogrammetric Engineering and Remote Sensing*, 50(6), 719-727.
The accuracy of thematic maps derived from remotely sensed data is examined in the context of a loss function. This function accounts for variations in sample size for different cover types and the relative significance of misclassification for different cover types. A classification algorithm is presented that minimizes the loss function. The loss function for different classifications can also be compared to identify the relative success of different classification systems. [**Classification accuracy**)

11. Aronoff, S. (1985). The minimum accuracy value as an index of classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 51(1), 99-111.
The minimum accuracy value is posited as one measure of thematic map accuracy. This value is defined as the lowest expected accuracy for a map given an observed number of misclassifications in a sample and an acceptable level of consumer's risk. The minimum accuracy value reflects the level of uncertainty associated with sample size and is therefore useful for comparing accuracy tests based on samples of different sizes. [**Classification accuracy**]

12. Barrett, A. N. & Rhind, D. W. (1974). Retrospective application of boundary constraints in automated contouring by digital spatial filtering. *Mathematical Geology*, 6(l), 83-89.
Most automated contouring procedures assume that no horizontal or vertical error exists in the coordinates of control points. This paper describes a contouring method that accounts for the presence of error. Undesirable values identified from Fourier analysis are deleted and a new field is constructed from the inverse Fourier transform. Results reveal improvements in fit and trend representation compared to standard contouring methods. [**Isometric mapping**]

13. Barrett, J. P. & Philbrook, J. S. (1970). Dot grid area estimates: Precision by repeated trials. *Journal of Forestry*, 68(3), 149-151.
This study examines the precision of area estimates derived from dot planimetry. The mean and standard error of area estimates may be obtained from repeated random placements of the dot grid. This is contrasted with the traditional method based on a single placement of the dot grid. Experimental results indicate that the traditional method yields inflated standard errors some three to five times larger than those obtained with the repeated trial technique. Results also indicate that standard errors decrease

with an increase in dot density. [**Area estimation**]

14. Bascomb, C. L. & Jarvis, M. G. (1976). Variability in three areas of the Denchworth soil map unit: L Purity of the map unit and property variability within it. *Journal of Soil Science*, 27(3), 420-437.
This study examines the variability of soil properties along profiles within a single mapping unit. The purity of the mapping unit, estimated as the percentage of profiles conforming to the definition of the soil series, was found to be on the order of 60 percent. The variability of individual soil properties, measured by the coefficient of variation, differed widely but was greater for chemical properties than physical ones. Since mapping units are designed to identify parcels of land that require similar management practices, the observed variations in soil properties are important in assessing the confidence with which different management practices can be applied. [**Soil properties**]

15. Battle, J. E. (1977). Coding and interpolation of topographic surfaces. *Proceedings, American Congress on Surveying and Mapping*, 288-300.
Interpolation accuracy is dependent on the weighting function and sampling method employed. The weighting function must reflect the nature of the underlying surface. Surface-specific sampling will often yield better results than grid sampling, due to the inclusion of important inflection points. In a set of experimental trials, variations in the weighting function produced small but systematic changes in accuracy. Grid sampling was observed to produce a more accurate interpolated surface than surface-specific sampling, but the frequency of large errors was higher for grid sampling. Grid orientation was also observed to have a slight impact on interpolation accuracy. [**Isometric mapping**]

16. Bauer, M. E., Hixson, M. M., Davis, B. J. & Etheridge, J. B. (1978). Area estimation of crops by digital analysis of Landsat data. *Photogrammetric Engineering and Remote Sensing*, 44(8), 1033-43.
Classification error may cause area estimates derived from pixel counts on satellite imagery to be inaccurate. This study proposes a method for correcting area estimates to account for classification error. This method involves multiplication of the normalized vector of area estimates by the inverse of the transposed classification error matrix. 'The variance of area estimates may also be computed and used to test hypotheses about the accuracy of area estimates. [**Area estimation; Classification accuracy**]

17. Baugh, 1. D. H. & Boreham, I R. (1976). Measuring the coastline from maps: A study of the Scottish mainland. *The Cartographic Journal*, 13(2), 167-171.
This study investigates errors in estimates of coastline length obtained from maps. Two methods of estimating length are compared. The first employs a mechanical wheel and the second involves summation of the distances between successive digitized points. Both methods yield larger estimates of coastline length as map scale increases, especially for highly convoluted coastlines. The digital approach yields more accurate estimates and, with the exception of relatively straight coastlines, also yields larger estimates. [**Generalization error; Length estimation**]

18. Beard, K. (1989). Use error: The neglected error component. *Proceedings, Auto Carto 9*, 808-817.
Commonly recognized map errors include those associated with data collection and representation, or source errors, and those associated with data manipulation, or process errors. However, it is also important to consider use errors, or errors associated with the misinterpretation and misapplication of maps. Use errors may result when map quality documentation is inadequate, symbology is unconventional or maps are outdated, highly generalized or of small scale. Use error may be reduced with GIS technology, which allows data to be stored in greater detail and at higher levels of precision, affords a means of documenting data quality and facilitates the development of methods for updating data and preventing illegal and illogical operations. [**Cartographic communication**]

19. Beckett, P. (1977). Cartographic generalization. *The Cartographic Journal*, 14(l), 49-50.
Cartographic generalization implies that a decrease in map scale will be associated with a decrease in the sinuosity of linear features and hence in estimates of line length. Experimental results indicate that the length of a line measured off a map of scale LRI is approximately (R2 / R,)" of its length measured off a map of scale R2. For measurements in kilometers, the value of -y may vary between 0 and 0.4. Results suggest a value of 0.017 for British road maps. [**Generalization error; Length estimation**]

20. Beckett, P. H. T. & Burrough, P. A. (1971). The relation between cost and utility in soil survey: IV. Comparison of the utilities of soil maps produced by different survey procedures, and to different scales. *Journal of Soil Science*, 22(4), 466-480.
This study compares the utility of a number of single-property and general-purpose soil maps produced by free and grid survey at scales ranging from 1:20 000 to 1:70 000. The proportions of different soil series on maps were found to be better estimated from grid than from free surveys. The purity of series mapping units was found to increase with survey effort and was highest when soil boundaries had clear external expression. Average purity was as high for grid as for free survey. Uniformity of soil

properties within mapping units was measured by the ratio of the within-unit variance to total variance. Free survey fared better at scales between 1:100 000 and 1:30 000, especially when soil boundaries had clear external expression. Reduction in the variability of soil properties requires special purpose or single-property maps of relatively large scale. [**Soil properties**]

21. Beckett, P. H. T. & Burrough, P. A. (1971). The relation between cost and utility in soil survey: V. The cost-effectiveness of different soil survey procedures. *Journal of Soil Science*, 22(4), 481-489.
This study compares the utility of a number of single-property and general-purpose soil maps produced by free and grid survey at scales ranging from 1:20 000 to 1:70 000. Utility may be defined in terms of the variability of soil properties of practical significance within mapping units. Soil series maps made by grid survey were found to have greater utility than those made by free survey, particularly at scales greater than 1:50 000. There is a ceiling on the utility of general-purpose soil series maps reached at a scale of approximately 1:25 000. This limit may be surpassed only by relatively large scale single-purpose maps made by grid survey. [**Soil properties**]

22. Bedard, Y. (1986). A study of the nature of data using a communication-based conceptual framework of land information systems. *The Canadian Surveyor*, 40(4), 449-460.
The development and implementation of a land information system (LIS) data base is examined from the context of a communication model, where the real world is repeatedly filtered in the construction of cognitive and physical models by data collectors, technological systems and pro-duct users. Uncertainty is introduced whenever one of these models is constructed. Uncertainty is associated with limitations inherent in modeling and the participants involved in the communication process. Uncertainty may be ignored, reduced or absorbed. Reduction of uncertainty occurs when standards of modeling are established. Absorption occurs when the model maker guarantees the model or the user accepts the consequences of using a model that is not guaranteed. [**Cartographic communication**]

23. Bedard, Y. (1986). Comparing registration systems using a communication-based criterion: Uncertainty absorption. *XVIII Congress, International Federation of Surveyors, Commission* 7, 565-579.
The development of land information system (LIS) data bases involves the construction of cognitive and physical models of reality. Uncertainty results from the limitations inherent in model building, including issues of technical accuracy, semantic precision and communication efficiency. Uncertainty may be reduced but not eliminated, and thus some uncertainty must be absorbed by the information producer and user. It is possible to reduce uncertainty associated with the existence of rights to land, the spatial location of these rights and the internal and external validity of property surveys. Uncertainty absorption refers to the guarantee of land title and specific property boundaries. [**Cartographic communication**]

24. Bedard, Y. (1987). Uncertainties in land information systems databases. *Proceedings, Auto Carto 8*, 175-184.
This study examines uncertainty in land information system (LIS) data bases in the context of a communication model. Uncertainty may be introduced into LIS data bases whenever a model of the real world is created. This occurs whenever information is observed by data collectors, encoded into the computer or decoded by users. Uncertainty results from inherent limitations of the models created at these stages and the limitations of model builders. The types of uncertainty that may result include conceptual (or fuzziness), descriptive (imprecision in attribute values), locational (imprecision in positional features) and meta-uncertainty. Uncertainty may be reduced by establishing modeling and communication standards, or absorbed by the users or producers of the information. [**Cartographic communication**]

25. Bennett, H. C. (1977). The geographic data base: Reliability or chaos? *Proceedings of the American Congress on Surveying and Mapping 37th Annual Meeting*, 675-680.
Geographic data bases must be based on a framework of accurately positioned field points. A detailed mapping program must be designed to establish standards for accuracy, scale, format, data type and symbolization. Without careful planning and maintenance of high accuracy standards the utility of the data base will suffer. [**Planimetric accuracy**]

26. Bethel, J. S. & Mikhail, E. M-. (1983). On-line quality assessment in DTM. Technical Papers, *American Congress on Surveying and Mapping-American Society of Photogrammetry Fall Convention*, 576-584.
This paper presents a technique for detecting gross errors in digital elevation models based on mathematical modeling of the terrain surface with tensor-product B-splines. Error detection is achieved by examining the residual values for a sample of points after surface modeling. Experiments performed on a number of synthetic surfaces with known mathematical functions indicate that this technique is most effective in the case of multiple errors of large magnitude. [**DEMs**]

27. Blakemore, M. (1983). Generalization and error in spatial data bases. *Cartographica*, 21(2/3), 131-139.
[See also: Blakemore, M. (1983). Generalization and error in spatial data bases. *Proceedings, Auto Cart 6*, 313-322.1
Inaccuracies in the positions of cartographic features can have serious implications for geographical modeling. Polygon

boundaries may be inaccurate due to imprecision in measurement,digitizing error, the artificiality of many boundaries and high levels of generalization for sinuousfeatures. The epsilon band concept may be used to assess the effects of error in polygon boundaries. Experiments were conducted on a point-in-polygon algorithm using epsilon bands ofvarious widths to represent different levels of positional error in polygon boundaries. Resultsindicate that even a modest amount of error can produce a high percentage of points that cannot be uniquely assigned to an individual polygon. The ability to assign points to polygons was observed to be affected by polygon size and shape, the proximity of points to polygon boundaries and the width of the epsilon band. [**Digitizing error; Generalization error**]

28. Blakemore, M. (1985). High or low resolution? Conflicts of accuracy, cost, quality and application in computer mapping. *Computers and Geo8cience8*, 11(3), 345-48.
Enthusiasm for the display capabilities of microcomputers has caused practitioners to ignore more fundamental aspects of the nature of spatial data in a computer environment. Data quality issues have been ignored despite the existence of published map accuracy standards by all major conventional mapping agencies. The low resolution display devices common to most microcomputers has caused an erosion of concern for data quality. Technology has not responded to data quality concerns. Rather, technological limitations have shaped data quality issues, as users have simply been forced to cope with the crude nature of display devices. [**Data quality issues**]

29. Blakney, W. G. G. (1968). Accuracy standards for topographic mapping. *Photogrammetric Engineering*, 34(10), 1040-1042.
One body of opinion suggests that the US National Map Accuracy Standards should be made more rigorous. This paper defends the current standards on the basis of the additional cost of more rigorous standards. Moreover, it is argued that if the standard error is to be adopted as an index of map accuracy, it should be used in conjunction with the maximum error estimate rather than replacing it altogether. [**Topographic maps**]

30. Braile, L. W. (1978). Comparison of four random to grid methods. *Computers and Geosciences*, 4(4), 341-349.
This study examines the relative accuracies of four spatial interpolation methods (weighted average of closest points, weighted average of the three closest points with azimuth control, local polynomial surface fitting and piecewise cubic polynomial interpolation along profiles). These methods were used to interpolate an aeromagnetic map based on the values at randomly selected points. Difference maps were constructed from the deviations between actual and interpolated values at grid intersection points. Root mean squared error was observed to be similar for all methods except polynomial surface fitting, for which it was significantly lower. None of the methods was able to adequately reproduce high-frequency variations. Minima and maxima were interpolated poorly by the two weighted average methods. Edge effects, while apparent on all four difference maps, were smaller for the local polynomial surface fitting method. [**Isometric mapping**]

31. Brusegard, D. & Menger, G. (1988). Real data with real Problems: Dealing with large spatial databases. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Many agencies develop and maintain large geographic data bases for use in market and social research applications. This paper details an number of conceptual and practical problems encountered in developing accurate data bases. Accuracy concerns include the integration of data from different sources and techniques for aggregation and disaggregation, map overlay and point-in-polygon searches. [**Data quality issues**]

32. Burrough, P. A. (1983). Multiple sources of spatial variation in soil: L The application of fractal concepts to nested levels of soil variation. *Journal of Soil Science*, 84(3), 577-597.
Variations in soil properties are often considered to be composed of functional variation that can be explained and random variation that is un-resolvable. This distinction is scale-dependent since an increase in scale often reveals structure in random variations. The nested, autocorrelated and scale-dependent nature of unresolved variation can be taken into account with a fractal model. Experimental results suggest that soil does not behave exactly as a Brownian fractal because variation is controlled by many independent processes causing abrupt transitions in space. [**Soil properties**]

33. Burrough, P. A. (1983). Multiple sources of spatial variation in soil: II. A non-Brownian fractal model and its application in soil survey. *Journal of Soil Science*, 34(3), 599-620.
This study proposes a non-Brownian, nested model to account for soil variability as a series of superimposed, independent, soil-forming processes acting at different scales. The model is explained and theoretical examples are given of its semivariograms and confidence limits. The ability to identify structure in variations that appear to be random at smaller scales allows more precise statements to be made about soil properties. [**Soil properties**]

34. Burrough, P. A. (1986). *Principles of geographical information systems for land resources assessment.* Oxford: Clarendon.
Chapter 6 ("Data quality, errors, and natural variations") addresses many issues pertaining to the accuracy of spatial data bases.

The issues reviewed in this chapter include obvious sources of error (age of data, areal coverage, scale effects, sampling density, limitations of surrogate variables, data formats, and cost and accessibility factors) and errors resulting from natural variations (positional and attribute accuracy and the effects of spatial variations). Errors arising through data processing are also addressed, including numerical round-off error, digitizing error and errors resulting from vector to raster conversion and map overlay. The indeterminacy of many natural boundaries, particularly soil boundaries, is discussed with reference to fractal theory and the statistical nature of boundaries. Propagation of * errors in attribute values is examined for both arithmetic and logical operations. [**Digitizing error; Generalization error; Geographic modeling; Map overlay; Soil properties; Vector to raster conversion]**

35. Burrough, P. A. (1987). Multiple sources of spatial variation and how to deal with them. Proceedings, *Auto Carto 8*, 145-154.
The paper reviews some methods for dealing with multiple sources of spatial variation. Spatial interpolation methods must be capable of accounting for variation at different spatial scales. What appears to be random variation at one scale may be resolvable at another. Optimal methods of interpolation, such as kriging, recognize that variation may be due to locally random but spatially correlated components. Classes of the semivariogrарn model used in kriging include transitive models, in which the semivariance appears to reach a plateau at a certain lag, and unbounded models, where the semivariance continues to rise as the lag increases. The choice of an appropriate model is critical to the success of kriging. A multi-scale model is presented for use in a one-dimensional case. The model assumes that randomly correlated variation can exist at all scales. [**Soil properties**]

36. Burtch, R. & Thapa, K. (1989). Developing accuracy standards for multipurpose cadastral systems. *Technical Papers, American Society of Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping Annual Convention*, 5, 67-75.
This paper reviews the accuracy needs for users of multipurpose cadastral systems and presents some criteria for devising meaningful accuracy standards. Existing map accuracy standards are reviewed and compared, including the US National Map Accuracy Standards and the standards proposed by the Federal Geodetic Control Committee, the American Society of Photogrammetry and Remote Sensing, the American Society of Civil Engineers, the American Congress on Surveying and Mapping and the American Land Title Association. Developing accuracy standards for multipurpose cadastral systems is difficult because the data are used in a myriad of applications by a variety of users. Accuracy requirements vary considerably depending on the needs of the user. Accuracy must therefore be at a level acceptable to the user with the most stringent requirements. Due to the costs of obtaining accurate information, users must realistically evaluate their accuracy requirements. [**Planimetric accuracy**]

37. Campbell, J. B. (1977). Variation of selected properties across a soil boundary. *Soil Science Society of America Journal*, 41(3), 578-82.
This study examines changes in soil properties across a boundary between two soil series. Variations in sand, silt and pH measurements are compared with three statistical models of variation (a random distribution, a trend surface and an abrupt boundary model). Results indicate that pH and silt measurements change slowly while sand measurements change abruptly at boundaries. These results suggest that different models of spatial variation may be required for different soil properties. [**Soil properties**]

38. Campbell, J. B. & Edmonds, W. J. (1984). The missing geographic dimension to soil taxonomy. *Annals of the Association of American Geographers*, 74(l), 83-97.
The US Soil Conservation Service's Soil Taxonomy contains a number of unique concepts to aid in soil classification and mapping, but the taxonomy is unable to adequately portray the spatial distribution of soils. The logical structure of taxonomy presents a problem for mapping because its primary focus is taxonomic rather than geographic. In theory, a mapping unit should correspond to a taxonomic unit, but limits on cartographic fidelity ensure that mapping units are never perfectly pure. Mapping units typically contain impurities associated with natural variations in soil-forming factors (continuous variation) and unlabelled patches of other soils (discontinuous variation). Application of the taxonomy may yield maps that are unreliable due to high pedological diversity over short distances. The often complex character of the spatial distribution of soil may more effectively be conveyed to map users by documenting mapping unit purity, a method applied in older (pre-1935) US soil surveys and modern surveys in some European countries. Alternate criteria for delineating mapping units are discussed that account more explicitly for the spatial distribution of soil. (**Soil properties**]

39. Card, D. H. (1982). Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing*, 48(3), 431-439.
This paper presents a method for more precisely assessing thematic map accuracy. Given the marginal proportions for the mapped cover types in a classification error matrix, the marginal proportions for the actual cover types may be estimated. These may be used in turn to improve estimates of the accuracy of individual cover types. Formulae are derived for obtaining these improved estimates and their asymptotic variances for both random and stratified-random samples. A numerical example is presented to illustrate the required computations. [**Classification accuracy**]

40. Carstensen, L. W., Jr. (1987). A measure of similarity for cellular maps. *The American Cartographer*, 14(4), 345-358.
This study examines the k coefficient as a means of comparing map patterns. Test subjects' subjective rankings of map similarity were found to be in agreement with le coefficient values. The k coefficient is proposed as a general index for measuring the agreement between maps and may also be used in assessing map accuracy. [**Classification accuracy**]

41. Carter, J. R. (1989). Relative errors identified in USGS gridded DENIS. *Proceedings, Auto Carto 9*, 255-265.
This paper discusses the nature of relative errors in US Geological Survey gridded digital elevations models (DENIS). Relative errors occur when an elevation value is inconsistent with neigh-boring values which as a group give an adequate definition of the form of the land surface. In contrast, global errors occur when the general form of the surface is adequately defined, but the total model departs significantly from the actual surface, Specific examples of relative errors occurring in both Gestalt Photomapper II and digital line graph DENIS are presented. This discussion of error is compared to the Geological Survey taxonomy of random errors, systematic errors and blunders. [**DEMs**]

42. Caruso, V. M. (1987). Standards for digital elevation models. *Technical Papers, American Society of Photogrammetry and Remote Sensing-American Congress on Surveying and Mapping Annual Convention*, 4, 159-166.
Standards for US Geological Survey digital elevation models (DEMs) are intended to facilitate DEM collection, editing, exchange and quality control. This paper reviews these standards, including basic DEM characteristics, geometry, logical format, errors and quality control. Errors in DEMs include blunders and random and systematic errors. Blunders are errors exceeding the maximum permissible error. Random errors result from measurement error and, unlike systematic errors, do not introduce bias. DEM accuracy is defined in terms of horizontal and vertical accuracy. [**DEMs**]

43. Chrisman, N. R. (1982). A theory of cartographic error and its measurement in digital data bases. *Proceedings, Auto Carto 5*, 159-168.
Positional errors in cartographic lines may arise from digitizing error, numerical round-off error and the width of lines on paper maps. The epsilon band concept may be used to model this error. This band represents an estimate of the positional error associated with a line. Large areas of uncertainty may result on a map even when the epsilon band is quite narrow. In an experiment performed on a US Geological Survey land use/land cover map, an epsilon band 20 m wide yielded an area of uncertainty greater than 7000 ha, or over 7 percent of the total map area. [**Digitizing error**]

44. Chrisman, N. R. (1982). Beyond accuracy assessment: Correction of a misclassification. *Proceedings, International Society for Photogrammetry and Remote Sensing Commission IV*, 123-132.
Common methods of thematic map accuracy assessment, such as the percentage of sample points correctly classified, rest on shaky statistical foundations and provide ambiguous results. The x statistic, which corrects for random correct classifications, is proposed as an alternate index of accuracy. The accuracy of thematic maps may be improved by correcting for classification error observed in a sample based on simple manipulation of the classification error matrix. Results from published data show significant changes in area estimates for different cover types following correction of classification errors. [**Area estimation; Classification accuracy**]

45. Chrisman, N. R. (1983). Epsilon filtering: A technique for automated scale changing. *Technical Papers of the 43rd Annual Meeting of the American Congress on Surveying and Mapping*, 322-331.
This paper describes an approach to automated map generalization based on the epsilon band concept. A filtering method is described that applies geometric criteria to reduce the number of points representing linear cartographic features. The method has been implemented in the WHIRLPOOL program of the ODYSSEY system. This program may be used to delete spurious polygons created in map overlay, by defining line intersections according to a tolerance level corresponding to the width of the epsilon band. [**Generalization error; Map overlay**]

46. Chrisman, N. R. (1983). The role of quality information in the long-term functioning of a geo-graphic information system. *Cartographica*, 21(2/3), 79-87.
[See also: Chrisman, N. R. (1083). The role of quality information in the long-term functioning of a geographic information system. *Proceedings, Auto Carto 6*, 303-312.1
Information on data quality must be incorporated into geographic information systems along with the data themselves. Information on data quality should enable users to make informed judgements about the appropriateness of the data for a given application. There are several important components of data quality. Lineage refers to the history of the data, including sources, editing changes and other transformations. Completeness refers to the criteria and definitions used in data classification and map construction. Positional and attribute accuracy refer to the accuracy of the spatial and aspatial components of the data base, respectively. Attribute accuracy may be assessed with logical consistency checks. Positional accuracy may be assessed by deductive estimates, internal evidence tests or independent confirmatory evidence. Geodetic control is important for ensuring positional accuracy, but a GIS must incorporate a means of adjusting encoded coordinate values. Temporal changes in data should also be incorporated into information systems. [**GBF standards**]

47. Chrisman, N. R. (1987). Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. *International Journal of Geographic Information Systems*, 1(3), 265-277.
The merits and shortcomings of different digitizing procedures are presented as a tutorial for those involved in map digitizing projects. Factors affecting digitizing accuracy and efficiency at various stages in the digitizing process are discussed, including data sources, personnel, equipment, digitizing methods and data structures. Procedures are defined for detecting and correct-ing errors, including visual inspection and topological consistency checks. The relative advantages of scanning technology are also discussed. [**Digitizing error**]

48. Chrisman, N. R. (1987). The accuracy of map overlays: A reassessment. *Landscape and Urban Planning*, 14(5), 427-439.
The reliability of composite maps produced by map overlay is affected by positional error, attribute error and error in assembling the overlay. Positional error arises from inaccuracies in the measured locations of features, as well as fuzziness due to the indeterminacy of certain types of boundaries. Attribute error includes identification error, or error associated with the assignment of cover types, and discrimination error, or error in separating adjacent cover types. Errors associated with the assembly of the overlay, such as the generation of spurious polygons, can be reduced by merging data from different sources and selecting the most accurate representation of the same cartographic features. Careful application of map overlay can yield significantly lower levels of error than theoretical predictions would suggest. [**Map overlay**]

49. Chrisman, N. R. (1988). Modeling error in overlaid categorical map8. *Paper presented at First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
This study focuses on the nature of statistical statements of error and their relevance to geographic information systems, with particular reference to the overlay of categorical coverages. Categorical coverages are categorical maps for which a system of classification determines the locations of polygon boundaries. Previous research on models of spatial error is reviewed, including contributions from surveying, remote sensing and stochastic and fractal geometry. Other important models include the epsilon band model and the phase space model proposed by Goodchild & Dubuc (1987). Errors in categorical coverages may be attributed to positional and attribute error, and an error model is proposed that is based on a set of stochastic processes operating simultaneously at specific scales. [**Map overlay**]

50. Chrisman, N. R. (1989). Error in categorical maps: Testing versus simulation. *Proceedings, Auto Carto 9*, 521-529.
Error in categorical maps may be modeled via stochastic simulation, but empirical tests provide a more useful indication of data quality. For categorical coverages, or categorical maps for which a system of classification determines the locations of polygon boundaries, positional and attribute errors may be difficult to disentangle. A comprehensive test based on the overlay of two categorical maps intended to be the same can provide estimates of both positional and attribute accuracy as well as evaluate scale effects. The framework for a mathematical model is proposed in which the total error is decomposed into a set of stochastic processes operating simultaneously at specific scales. [**Map overlay**]

51. Clarke, K. C. (1985). A comparative analysis of polygon to raster interpolation methods. *Photogrammetric Engineering and Remote Sensing*, 51(5), 575-582.
This study examines the errors resulting from the vector to raster conversion operation. Five surfaces of varying complexity were generated and mapped onto six sets of polygons of varying size and shape. These maps were then used as inputs to five vector to raster conversion algorithms. Errors were computed from the deviations between - the original and interpolated surfaces. Fourteen variables related to surface complexity and polygon size and shape were employed in a multiple regression model to predict absolute error. The predictive power of the derived regression equations was found to be high, but dropped significantly when variables describing surface complexity were excluded. Surface complexity was thus identified as the dominant factor affecting accuracy. Interpolation algorithms should be selected on the basis of the characteristics of the underlying surface. [**Vector to raster conversion**]

52. Congalton, R. G. (1988). Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 54(5), 587-592.
This study examines the effects of spatial autocorrelation on accuracy assessment methodologies for thematic maps. For three data sets of varying complexity, a difference image was constructed by comparing the actual cover types to those obtained through the classification of satellite data. The autocorrelation for each difference image was computed at various spatial lags from joint count statistics. The level of autocorrelation was observed to depend on the spatial distribution of cover types. Such analysis may assist in selecting an appropriate sampling methodology for performing an accuracy assessment. [**Classification accuracy**]

53. Congalton, R. G. (1988). A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 54(5), 593-600.

Five different sampling schemes (random, stratified random, cluster, systematic and stratified systematic unaligned) were compared in terms of their efficiency in thematic map accuracy assessment. For three data sets of varying complexity, a difference image was constructed by comparing the actual cover types to those obtained through the classification of satellite data. The mean and variance in the proportion of misclassified pixels were calculated for each difference image and served as the population parameters in the analysis. Random sampling of the difference images was observed to provide consistently adequate estimates of the population parameters, as long as the sample size was sufficiently large. Stratified random sampling also performed well, especially when cover types were small in area. For less complex data sets, systematic and stratified systematic unaligned sampling greatly over-estimated the population parameters. Cluster sampling performed adequately as long as clusters were relatively small. [**Classification accuracy**]

54. Congalton, R. G. & Mead, R. A. (1983). A quantitative method to test for consistency and correctness in photo-interpretation. *Photogrammetric Engineering and Remote Sensing*, 49(l), 69-74.

[See also: Congalton, R. G. & Mead, R. A. (1981). A quantitative method to test for similarity between photo interpreters. Technical Papers, American Society of Photogrammetry 47th Annual Meeting, 263-266.1

A method is developed for testing the agreement between different photo-interpretations. A classification error matrix was generated for different interpretations. The ic statistic was computed for each matrix and compared across interpretations. [**Classification accuracy**]

55. Congalton, R. G., Oderwald, R. G. & Mead, R. A. (1983). Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineer-ing and Remote Sensing*, 49(12), 1671-1678.

In this study, classification error matrices are analyzed with three multivariate statistical techniques. The first method involves matrix normalization, which permits direct comparison of cell values by eliminating the effects of sample size. The second method involves computation of a measure of agreement between different classification error matrices based on the /C coefficient. The third method allows for comparisons of more than two matrices based on the log-linear model, and allows for testing of factors hypothesized to affect accuracy. Numerical examples are presented using data derived from previous studies. [**Classification accuracy**]

56. Cook, B. G. (1983). Geographic overlay and data reliability. *Proceedings, United States/Australia Workshop on Design and Implementation of Computer-Based Geographic Information Systems*, 64-69.

A model is presented for identifying spurious polygons on composite maps produced by map overlay. The model is based on the assumption that the validity of a composite map polygon is a function of its size. This relationship may be expressed as a size-probability function (SPF). Methods are presented for computing the SPFs of various classes of input data. Assuming independence between input data layers, the joint SPF for the composite map can be computed as the product of the SPFs for each data layer. The model can also incorporate knowledge about the level of map generalization and its accuracy. [**Map overlay**]

57. Cook, H. R. (1976). Accuracy/quality control of digital data in the digital topographic information bank. *Proceedings of the American Congress on Surveying and Mapping Annual Meeting*,483-488.

This paper presents two general methods that may be employed in the determination of the accuracy of digital topographic data. The first is a mathematical model of the automated data production process that accounts for the errors introduced at each step and the mechanisms whereby this error is propagated. The second is a controlled sampling method using a random control net (RCN) of points with known values and associated accuracies. The computed values at each step in the automated data production process may be compared to the RCN to give a statistical appraisal of the product as it moves through the production system. [**Topographic maps**]

58. Court, A. (1970). Map comparisons. *Economic Geography*, 46(2) (supplement), 435-438.

Quantification of the level of agreement between two maps can be achieved by mapping equivalent isopleths, which are defined such that the areas between successive isopleths are equal. Two maps with the same number of equivalent isopleths may be superimposed and the areas of agreement and disagreement measured. These areas are then entered into a resemblance matrix to compute a coefficient of quantile correlation. This coefficient may be used to quantify the extent of agreement between two different maps, two versions of the same map or one map and a more accurate standard. [**Classification accuracy**]

59. Crackel, T. J. (1975). The linkage of data describing overlapping geographical units: A second iteration. *Historical Methods Newsletter*, 8(3), 146-150.
This paper discusses some deficiencies in the areal interpolation techniques presented by Markoff & Shapiro (1973). These deficiencies result from the possibility that the total area of source zone overlap for a given target zone might be larger or smaller than the target zone itself. Equations are presented that may be used if this problem is encountered. [**Areal interpolation**]

60. Crapper, P. F. (1980). Errors incurred in estimating. an area of uniform land cover using Landsat. *Photogrammetric Engineering and Remote Sensing*, 46(10), 1295-1301.
Inaccuracies in estimates of polygon area based on pixel counts may arise from errors of omission or commission within pixels at polygon boundaries. This paper develops a formula for calculating the standard error of polygon area estimates obtained from counts of Landsat pixels. Results indicate that considerable inaccuracy may occur when polygons are small or boundaries are highly contorted. [**Area estimation; Vector to raster conversion**]

61. Crapper, P. F. (1981). Geometric properties of regions with homogeneous biophysical characteristics. *Australian Geographical Studies*, 19(l), 117-124.
Rasterization of polygon maps can result in inaccuracies in polygon area estimates due to errors of omission or commission within pixels at polygon boundaries. It has been suggested that the level of inaccuracy might be estimated using the relationship between boundary length and polygon area. However, length and area are not perfectly correlated due to variations in polygon shape. Empirical results for biophysical regions indicate that as polygons become larger, their shapes become more contorted. Hence the level of accuracy depends on the polygon size distribution. [**Area estimation; Vector to raster conversion**]

62. Crapper, P. F. (1981). The relationship between region shape and error variance. *Proceedings, Second Australian Remote Sensing Conference*, 6.3.1-6.3.5.
Inaccuracies in estimates of polygon area based on pixel counts may arise from errors of omission or commission within pixels at polygon boundaries. A formula is derived that gives the variance of area estimates for Landsat pixels. The formula indicates that the relative error declines non-linearly as area increases. [**Area estimation; Vector to raster conversion**]

63. Crapper, P. F. (1984). An estimate of the number of boundary cells in a mapped landscape coded to grid cells. *Photogrammetric Engineering and Remote Sensing*, 50(10), 1497-1503.
A formula is developed for estimating the number of cells containing a polygon boundary segment in a thematic map composed of grid cells. The number of boundary cells is dependent on grid cell size, polygon shape and the frequency distribution of polygon areas. The formula facilitates computation of the standard error of polygon area estimates. [**Area estimation; Vector to raster conversion**]

64. Crapper, P. F., Walker, P. A. & Manninga, P. M. (1986). Theoretical prediction of the effect of aggregation on grid cell data sets. *Geo-Processing*, 3(2), 155-166.
This study focuses on the misclassification that occurs when thematic map polygons are converted to grid cells. Map accuracy is related to the number of boundary cells since misclassification occurs in cells located on the boundaries of polygons. A theoretical model is presented for estimating the relationship between cell size and the number of boundary cells. The model shows that the proportion of boundary cells will be greatest for small or highly contorted regions, especially when cell size is large in relation to average polygon size. [**Area estimation; Vector to raster conversion**]

65. Crosson, L. S. & Protz, R. (1974). Quantitative comparison of two closely related soil mapping units. *Canadian Journal of Soil Science*, 54(l), 7-14.
This study focuses on the validity of statistical comparisons of soil properties in different map-ping units. Empirical results indicate that even when soil properties are significantly different, many properties may not exhibit differences in modal values among mapping units. Separation of mapping units may be unreliable if based on soil properties alone. Landscape position may be used as an additional discriminating variable. [**Soil properties**]

66. Curran, P. J. & Williamson, H. D. (1985). The accuracy of ground data used in remote-sensing investigations. *International Journal of Remote Sensing*, 010), 1637-1651.
This study examines the reliability of ground data collected for remote sensing applications, including thematic map accuracy assessment. The design of a sampling scheme for collecting ground data must account for the number of ground resolution elements, the number of samples to be taken from each element, the area of each sample, the method of processing sample data and the number and training of field personnel. An empirical example illustrates that variability in ground data can be attributed to spatial variation, sample processing error and inconsistency in field interpretation. Errors- in ground data may be larger than those in remotely sensed data, especially when small samples are employed. [**Classification accuracy**]

67. Dahlberg, R. E. (1984). The public land survey system: The American rural cadastre. *Computers, Environment and Urban Systems*, 9(2/3), 145-153.
This paper examines some of the issues involved in modernizing the US public land survey sys-tem. Accuracy standards are required for the acquisition of digital coordinates from existing files, the creation of digital coordinate files from surveys and the development of the multi-purpose cadastre. Cadastral parcel data are not comparable with natural resource data, since levels of spatial and taxonomic resolution are different and many cadastral maps are not tied to a general coordinate system. The ability to merge spatial data from different sources is hampered by a lack of common standards, in part because different data sets were conceived and built prior to the advent of computerized information systems. 'Recent technological developments suggest that compatibility between parcel and natural resource data bases has begun to con-verge. [**Data quality issues; Planimetric accuracy**]

68. Dahlberg, R. E. (1986). Combining data from different sources. *Surveying and Mapping*, 46(2), 141-149.
Problems of incompatibility may occur when land information system (LIS) data bases are derived from a variety of sources. Compatibility is affected by positional accuracy, classification methods, data quality, the degree of spatial generalization, the level of spatial and taxonomic resolution, and the coordinate system, projection and data structure employed. Incompatibility between parcel and natural resource data might be solved by aggregating parcel data or acquiring Tore detailed natural resource data. Cartographic specifications and data standards should be explicit for natural resource inventories. Data quality information should be incorporated into LIS data bases. The LIS user community might also mobilize to develop more accurate base maps. Effective algorithms are needed to facilitate integration of data at various levels of generalization. [**Data quality issues; Planimetric accuracy**]

69. Dangermond, J. (1988). A review of digital data commonly available and some of the practical problems of entering them into a GIS. *Technical Papers, American Congress on Surveying and Mapping-American Society of Photogrammetry and Remote Sensing Annual Convention*, 5, 1-10.
Map accuracy may be improved at the pre-automation, automation or post-automation stages of data base compilation. A pre-automation stage technique is the identification of integrated terrain units (ITUs), or polygons possessing multiple attributes derived from several data layers. ITUs ensure consistency by avoiding the problem of alternate representations of the same cartographic feature on different data layers. An automation stage technique is templating, in which previously encoded information is reused to ensure consistency when other data layers are entered. Other examples include on-line transformation of digitized coordinates to geodetic coordinates, automated feature annotation and automatic checking of topological consistency. Post-automation techniques include manual editing, interactive rubber-sheeting, attribute consistency checking, line snapping and fuzzy tolerance operations. [**Data quality issues**]

70. Davis, F. (1988). Impacts of map errors on predictive land classification. *Paper presented at First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
The reliability of predictive land classification depends on both cartographic and ecological factors. Cartographic factors include the resolution, accuracy and bias in terrain maps. Ecological factors include the strength, consistency and scale-dependence of the relationships between the terrain and predicted variables. This paper presents some statistical techniques for assessing predictive error due to cartographic factors and provides two examples to illustrate the impact of such errors on predictive land classification. The first example uses information analysis to classify land, based on the association between mapped natural vegetation classes and categorical environmental variables. 'The second example uses a polychotomous logic model to predict site potential for vegetation. Non-random scale-dependent errors in topographic and vegetation maps were observed to lower the information content and reliability of predictive models. [**Geographic modeling**]

71. Dicks, S. E. & Lo, T. H. C. (1989). Evaluation of thematic map accuracy in a land use and cover mapping program. *Technical Paper8, American Society of Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping Annual Convention*, 4, 7-20.
This paper presents a method for assessing the accuracy of thematic maps derived from remotely sensed data. The cost associated with accuracy assessment can be minimized by exploiting the digital character of these data. This paper describes a computerized method of sample selection based on the overlay of the thematic map with a computer-generated grid. Map accuracy is assessed with binomial and x coefficients calculated from a classification error matrix. [**Classification accuracy**]

72. Digital Cartographic Data Standards Task Force. (1988). Draft proposed standard for digital cartographic data. *The American Cartographer*, 15(l).
This report contains a description of the data quality standards proposed by the Digital Carto-graphic Data Standards Task Force. Important components of quality include information on lineage, positional and attribute accuracy, logical consistency and completeness. Lineage refers to the history of the data, including sources, editing changes and other transformations. Positional and attribute accuracy refer to the accuracy of the spatial and aspatial components of the data base, respectively. Positional accuracy may be assessed with deductive or internal evidence, or with reference to a document of higher quality.

Attribute accuracy may be evaluated with deductive estimates or with tests based on independent samples or polygon overlay. Logical consistency refers to the fidelity of the relationships encoded in the data. This may be evaluated with graphical and topological tests. Completeness refers to the criteria and definitions used in data classification and map construction. [**GBF standards**]

73. Donahue, J. G. (1988). Land base accuracy: Is it worth the cost? *American Congress on Surveying and Mapping Bulletin*, 117, 25-27.
The use of different coordinate systems can affect the compatibility of data from different sources. Cartographic systems are usually based on latitude and longitude, while cadastral surveys often rely on local coordinate systems. Positional accuracy may be improved with enhanced geodetic control, but the quality and density of geodetic control depends on user requirements. Various methods exist for obtaining control points for geodetic networks, and these vary widely in cost. Obtaining more accurate geodetic control is costly but ensures long-term compatibility between data bases. Working out compatibility problems on a case-by-case basis is less expensive but is only a short-term solution. [**Data quality issues; Planimetric accuracy**]

74. Dougenik, J. (1980). Whirlpool: A geometric processor for polygon coverage data. *Proceedings, Auto Carto 4, 2,* 304-311.
The WHIRLPOOL module of the ODYSSEY system performs map overlay using a measure of tolerance for defining line intersections. This approach is useful for handling alternate versions of the same cartographic feature on different data layers and may reduce the number of spurious polygons on the composite map. Although errors larger than the defined tolerance cannot be handled, the tolerance level can be adjusted to reflect the degree of generalization present in the data. [**Generalization error; Map overlay**]

75. Dozier, J. & Strahler, A. H. (1983). Ground investigations in support of remote sensing. *In* R. N. Colwell (Ed.), *Manual of Remote Sensing* (2nd ed.) (pp. 959-986). Falls Church, VA: American Society of Photogrammetry.
This chapter focuses on methods of accuracy assessment for thematic maps derived from remotely sensed data. Errors in thematic maps may be attributed to control point location error, classification error and boundary line error. Techniques for measuring each of these types of error are reviewed and evaluated. [**Classification accuracy; Planimetric accuracy**]

76. Dutton, G. (1984). Truth and its consequences in digital cartography. *Technical Papers, Forty-Fourth Annual Meeting, American Congress on Surveying and Mapping*, 273-283.
This paper offers some new perspectives on issues of accuracy and quality standards for digital cartographic data bases. Data bases should contain data quality parameters and software should be capable of partial error estimation. There is also a need for ground control and documentation of the coordinate systems used in data bases. Data structures can be developed that would enable ground control to be self-documenting and systematically geocoded. The epistemology of space must be addressed in order to recognize how the organization of spatial data reflects specific conceptions of reality. The exact locations of certain geographical boundaries are indeterminate while others exist only by virtue of cultural or institutional consensus. [**Data quality issues**]

77. Dutton, G. (1988). Modeling locational uncertainty in hierarchical tessellations. *Paper presented at First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Maintaining coordinate data in a hierarchical tessellated framework can assist in documenting the precision with which points have been encoded. Points can also be retrieved at the level of precision required. These inherent properties of hierarchical data structures have not been addressed extensively in the literature. This paper explores some properties of hierarchical triangular tessellations, which tessellate a sphere better than rectangular tessellations and produce planar facets. [**Planimetric accuracy**]

78. Estes, J. (1988). GIS product accuracy. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
In research programs in which geographic information system (GIS) technology is applied, it is often necessary to use data for which the degree of spatial and thematic error is unknown. In evaluating different methods of establishing the accuracy of GIS output products, the ease and efficiency of a given method must be balanced against its reliability, cost and robustness. Methods combining field sampling and remotely sensed data hold considerable promise in this area. While there is a need to understand the theory of error propagation in a GIS context, there is also a need to devise methods for directly assessing the accuracy of GIS output pro-ducts. [**Classification accuracy**]

79. Fairchild, D. (1981). The effect of enumeration unit shape on isopleth map accuracy. *Technical Papers of the American Congress on Surveying and Mapping*, 423-432.
This study examines the effects of enumeration unit compactness and surface complexity on the accuracy of isopleth maps. Six enumeration unit area patterns were superimposed over three topographic surfaces of varying complexity. The average elevation value within each enumeration unit was used to interpolate an estimated surface. Accuracy was measured as the correlation between the elevation values on the actual and estimated surfaces. Accuracy was observed to be higher for more

compact enumeration units. No clear relationship was detected between accuracy and surface complexity. [**Isoplethic mapping**]

80. Fisher, P. F. (1988). Knowledge-based approaches to determining and correcting areas of unreliability in geographic databases. *Paper presented at First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Unreliability in the mapping of soils is attributable to cartographic production techniques, natural complexity and variation, human- error and temporal change. Unreliability may be portrayed with special symbols (e.g., dashed lines) and with map legends or accompanying reports that quantify mapping unit purity. Knowledge of the spatial relationships between soils and other variables that limit the distribution of soils can be used to identify and correct errors on soil maps. This approach is analogous to the dasymetric mapping method. Knowledge engineering may be applied to identify the salient relationships between soil and other limiting variables. [**Soil properties**]

81. Fitzpatrick-Lins, K. (1978). Accuracy and consistency comparisons of land use and land cover maps made from high-altitude photographs and Landsat multi-spectral imagery. *Journal of Research, US Geological Survey*, q1), 23-40.
Two methods of thematic map accuracy assessment are presented. The first compares the total area classified into each cover type for maps compiled at different scales. The second compares the mapped cover types of sample points with data obtained from ground survey. For the second method, two-way analysis of variance may be used to identify differences in accuracy between maps of different scales or among different cover types. [**Classification accuracy]**

82. Fitzpatrick-Lins, K. (1978). Accuracy of selected land use and land cover maps in the greater Atlanta region, Georgia. *Journal of Research, US Geological Survey*, 6(2), 169-173.
Thematic map accuracy may be assessed by comparing the mapped cover types for a sample of points with the actual cover types as determined from ground survey. Statistical analysis is based on the normal approximation to the binomial distribution. Accuracy is expressed as the proportion of sample points correctly classified and confidence limits are computed as a function of sample size. [**Classification accuracy**]

83. Fitzpatrick-Lins, K. (1981). Comparison of sampling procedures and data analysis for a land-use and land-cover map. *Photogrammetric Engineering and Remote Sensing*, 47(3), 343-351.
Thematic map accuracy may be assessed by comparing the mapped cover types for a sample of points with the actual cover types as determined from ground survey. Statistical analysis is based on the normal approximation to the binomial distribution. Empirical results indicate that stratified systematic unaligned sampling may under-represent certain cover types. Stratified systematic unaligned sampling followed by random sampling ensures that all cover types are adequately represented. [**Classification accuracy**]

84. Flowerdew, R. (1988). Statistical methods for inference between incompatible zonal systems. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and-Analysis*, Montecito, California.
[See also: Flowerdew, R. (1988). *Statistical methods for areal interpolation: Predicting count data from a binary variable* (Research Report No. 16). Northern Regional Research Laboratory.]
This study describes a method of areal interpolation that yields more accurate target zone estimates than more conventional approaches. This method is based on the identification of factors limiting the spatial distribution of the variable of interest, and is analogous to the dasymetric mapping method. These limiting factors may be incorporated into a Poisson regression model to estimate target zone values. An example is presented in which demographic variables for target zones are estimated from source zone values using information about political party affiliation. Results suggest substantial improvements in accuracy over the direct overlay method. [**Areal interpolation**]

85. Flowerdew, R. & Openshaw, S. (1987). *A review of the problems of transferring data from one act of areal units to another incompatible set* (Research Report No. 4). Northern Regional Research Laboratory.
This paper focuses on the data problems associated with areal interpolation. A typology of potential problems is developed based on the nature of the areal unit, the scale of measurement and the relationship between the areal unit and the variable of interest. Some solutions to these problems are also discussed. [**Areal interpolation**]

86. Frazier, B. E. & Shovic, H. F. (1980). Statistical methods for determining land-use change with aerial photographs. *Photogrammetric Engineering and Remote Sensing*, 46(8), 1067-1077.
This study is concerned with the effects of sampling methodology on the accuracy of Planimetric area estimates. Three sampling methods (random, stratified random and systematic) were used to acquire estimates of the area having undergone land use change. Similar estimates were obtained with random and stratified random methods. Accuracy was lower for systematic sampling due to periodicity in the spatial distribution. [**Area estimation**]

87. Frolov, Y. S. & Maling, D. H. (1969). The accuracy of area measurement by point counting techniques. *The Cartographic Journal*, 6(l), 21-35.
A model is developed for assessing the accuracy of polygon area estimates derived from dot or grid cell planimetry. The model is based on the possible locations of polygon boundary segments passing through individual grid cells. The standard error of area estimates is shown to be a function of polygon size. Relative error declines as polygon size or the number of grid cells increases. This relationship may be used to select an appropriate cell size for estimating the area of a given polygon. [**Area estimation; Vector to raster conversion**]

88. Galloway, R. W. & Bahr, M. E. (1979). What is the length of the Australian coast? *Australian Geographer*, 14(4), 244-247.
The accuracy of estimates of coastline length depends on map scale, measurement precision and the degree to which small islands and inlets are excluded from measurement. Coastline length estimates increase with map scale due to a decrease in the level of cartographic generalization. Length estimates increase as measurement precision rises according to a double-log relationship. The exclusion of islands and inlets may significant reduce length estimates, especially for highly indented coastlines where these features account for a large proportion of coastline length. [**Generalization error; Length estimation**]

89. van Genderen, J. L. & Lock, B. F. (1977). Testing land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*, 48(9), 1135-1137.
A statistical sampling procedure is described for testing the accuracy of thematic maps. Accuracy testing is based on the probability of finding no errors in the sample. Sample size is deter-mined on the basis of the required level of accuracy. For small samples, there is a relatively high probability of finding no errors in the sample even when the accuracy of the map is quite low. A table is presented which facilitates calculation of the minimum sample size needed to meet a required level of accuracy at a specified level of confidence. [**Classification accuracy**]

90. van Genderen, I L., Lock, B. F. & Vass, P. A. (1978). Remote sensing: Statistical testing of thematic map accuracy. *Remote Sensing of Environment*, AI), 3-14.
A method is described for testing thematic map accuracy based on the probability of finding a certain number of misclassifications in a sample drawn from each cover type. The optimal sample size for each cover type is estimated from the binomial probability density function as the minimum number of sample points that must be checked in order to meet a required level of accuracy at a specified level of confidence. The use of small samples can be misleading since the number of observed misclassifications may be quite small even if the accuracy of the map is low. [**Classification accuracy**]

91. Gering, L. R., Bailey, R. L. & Shain, W. A. (1985). Estimation of area using optimum dot-grid density with aerial photographs. *Technical Papers, American Congress on Surveying and Mapping-American Society of Photogrammetry Fall Convention*, 811-820.
This study examines how area estimates derived from dot planimetry are affected by the density of the dot grid. Empirical results indicate that different densities yield unbiased area estimates with variances that decrease as density increases. The optimal dot density for a given application reflects the trade-off between the lower variance and higher cost of data acquisition for denser grids. [**Area estimation**]

92. Gersmehl, P. J. (1981). Maps in landscape interpretation. *Cartographica*, 18(2), 79-115.
Map compilation often entails a significant loss of information due to the processes of observation, translation and cartographic representation. The level of information loss is affected by classification methods, criteria for selecting areal units for the collection and storage of data, variable scales of spatial variation in the phenomenon being mapped and the representation of continuously varying phenomena as a set of discrete areal units. Resource maps for planning and policy making have proliferated without effective quality control safeguards. In most cases the penalty for error falls on the public, not the map producer. Mapping techniques are needed that enable cartographers to convey ambiguity and uncertainty, including uncertainty bands around data values, locationally ambiguous symbols and fuzzy data classes with overlapping boundaries. [**Cartographic communication**]

93. Gersmehl, P. J. (1985). The data, the reader, and the innocent bystander: A parable for map users. *The Professional Geographer*, 87(3), 329-334.
The map producer has three primary responsibilities. The first is to ensure accurate representation of the spatial distribution being mapped. The second is to educate the map reader about the meaning of the map. The third is to protect the innocent bystander from misinterpretations made by the map reader. Standard communication models must be extended to include anticipated misinterpretations and their possible consequences. This issue is explored with reference to a set of general, coarse-resolution soil maps published by the author which, through a series of misinterpretations, eventually appeared in a planning document and were employed in a policy making context. [**Cartographic communication**]

94. Ginevan, M. E. (1979). Testing land-use map accuracy: Another look. *Photogrammetric Engineering and Remote Sensing*, 45(10), 1371-1377.
This paper presents a method of testing thematic map accuracy based on field checking of a sample of points. Statistical tests are based on the minimum and maximum acceptable accuracy and specified levels of producer's and consumer's risk. [**Classification accuracy**]

95. Goodchild, M. F. (1978). Statistical aspects of the polygon overlay problem. *In* G. Dutton (Ed.), *Harvard Papers on Geographic Information Systems*,. 6. Cambridge, MA: Laboratory for Computer Graphics and Spatial Analysis, Harvard University.
Spurious polygons may be generated in map overlay when the same cartographic feature is encoded differently on successive data layers. The generation of spurious polygons may be modeled in terms of the points of intersection between a true cartographic line and the digitized representations of this line on different data layers. The number of spurious polygons generated under worst-case and random conditions may be computed from the number of polygon vertices on each data layer. Paradoxically, the inclusion of additional vertices to improve cartographic line representation on individual data layers tends to increase the number of spurious polygons on the composite map. Some strategies for deleting spurious polygons are discussed. [**Map overlay**]

96. Goodchild, M. F. (1980). Fractals and the accuracy of geographical measures. *Mathematical Geology*, 12(2), 85-98.
This study examines the inaccuracies associated with estimates of point, line and area characteristics derived from maps. Previous research indicates that estimates of line length decrease non-linearly with map scale. Inaccuracies in estimates of area based on grid cell planimetry result from errors of omission and commission in cells located at polygon boundaries. The error variance of area estimates is dependent in part on cell size. Errors of omission and commission also affect the reliability with which the characteristics of randomly selected points within a given grid cell may be estimated. Inaccuracies in point, line and area characteristics are shown to be related to fractional dimensionality, which allows for prediction of sampling effects. [**Area estimation; Generalization error; Length estimation; Vector to raster conversion**]

97. Goodchild, M. F. (1980). The effects of generalization in geographical data encoding. *In* H. Freeman & G. G. Pieroni (Eds.), *Map data processing* (pp. 191-205). New York, NY: Academic Press.
This paper reviews methods for estimating the errors associated with the generalization inherent in cartographic representation. For raster systems, errors are associated mainly with the estimation of the area of a patch of land with a particular set of characteristics. These errors may be assessed with methods derived from dot and grid cell planimetry. The applicability of Switzer's method (Switzer, 1975) and fractals is also discussed. For vector systems, errors are associated primarily with the generation of spurious polygons in map overlay due to differences in the representations of the same cartographic feature on successive data layers. The number of spurious polygons can be modeled as a function of the number of vertices on each data layer. [**Generalization error; Map overlay; Vector to raster conversion**]

98. Goodchild, M. F. (1982). Accuracy and spatial resolution: Critical dimensions for geoprocessing. *In* D. H. Douglas & A. R. Boyle (Eds.), *Computer assisted cartography and geographic information processing: Hope and Realism* (pp. 87-90). Ottawa, Ontario: Canadian Carto-graphic Association.
This paper focuses on the implications of accuracy and spatial resolution for geoprocessing operations. Spatial data contain conventional errors, which are also present in non-spatial data, and locational errors, or errors in the positions of points, lines and areas. Locational errors result from inaccuracies in measurements of location and the inherent generalization associated with cartographic representation. The effects of accuracy and spatial resolution are discussed with reference to map overlay, vector to raster conversion and spatial interpolation. Automated geoprocessing systems tend to expose data inadequacies in a way that manual systems do not. Automation also requires the translation of subjective processes into objective ones, which may over-simplify these processes and result in error. [**Data quality issues**]

99. Goodchild, M. F. (1988). Towards models of error for categorical maps. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
In natural resource data, polygon boundaries are often derived from the data themselves rather than defined a priori. However, polygons are not necessarily homogeneous, especially at transition zones between polygons. Two models of this distortion process are proposed for raster systems. The first regards the map as truth and introduces distortion by randomizing the process that created it. The second is a stochastic approach in which pairs of maps are generated and the differences between them are regarded as distortion. These maps are generated using a set of probability surfaces describing the likelihood that a given grid cell is of a given cover type. [**Choroplethic mapping**]

100. Goodchild, M. F. & Dubuc, 0. (1987). A model of error for choropleth maps, with applications to geographic information systems. *Proceedings, Auto Carto 8*, 165-174.
In natural resource data, polygon boundaries are often derived from the data themselves rather than defined a priori. Given m

continuous variables distributed over space, the m-dimensional phase space may be divided into domains associated with a set of n classes. These domains define the function that transforms a set of rn continuous variables into a set of n classes. Since the rn variables are continuous, two polygons share a boundary in phase space if they also share a boundary on the map. This concept may be applied to simulated data, such as fractional Brownian surfaces, to examine the relationships between the accuracy of spatial data and the reliability of measurements derived from them. Sources of error in choropleth maps may also be examined and the effects of error on the output of data processing operations may be observed. [**Choroplethic mapping**]

101. Goodchild, M. F. & Lam, N. S.-N. (1980). Areal interpolation: A variant of the traditional spatial problem. *Geo-Processing*, 1(3), 297-312.
This paper focuses on the direct overlay method of areal interpolation. In matrix form this method may be represented as the product of matrices W and U. Element wjj is the area of intersection between target zone i and source zone j (normalized by the total area of either zone i or j) and element uj is the value of the variable of interest for source zone j. As many of the small, non-zero elements of W may be spurious, the degree of error in the target map is reflected in the structure of the W matrix. Another source of error is the assumption that source zones are homogeneous. Although this assumption is often unrealistic, empirical results show that in many cases target zone estimates may still be quite accurate. Other methods of areal interpolation are discussed that may yield more accurate results or are computationally less demanding. [**Areal interpolation**]

102. Goodchild, M. F. & Wang, M.-H. (1988). Modeling error in raster-based spatial data. *Proceedings, Third International Symposium on Spatial Data Handling*.
Spatial data include source errors, or discrepancies between reality and its mapped representation, and processing errors, or errors introduced by digitizing and data manipulation. This study is concerned with methods of characterizing source errors for raster systems. These methods are based on pixel probability vectors, or the probabilities with which a given pixel may be assumed to belong to each of a set of cover types. Simulated data can be generated from these probability vectors in order to model the effects of error. For example, generation of a set of realizations of the position of a line between two polygons provides a direct link between pixel probabilities and polygon homogeneity. This approach can be used to obtain information about uncertainty given an appropriate stochastic model of the process whereby cover types are obtained from probability vectors. [**Choroplethic mapping**]

103. Goodchild, M. F. & Wang, M.-H. (1989). Modeling errors for remotely sensed data input to GIS. *Proceedings, Auto Carto 9*, 530-537.
This paper is concerned with the different concepts of accuracy in the fields of remote sensing and geographic information processing. Accuracy in remote sensing is modeled using probabilities of class membership for each pixel. For vector-based geographic information systems it is modeled using concepts such as the epsilon band. The paper posits a linkage between these two views of accuracy based on stochastic process models that account for prior and posterior probabilities and levels of spatial dependence. Two such models are described. The first involves the random selection of realizations to obtain a closer match to a specified target level of spatial dependence. The second implements a spatially autoregressive model that accounts for the spatial dependence between nearby pixels. [**Choroplethic mapping**]

104. Grady, R. K. (1988). Data lineage in land and geographic information systems (LIS/GIS). *Proceedings, GISILIS '88*, 722-730.
This paper examines the importance of data lineage in evaluating the quality of spatial data bases. Information about the sources and processing history of data bases should be provided by producers, but users must also be cognizant of suitability standards that may be application-specific. An analogy exists with management information systems, where transactions that modify a data base are recorded in a transaction history file which serves as an audit trail. A data dictionary, which applies rules to transactions before new data are entered or existing data are modified, ensures referential integrity and product consistency. Deferred checking of new or transformed data can supplement such input checks. The mandate or purpose behind data collection should also be included as part of the lineage record. Lineage records are of particular importance when data from a variety of sources are to be integrated. [**GBF standards**]

105. Greenland, A., Socher, R. M. & Thompson, M. R. (1985). Statistical evaluation of accuracy for digital cartographic data bases. *Proceedings, Auto Carto 7*, 212-221.
The x statistic is used to quantify the level of agreement between manual and digital versions of soil, slope and vegetation maps. The approach consists in superimposing the manual and digital versions and computing the total area for which cover types are in agreement and disagreement. Agreement was observed to be high for all three types of maps, but dropped as spatial resolution levels declined. The approach was also applied to ordinal data by weighting the disagreement between ordinal classes as a function of their relative significance. Agreement was poor and varied widely for different weighting schemes. [**Classification accuracy**]

106. Griffith, D. (1988). Distance calculations and errors in geographic data bases. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Distance or separation measurements for network link lengths or between areal unit centroids are used in numerous applications, including minimum path calculations and location-allocation modeling. To date the errors in such measurements have largely been ignored in the literature. This paper explores some properties of this problem and studies the expected values and variances of both additive and proportional error structures for selected error distributions in the presence or absence of bias. [**Length estimation**]

107. Guptill, S. C. (1988). Inclusion of accuracy data in a feature based, object oriented data model. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
In a digital spatial data base, entities are individual phenomena in the real world, features define classes of entities and a set of objects represents a single occurrence of a feature. Spatial objects contain locational information and feature objects contain aspatial information. Objects have attributes and relationship values, which in turn also have attributes. This recursive schema allows for the inclusion of accuracy information at any level of the geographical model. Measures of locational accuracy can be assigned to each spatial object and measures of uncertainty to each attribute. Mechanisms exist in data structures to handle such accuracy data. [**Planimetric accuracy**]

108. Gustafson, G. C. (1977). Koppe errors in automatic contouring. *Proceedings of the American Congress on Surveying and Alapping* 37th Annual Meeting, 305-317.
This study examines errors in topographic map contours. A strong correlation exists between terrain gradient and root mean squared vertical error for topographic surveys based on field methods, stereoplotting techniques and computer interpolation from orthophoto data. Thus terrain gradient may significantly affect the accuracy of contouring techniques. The Koppe formula gives a good approximation of expected vertical error on any part of the topographic map, since the formula accounts for the effects of terrain gradient. [**Topographic maps**]

109. Gustafson, G. C. & Loon, J. C. (1981). Updating the national map accuracy standards. *Technical Papers of the American Congress on Surveying and Mapping,* 466-482.
For contour maps, the magnitude of a vertical error is dependent on the horizontal error of the contour and the local terrain gradient. The US National Map Accuracy Standards (NMAS) are based on compliance testing and do not express map accuracy in a statistical form. This makes map accuracy comparisons -difficult. In contrast, European accuracy standards are typically based on the Koppe formula, which incorporates vertical and horizontal error into a single formula. This formula may be used to evaluate a given map or map series with the official standard, compare the accuracy of different maps and estimate the expected error when interpolating elevations from a contour map. Empirical tests suggest that US Geological Survey topographic maps are more accurate than NMAS requirements, but less accurate than many European maps. The NMAS are outdated and no longer reflect the accuracy that topographic surveys may attain. [**Topographic maps**]

110. Hakanson, L. (1978). The length of closed geomorphic lines. *Mathematical Geology*, 10(2), 141-167.
This paper examines the accuracy of estimates of lake shoreline length derived from maps rang-ing in scale from 1:10 000 to 1:1000 000. A formula is presented which defines the functional relationship between map scale, shoreline irregularity, shoreline length and lake area. The formula affords a high degree of precision and allows for estimation of the length of any closed geomorphic line independently of map scale. [**Generalization error; Length estimation**]

111. Hannah, M. J. (1979). Error detection and rectification in digital terrain models. *Joint Proceedings of the American Society of Photogrammetry and American Congress on Surveying and Mapping* Fall Meeting, 152-164.
This paper presents a set of algorithms for detecting and correcting gross errors in grid-based digital elevation models. These algorithms are based on constraints on the allowable slope and change-in-slope in local areas around each grid point. Change-in-slope tests are performed at each point for four local neighbors and eight distant neighbors, resulting in a reliability measure for each point. Slope tests are then performed for the eight neighboring points and combined with the change-in-slope reliability measure to yield an overall reliability index. The error correction algorithm alters elevation values by minimizing local changes in slope subject to constraints on reliability and the variance of the elevations of the eight neighboring points. **[DEMs**]

112. Havens, K. A., Minter, T. C. & Thadani, S. G. (1977). Estimation of the probability of error without ground truth and known a priori probabilities. *IEEE Transactions on Geoscience Electronics*, GE-15(3), 147-152.
This paper describes two methods for estimating thematic map classification accuracy when field data are unavailable. The first method estimates the probability of error analytically, using an a posteriori density function. The second method, called the majority-rule method, is based on the identification of the dominant cover type. [**Classification accuracy**]

113. Hay, A. M. (1979). Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*, 45(4), 529-533.
Probabilistic interpretation of thematic map classification accuracy requires that sample size be sufficiently large to allow accuracy to be assessed independently for each cover type. A minimum sample size of 50 is suggested. Stratified sampling ensures that all cover types are adequately represented in the sample. [**Classification accuracy**]

114. Hay, A. M. (1988). The derivation of global estimates from a confusion matrix. *International Journal of Remote Sensing*, 9(8), 1395-1398.
This paper presents a method for manipulating a classification error matrix such that the land area encompassed by each cover type can be corrected for classification error. The ratio method usually used in this capacity is inadequate since the estimated areas may not sum to the correct total. Unlike the ratio method, the method proposed in this paper does not assume independence in classification error for different cover types. [**Area estimation; Classification accuracy**]

115. Herzog, A. (1988). Modeling reliability on statistical surfaces by polygon filtering. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
The reliability of cartographic communication is affected by errors in data collection, encoding, manipulation, analysis and visualization. Visualization ' incorporates aspects of technical methods, generalization levels and perceptual factors. Visualization is an important component of reliability since maps represent the human interface with geographic information systems (GIS). Map generalization is necessary for efficient cartographic communication. GIS packages should offer more generalization options with better control over their application. Traditional cartographic tools must be sharpened to enhance the effectiveness of communication. [**Carto-graphic communication**]

116. Hixson, M. M. (1981). Techniques for evaluation of area estimates. *1981 Machine Processing of Remotely Sensed Data Symposium*, 84-90.
This study focuses on methods of area estimation in the context of remotely sensed data. Inaccuracies in area estimates may arise from classification error. Methods for evaluating classification error must account for the effects of sampling methodology. Once a classification error matrix has been constructed, a variety of methods may be used to correct for classification error. The accuracy of area estimates may be assessed by statistical hypothesis tests or comparison with more accurate reference data. [**Area estimation; Classification accuracy**]

117. Hixson, M. M., Bauer, M. E. & Davis, B. J. (1979). Sampling for area estimation: A comparison of full-frame sampling with the sample segment approach. *1979 Machine Processing of Remotely Sensed Data Symposium*, 97-103.
The objective of this study is to evaluate the effects of sampling methodology on the accuracy of area estimates derived from pixel counts on satellite imagery. Four sampling schemes were applied in an empirical test in which samples were selected from sets of rectangular blocks of pixels called sampling units. The total number of sample pixels was held constant while sampling unit size and the number of sampling units were allowed to vary. Results indicate that less precision and greater maximum absolute bias occur in area estimates as sampling unit size increases. [**Area estimation**]

118. Honeycutt, D. M. (1986). *Epsilon, generalization and probability in spatial data bases*. Unpublished manuscript.
This study seeks to measure the positional error of cartographic lines resulting from generalization and digitizing error, build a statistical model of this error based on derived probabilities and use this model as a means of identifying spurious polygons generated in map overlay. Examination of the digitized representations of cartographic lines reveals a bimodal distribution of error around the mathematical center of each line. This may be attributed in part to the tendency for human operators to undercut and overshoot while digitizing. This error distribution may be used to calculate the probability that a random point in a digitized polygon is actually located within that polygon. Probabilities may also be computed for composite map polygons generated by the map overlay operation, which may be used to identify spurious polygons. [**Digitizing error; Generalization error; Map overlay**]

119. Hopkins, L. D. (1977). Methods of generating land suitability maps: A comparative evaluation. **American Institute of Planners Journal**, 4,9(4), 386-400.
This paper compares different methods of generating land suitability maps using map overlay. The ordinal combination method is based on the superimposition of data layers that display ordinal suitability rankings. In the linear combination method, suitability rankings are differentially weighted on each data layer according to the importance of the variable portrayed on the layer. These two methods are inappropriate if the data layers are not independent. If interdependence exists, land suitability maps may be constructed using cluster analysis, gestalt approaches and the factor combination method (in which suitability rankings are determined independently for each unique combination of cover types on the composite map). [Map over-lay]

120. Hord, R. M. & Brooner, W. (1976). Land-use map accuracy criteria. *Photogrammetric Engineering and Remote Sensing*, 42(5), 671-677.
Thematic maps include classification error, boundary line error and control point location error. The normal approximation to the binomial distribution may be used to assess classification error based on the number of misclassifications in a sample of points. Boundary line error may be assessed by calculating the level of agreement between two raster representations of the same set of cartographic features. Control point location accuracy may be assessed by constructing a binomial variable indicating whether or not a given point is correctly positioned. Error may be summarized as the root mean squared error of the lower 95 percent confidence limits of the three sources of error. [**Classification accuracy; Planimetric accuracy**]

121. Hsu, M.-L. (1975). Filtering process in surface generalization and isopleth mapping. *In* J. C. Davis & M. J. McCullagh (Eds.), *Display and analysis of spatial data* (pp. 115-129). London: Wiley.
Method-produced errors in isoplethic mapping are examined in the context of spatial filtering. Isoplethic mapping may be seen as a process in which an underlying surface is filtered to pro-duce the values for a set of enumeration areas. The differences between the original and filtered surfaces gives the residual surface. The ratio of the residual values to the original values is an index of the relative importance of the residual. This ratio may be mapped at enumeration area centroids in order to identify patterns in the resulting surface. [**Isoplethic mapping**]

122. Hsu, M.-L. & Robinson, A. H. (1970). *The fidelity of isopleth maps: An experimental study*. Minneapolis, MN: University of Minnesota Press.
Isopleth map accuracy may be defined in terms of the discrepancies between the actual and mapped values at a set of sample points. Accuracy is largely a function of the characteristics of the underlying surface, the sizes and shapes of the enumeration units and the number of control points on which the map is based. Accuracy tends to be higher for simple surfaces and when enumeration units are of uniform shape. Accuracy is also dependent on the interaction between the characteristics of the surface and the number and location of sample points. The characteristics of the surface and control points affects the quantitative aspects of isopleth maps, while the surficial aspects are affected by the size and shape of the enumeration units. [**Isoplethic mapping**]

123. Hsu, S.-Y. (1978). The national map accuracy standards in the context of mapping with imagery data. *Proceedings of the American Congress on Surveying and Mapping* Fall Convention, 202-207.
This paper examines the implications of the National Map Accuracy Standards for maps derived from satellite data. The majority of large-scale planimetric and topographic maps produced by field survey or photogrammetric methods can meet this standard. Thematic maps have traditionally not been constructed with the aim of complying with such standards. How-ever, such standards can be attained in thematic mapping if base maps are derived from planimetric or topographic maps or processed satellite imagery. Thematic maps should contain information about the accuracy of the base map and source materials used. [**Planimetric accuracy**]

124. Hudson, D. (1988). Some comments on data quality in a GIS. Technical Papers, *American Congress 'on Surveying and Mapping-American Society of Photogrammetry and Remote Sensing* Annual Convention, 2, 203-210.
Indices of data quality should be computed for all data in a geographic information system in order to provide users with an indication of data reliability. Uncertainties associated with data transformations should be reflected in the graphical representation of the transformed product. Methods for modeling uncertainty must also be developed. Data quality reports should contain information about data lineage and positional accuracy. [**GBF standards**]

125. Hudson, W. D. & Ramm, C. W. (1987). Correct formulation of the kappa coefficient of agreement. *Photogrammetric Engineering and Remote Sensing*, 53(4), 421-422.
This paper presents the correct formulation of the ic coefficient as an index of thematic map classification accuracy. Incorrect formulae have often been used in previous studies. [**Classification accuracy**]

126. Imhof, 1. (1982). *Cartographic relief presentation*. New York, NY: de Gruyter.
Chapter II ("The topographic foundation") of this book gives an overview of methods of topographic survey and describes the accuracies achievable in such surveys. Accuracy issues include the horizontal and vertical accuracy of surveyed points, the positional accuracy of edge lines and the accuracy of contour lines. Contour lines may possess geometric errors (including horizontal and vertical errors as well as errors in shape, direction, curvature, length and slope) and systematic, random and gross errors. Koppe's formula and other supplemental tests are described for assessing the accuracy of contour lines. Accuracies attainable in modern topographic surveys are presented and compared. [**Topographic maps**]

127, Jenks, G. F. (1967). The data model concept in statistical mapping. *International Yearbook of Cartography*, 7, 182-190.
The accuracy of a choropleth map can be assessed with the data model concept, in which the stepped statistical surface corresponding to the data model is sliced through by planes corresponding to choropleth class intervals. Choropleth maps

therefore represent a generalized statistical surface. The discrepancies in enumeration unit values between the data model and the generalized surface may be described as a blanket of error. In order to minimize error, the blanket must be distributed uniformly over the entire surface. In order to minimize relative error, it must be distributed such that it is thickest for those enumeration units with the highest values. [**Choroplethic mapping**]

128. Jenks, G. F. (1970). Conceptual and perceptual error in thematic mapping. *30th Annual Meet-ing, American Congress on Surveying and Mapping*, 174-188.
This paper identifies various types of conceptual and perceptual error in thematic mapping. Conceptual error results from the transformation of data and concepts into graphic form. Important components of conceptual error include symbology, the level of generalization and the ability to identify the salient features of the underlying surface. Perceptual error occurs when the map user does not precisely duplicate the concepts that the map producer intended to communicate. [**Cartographic communication**]

129. Jenks, G. F. (1976). Contemporary statistical maps: Evidence of spatial and graphic ignorance. *The American Cartographer*, 8(1), 11-19.
Deficiencies in statistical map making are identified and some strategies are suggested for improving the accuracy of such maps. Deficiencies are associated with such factors as map function, graphic language vocabulary, the relationship between data processing and the fidelity of map information, and the subtleties of graphic communication. Estimates of horizontal and vertical accuracy alone fail to incorporate all relevant factors affecting map accuracy. Accuracy in cartography involves conceptual as well as statistical concepts. [**Cartographic communication**]

130. Jenks, G. F. (1981). Lines, computers, and human frailties. *Annals of the Association of American Geographers*, 71(1), 1-10.
This paper examines the implications of human digitizing errors (logical, psychological and physiological errors) and errors associated with cartographic generalization. Different types of error arise from different digitizing methods (stream or point) and modes of generalization (point elimination or smoothing). Digitizing error may be reduced by operator training and data edit-ing. Inaccuracies associated with generalization are less well understood. Specific guidelines need to be established for performing cartographic generalization. [**Digitizing error; Generalization error**]

131. Jenks, G. F. & Caspell, F. C. (1971). Error in choroplethic maps: Definition, measurement, reduction. *Annals of the Association of American Geographers*, 61(2), 217-244.
Error in choropleth maps lowers the reliability of these maps as communication media. A choropleth map is conceptually equivalent to a stepped statistical surface that has been generalized, or sliced through by planes corresponding to choropleth class intervals. Tabular error refers to the discrepancies between the enumeration unit values for the actual and generalized surfaces. Overview error refers to the total volumetric discrepancy between the actual and generalized surfaces. Boundary error refers to inaccuracies in the representation of large differences in the values of adjacent enumeration units. Procedures are described for estimating these errors and obtaining near-optimal map accuracies. Accuracy is shown to increase at a decreasing rate as the level of generalization decreases. [**Choroplethic mapping**]

132. Keefer, B. J., Smith, J. L. & Gregoire, T. G. (1988). Simulating manual digitizing error with statistical models. *Proceedings, GIS/LIS '88*, 475-483.
This paper examines the error associated with manual, stream-mode digitizing. Digitizing error includes both line-following and line generalization error. Empirical results suggest that for stream-mode digitizing, digitizing error exhibits serial correlation due to the tendency for operators to over-compensate for line-following errors. Digitizing error may be simulated using an autoregressive model that allows the effects of different levels of error to be assessed. [**Digitizing error; Generalization error**]

133. Kellie, A. C. & Bryan, D. G. (1981). A comparison of field methods for testing the vertical accuracy of topographic maps. *Technical Papers of the American Congress on Surveying and Mapping*, 275-284.
This study compares topographic map vertical accuracy tests based on map profiles and con-strained and unconstrained random sampling. Results indicate that map profiles and con-strained random sampling do not produce significantly different estimates of map accuracy than unconstrained random sampling. [**Topographic maps**]

134. Kenk, E., Sondheim, M. & Yee, B. (1988). Methods for improving accuracy of thematic mapper ground cover classifications. *Canadian Journal of Remote Sensing*, 14(l), 17-31.
This study describes a procedure for normalizing a classification error matrix such that errors of omission and commission may more easily be assessed. Producer's risk and consumer's risk are also shown to be useful components of the map accuracy statement. [**Classification accuracy**]

135. Kennedy, S. (1988). The small number problem and the accuracy of spatial databases. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
The small number problem occurs whenever a percentage, ratio or rate is calculated for either a small geographical area with a

small denominator or a large geographical area that is sparsely populated. In either case small random fluctuations in the numerator may cause large fluctuations in the resulting percentage, ratio or rate. This paper reviews the characteristics and implications of the small number problem. [**Geographic modeling**]

136. Lam, N. S.-N. (1982). Areal interpolation using map overlay. *Modeling and Simulation*, 13, 953-959.
This paper examines the accuracy of the direct overlay method of areal interpolation. Accuracy is a function of the differences in the values of neighboring source zones and the number and size of split source zones (i.e., source zones that overlap more than one target zone). An error model incorporating these factors was tested for four fractal surfaces. Strong correlations were observed between actual and predicted levels of error for the surfaces. The highest correlations occurred for surfaces of low dimensionality when the number of source zones was large relative to the number of target zones. [**Areal interpolation**]

137. Lam, N. S.-N. (1982). An evaluation of areal interpolation methods. *Proceedings, Auto Carto 5*, 471-479.
This study examines the factors affecting the accuracy of the direct overlay and pyenophylactic methods of areal interpolation. Accuracy is a function of the differences in the values of neigh-boring source zones and the number and size of split source zones (i.e., source zones that over-lap more than one target zone). An error model incorporating these factors was tested for four fractal surfaces. The correlations between actual and predicted levels of error for these surfaces was higher for the direct overlay method than the pyenophylactic method. [**Areal interpolation**]

138. LaMacchia, R. A. (1982). Geographic areas and computer files from the 1980 decennial census. *Proceedings, Auto Carto 5*, 491-498.
US Bureau of the Census DIME (Dual Independent Map Encoding) files contain certain clerical errors. At the tract and block levels, agreement between DIME files and 1980 census publication maps is about 98 percent. For any single DIME file, disagreement may be as high as 7 percent. Inconsistencies also exist in block numbers between DIME files and master area reference files. [**Planimetric accuracy**]

139. Laskowski, P. H. (1987). Map distortions and singular value decomposition. *Technical Papers, American Society of Photogrammetry and Remote Sensing-American Congress on Surveying and Mapping* Annual Convention, 4, 42-51.
This paper establishes a connection between Tissot's Indicatrix, used to assess projection-related map distortions, and the singular value decomposition of a linear transformation between a datum surface and a projection surface. The axes of Tissot's Indicatrix are shown to be the singular values of the appropriately scaled Jacobean matrix of the projection equations. This approach requires linearization of the projection equations, but provides a simple one-step algorithm for obtaining the parameters of the Indicatrix. [**Map projections**]

140. Laskowski, P. H. (1988). The traditional and modern took at Tissot's Indicatrix. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
An alternate conception of Tissot's theory of distortions is presented and a new method is pro-posed for computing the parameters of Tissot's Indicatrix, which is used to analyse distortions associated with map projections. The new method is based on the singular value decomposition of the appropriately scaled Jacobian matrix of the mapping equations. The advantages of this method are outlined and the traditional conception of Tissot's Indicatrix is presented for comparative purposes. [**Map projections**]

141. Lee, T., Richards, J. A. & Swain, P. H. (1987). Probabilistic and evidential approaches for multi-source data analysis. *IEEE Transactions on Geo8cience and Remote Sensing*, GE-25(3), 283-293.
Two methods are described for integrating raster data derived from multiple sources. The first is a probabilistic framework employing a global membership function and is computed from all data sources. The second is an evidential calculus based on Dempster's orthogonal sum combination rule. Information about uncertainty can be incorporated into either method. For the first method, the resulting reliability measure is identical for all cells, while for the second method it varies from cell to cell. [**Modeling uncertainty**]

142. Lee, Y. C. (1985). Comparison of planimetric and height accuracy of digital maps. *Surveying and Mapping*, 45(4), 333-340.
This paper proposes several methods for measuring the horizontal and vertical accuracy of con-tour maps. Horizontal accuracy may be measured with error polygons formed by superimposing the cartographic features to be tested with a more accurate standard. Alternatively, the end points of the standard and test features may be assumed to coincide, and the deviations measured between points generated along the lengths of these features. Empirical tests demonstrate that error is under-estimated by the former method and overestimated by the latter. Methods of measuring vertical accuracy based on local interpolation routines are also described. [**Topographic maps**]

143. Little, A. R. (1989). An evaluation of selected computer-assisted line simplification algorithms in the context of map accuracy standards. *Technical Papers, American Society of Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping* Annual Convention, 5, 122-132.
This study examines the performance of line simplification algorithms in the context of two map accuracy standards, the National Map Accuracy Standards (NMAS) and the American Society of Photogrammetry and Remote Sensing (ASPRS) spatial accuracy specifications for large scale maps. Experiments were performed on five line simplification algorithms (independent point, local, global, and constrained and unconstrained extended local routines). Performance was measured in terms of the percentage of points exceeding the requirements of the accuracy standard. Significant differences in performance were observed. The local routine produced the poorest results while the global routine produced the best. The ASPRS standards, being more rigorous, were more difficult to satisfy than the NMAS. In general, performance was not significantly affected by line sinuosity. [**Generalization error**]

144. Lloyd, P. R. (1976). Quantisation error in area measurement. *The Cartographic Journal*, 13(l), 22-25.
An alternate criterion is presented for calculating several of the parameters used in the error model developed by Frolov & Maling (1969). This criterion yields larger parameter estimates but does not significantly alter the results obtained by the model. [**Area estimation; Vector to raster conversion**]

145. Lodwick, W. (1988). Confidence limits in geographic analysis-suitability analysis. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
This paper suggests that sensitivity analysis might be used to examine the effects of variations in input data on output product reliability. One area of application is suitability analysis, in which a final suitability map is produced by superimposing a set of weighted data layers. Using sensitivity analysis, the effects of different weights or the exclusion of individual data layers can be computed. [**Geographic modeling**]

146. Loelkes, G. L., Jr. (1977). *Specifications for land use and land cover and associated maps* (Open File No. 77-555). Reston, VA: US Geological Survey.
This report details the specifications for maps prepared by the Land Information and Analysis Office and Topographic Division of the US Geological Survey. Specifications are outlined for source materials, compilation procedures, shoreline delineation methods, methods of classification and mapping formats. Quality control specifications are also outlined, including guidelines for field checking and map editing. [**Classification accuracy; Planimetric accuracy**]

147. MacDougall, E. B. (1975). The accuracy of map overlays. *Landscape Planning*, 2(l), 23-30.
The accuracy of a composite map created by map overlay is a function of polygon purity, horizontal error and error introduced by the map overlay operation itself. The lower limit of composite map accuracy may be computed from the operational error, the sum of the horizontal errors on each data layer and the product of the polygon purities. The upper limit of map accuracy may be computed from the operational error, the average horizontal error and the minimum polygon purity. Equations are presented for calculating horizontal error for individual data layers based on the assumed accuracy of different types of cartographic lines. [**Map overlay**]

148. MacEachren, A. M. (1982). Choropleth map accuracy: Characteristics of the data. *Proceedings, Auto Carto 5*, 499-507.
The accuracy of choropleth maps is dependent on the degree of enumeration unit homogeneity and the method of classification. Homogeneity is a function of enumeration unit size and compactness, as well as the complexity of the underlying surface. Empirical tests were performed on a sample of US counties in which maps were constructed for each county using four underlying surfaces of varying complexity. Map accuracy was expressed as the coefficient of variation for values measured at grid intersection points within each county. Accuracy was observed to be directly related to county compactness and inversely related to county size and surface complexity. [**Choroplethic mapping**]

149. MacEachren, A. M. (1982). Thematic map accuracy: The influence of enumeration unit size and compactness. *Technical Papers, American Congress on Surveying and Mapping*, 512-521.
This study examines the effects of enumeration unit size and compactness on the homogeneity of enumeration units on choropleth maps. Empirical tests were performed on a sample of US counties in which maps were constructed for each county using four underlying surfaces of varying complexity. Map accuracy was expressed as the coefficient of variation for values measured at grid intersection points within each county. Relatively strong correlations were observed between accuracy and both county size and compactness. Correlations declined as the complexity of the underlying surface increased. [**Choroplethic mapping**]

150. MacEachren, A. M. (1985). Accuracy of thematic maps: Implications of choropleth symbolization. *Cartographica*, 22(l), 38-58.
Thematic map accuracy is a function of data representation accuracy as well as positional accuracy. Data representation accuracy depends on enumeration unit size and compactness and the complexity of the underlying surface. Experimental results indicate that accuracy is higher when the underlying surface is simple and enumeration units are small and compact. Multiple regression results indicate that surface complexity and enumeration unit size and compactness account for a large percentage of the total variation in enumeration unit error. [**Choroplethic mapping**]

151. MacEachren, A. M. & Davidson, J. V. (1987). Sampling and isometric mapping of continuous geographic surfaces. *The American Cartographer*, 14(4), 299-320.
In isometric mapping of continuous geographic variables, values on the surface are often estimated by interpolating from a set of control points with known values. The accuracy of these intermediate value estimates is affected by data measurement accuracy, surface complexity, the interpolation method and the density and spatial distribution of control points. This study examines the relationship between accuracy and control point density for six surfaces of varying complexity. Results indicate that as density increases, both the error in the intermediate estimates and the randomness of the spatial distribution of this error decreases at a decreasing rate. The absolute value of error is higher for more complex surfaces. [**Isometric map-ping**]

152. McKay, C. J. (1967). The pattern map: A reliability standard for area sampling methods. *The Cartographic Journal*, 4(l), 114.
The pattern map is composed of geometric shapes designed for simple and accurate area estimation. The pattern map can be used to compare area estimates obtained with different planimetric techniques. [**Area estimation**]

153. Maffini, G. (1988). Observations and comments on the generation and treatment of error in digital GIS data. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis,* Montecito, California.
This paper examines the errors in digital cartographic data resulting from the digitization process. Digitizing error may be attributed largely to measurement error, the inherent indeterminacy of many boundaries and the inadequacy of models used to portray point locations. Trials in which geometric figures were digitized repeatedly show that error varies with map scale and the speed of digitizing. The dispersion of digitizing error around cartographic features suggests that polygon boundaries may be mapped as a probability surface rather than a single line. Other output products, such as composite maps derived from map overlay, might also be represented as a probability surface. [**Digitizing error**]

154. Maling, D. H. (1968). How long is a piece of string? *The Cartographic Journal*, 5(2), 147-156.
This paper discusses the errors associated with the estimation of line length from maps. The errors that might result from using different measuring instruments are reviewed and techniques are presented for correcting these errors. Other factors that are discussed include the dimensional instability of paper maps, projection-related distortions and the effects of cartographic generalization. [**Generalization error; Length estimation**]

155. Mandelbrot, B. (1967). How long is the coast of Britain? *Science*, 156, 636-638.
This paper presents the theoretical basis for the observation that estimates of line length derived from maps vary as a function of measurement precision. For self-similar lines, an analytical formula exists which relates line length to measurement precision. Self-similar lines are characterized by an exponent referring to their fractional dimensionality. Lines on maps often display statistical self-similarity. [**Length estimation**]

156. Markoff, J. & Shapiro, G. (1973). The linkage of data describing overlapping geographical units. *Historical Methods Newsletter*, 7(l), 34-46.
This paper outlines some techniques for performing areal interpolation and identifies some of the assumptions, problems and errors associated with these techniques. In the common area technique, target zone values for a variable of interest are estimated from source zone values weighted by the overlapping area of the source and target zones. This technique assumes that the variable is distributed uniformly, which is unrealistic for variables such as population. The common population technique accounts for non-uniformity by using the actual locations of cities as a basis for apportioning population to target zones. Empirical results indicate that the common population technique yields more accurate target zone population estimates than the common area technique. Estimation is also more accurate when target zones are larger than source zones. An approach analogous to the dasymetric mapping method is suggested as a possible improvement to these techniques. Factors limiting the spatial distribution of the variable of interest could be used to improve the accuracy of target zone estimates. [**Areal interpolation**]

157. Maxim, L. D. & Harrington, L. (1983). The application of pseudo-Bayesian estimators to remote sensing data: Ideas and examples. *Photogrammetric Engineering and Remote Sensing*, 49(5), 649-658.
In assessing thematic map accuracy with classification error matrices, a problem exists when a large number of matrix cells contain zeros. Some of these zeros are attributable to sampling variability and can complicate statistical analysis. This paper explores the use of pseudo-Bayesian estimation as a means of circumventing this problem. Prior unconditional probabilities for all matrix cells are calculated using a maximum entropy model that removes sampling zeros. [**Classification accuracy**]

158. Maxim, L. D., Harrington, L. & Kennedy, M. (1981). Alternative "scale-up" estimators for areal surveys where both detection and classification errors exist. *Photogrammetric Engineering and Remote Sensing*, 48), 1227-1239.
This paper details four statistical estimators to improve cartometric estimates derived from aerial surveys where both detection and classification errors are present. Numerical examples and tables are provided to illustrate the properties of these estimators. The choice of an estimator depends on the availability of data from ground survey. If data exist to support more than one estimator, the estimator yielding the lowest variance may be selected. [**Area estimation; Classification accuracy**]

159. Mead, R. A. & Szajgin, J. (1982). Landsat classification accuracy assessment procedures. *Photogrammetric Engineering and Remote Sensing*, 48(l), 139-141.
This paper outlines a set of proposed standards for thematic map classification accuracy. Accuracy testing procedures and reporting formats should be standardized. Accuracy assessments should account for the spatial distribution of error. Samples should be distributed over space and each cover type should be adequately sampled. Factors affecting sampling reliability should be considered, including sample size, the number of cover types, the confidence level desired and the relative significance of different misclassifications. The usefulness of data bases is often application-specific and not necessarily tied to quantitative accuracy. The applicability of unconventional sampling and statistical methods to thematic map accuracy assessment should also be examined. [**Classification accuracy**]

160. Merchant, D. C. (1982). Spatial accuracy standards for large scale line maps. *Technical Papers of the American Congress on Surveying and Mapping*, 222-231.
This paper outlines the American Society of Photogrammetry proposed accuracy standards for large scale line maps. The standards pertain to topographic and planimetric line maps at scales of 1:20 000 or more. The paper summarizes the accuracy compliance tests that are to be applied in assessing horizontal and vertical error. The format of the map accuracy report is also detailed. [**Planimetric accuracy; Topographic maps**]

161. Merchant, D. 0. (1987). Spatial accuracy specification for large scale topographic maps. *Photogrammetric Engineering and Remote Sensing*, 53(7), 958-61.
[See also: Merchant, D. 0. (1987). Spatial accuracy specification for large scale topographic maps. Technical Papers, American Society of Photogrammetry and Remote Sensing-American Congress on Surveying and Mapping Annual Convention, 2, 200-207.1
This paper outlines an accuracy standard for large scale topographic maps proposed by the Standards and Specifications Committee of the American Society of Photogrammetry and Remote Sensing. It is intended as an alternative to the US National Map Accuracy Standards (NMAS). The proposed standard differs from NMAS in that spatial accuracy is stated at ground scale, statements of precision are given in terms of root mean squared error and procedures for testing accuracy are fully specified. The proposed standard incorporates aspects of horizontal and vertical accuracy, designation of map accuracy classes, procedures for testing horizontal and vertical accuracy, and map accuracy statements. [**Topographic maps**)

162. Moellering, H. (1982). The goals of the National Committee for Digital Cartographic Data Standards. *Proceedings, Auto Carto 5*, 547-554.
The National Committee for Digital Cartographic Data Standards includes a working group on data base quality. Important data quality issues include metric and topological fidelity, reliability coding, temporal effects, data lineage and methods of accuracy assessment. Data producers should provide full information about data quality in order that users can reliably assess the fitness of the data for a given application. [**GBF standards**]

163. Monmonier, M. S. (1975). *Maps, distortion, and meaning* (Resource Paper No. 75-4). Washington, DC: Association of American Geographers.
Distortions in maps are inevitable since maps are abstractions of reality. Much of this distortion is intentional as it is designed to facilitate cartographic communication. This study explores the interface between the map producer and map user, and seeks to identify the ways in which cartographic communication may be enhanced through the use of appropriate map projections, generalization techniques and symbology. Map projections must be selected care-fully to minimize distortion in the message that is to be conveyed. Generalization is inevitable in map construction. For choropleth maps, a trade-off exists between generalization and accuracy as a function of the selection of class intervals. Map symbology must be designed with a clearly

defined and feasible objective in mind. Map producers interested in effective cartographic communication need to be aware of the message they wish to convey, the map user's limitations and the basic elements of maps. [**Cartographic communication**]

164. Monmonier, M. S. (1982). Flat laxity, optimization, and rounding in the selection of class intervals. *Cartographica*, 19(l), 16-27.
Class breaks on choropleth maps may be selected with optimization methods to reduce the discrepancy between the actual and mapped statistical surfaces. However, such class breaks often contain more significant digits than is warranted by the needs of the user or the precision of the data. Use of round number breaks may solve this problem without a significant loss of optimality. This principle, known as flat laxity, is the basis for a method of selecting class breaks that incorporates a rounding constraint in the optimization routine. Inherently meaningful breaks, such as the regional mean of the mapped variable, can also be incorporated as constraints. [**Choroplethic mapping**]

165. Morehouse, S. (1988). Experimental model of GIS. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Important data quality issues for geographic information systems include data availability and utility. Data utility refers to the level of accuracy required. Error tracking models must account for the entire sequence of operations applied to spatial data in the generation of out-put. Error tracking capabilities may be enhanced through documentation of data models and the use of standard methods of sampling and analysis. [**Data quality issues**]

166. Morrison, J. L. (1969). Control point spacing as an indicator of the accuracy of isarithmic maps. Papers from the *29th Annual Meeting of the American Congress on Surveying and Mapping*, 187-194.
The accuracy of isarithmic maps is affected by measurement precision, the interpolation method used, the spatial resolution of the output medium and the number and spatial distribution of control points. This study focuses on the effects of variations in the spatial distribution of control points. Empirical results indicate that clustered distributions of control points tend to produce less reliable isarithmic maps than dispersed distributions. The accuracy of an isarithmic map may be inferred by computing the nearest neighbor statistic for the set of control points used to construct the map. [**Isometric mapping**]

167. Morrison, J. L. (1971). *Method-produced error in isarithmic mapping* (Technical Monograph No. CA-5). Washington, DC: American Congress on Surveying and Mapping.
This study examines method-produced error associated with the isarithmic mapping technique. Method-produced error is defined as the discrepancy between the actual and mapped data values introduced by the techniques employed in map construction. Method-produced error is a function of the spatial distribution and density of control points and the interpolation method employed. Empirical results indicate that error is minimized when the distribution of control points is more uniform than a random distribution but less uniform than a regular lattice. Error also rises as surface complexity increases. Control point density affects the level of error, but a threshold exists beyond which the addition of control points does not improve accuracy. [**Isometric mapping**]

168. Muller, J.-C. (1977). Map gridding and cartographic errors: A recurrent argument. **Canadian Cartographer**, 14(2), 152-167.
This study is concerned with the effects of vector to raster conversion on map accuracy. Previous research on this topic is reviewed and some new results are presented. Empirical results demonstrate that the misclassified area on raster maps is linearly related to cell size. An error model proposed by Switzer (1975) is shown to yield accurate estimates of misclassified area. [**Vector to raster conversion**]

169. Muller, J.-C. (1987). The concept of error in cartography. *Cartographica*, 24(2), 1-15.
This paper examines the evolution of the concept of map error in the field of cartography. A distinction is drawn between controlled and uncontrolled error, since certain types of error, such as that associated with cartographic generalization, are often necessary for effective communication. For detailed, multi-purpose reference maps, accuracy may be determined in terms of discrepancies between the map and the real world. For generalized, single-purpose communication maps, error may be intentional since the map is designed to communicate an idea rather than depict the locations of features precisely. The advent of digital maps has changed the conception of map error. These maps have also introduced new types of error, including those associated with the encoding, manipulation and merging of spatial data. [**Data quality issues**]

170. Napton, D. E. & Luther, J. (1981). Transferring resource interpretations: Limitations and safe-guards. Proceedings, *Pecora V11 Symposium*, 175-186.
Surrogate variables are often used in modeling when the variable of interest cannot be measured. This paper shows that the reliability of surrogate variables is non-uniform over space. Safeguards against the inappropriate application of surrogate variables include recognition of the limitations of spatial data bases, awareness of data lineage and documentation of the criteria used in classification. [**Geographic modeling**]

171. Neumyvakin, Yu. K. & Panfilovich, A. 1. (1982). Specific features of using large-scale mapping data in planning construction and land farming. *Proceedings, Auto Carto 5*, 733-738.
The utility of spatial data for decision making depends on the accuracy of functions of coordinate values, including slopes, elevations and areas. The accuracy of such functions may be computed from the partial derivatives of the function, the standard error of measurement and the correlations between function parameters. The sensitivity of results to errors in coordinate values may be assessed with simulation methods. A numerical example is presented in which the accuracy of polygon area estimates is assessed. [**Geographic modeling**]

172. Newcomer, J. A. & Szajgin, J. (1984). Accumulation of thematic map errors in digital overlay analysis. *The American Cartographer*, 11(l), 58-62.
A method is presented for evaluating the accuracy of composite maps produced by map over-lay. The method is applicable to raster data when Boolean operators are applied. It is based on the probability that cells are misclassified on one or more data layers. Examples are presented and the upper and lower limits of composite map accuracy are defined. [**Map overlay**]

173. Newton, R. (1973). A statistical prediction technique for deriving contour maps from geophysical data. *Mathematical Geology*, 5(2), 179-189.
A statistical technique is presented for automated contouring of geophysical data. The value associated with a given control point is assumed to have been drawn from some probability distribution. The probability distribution may be changed to reflect the variability or reliability of the control point values. Advantages of this technique include elimination of edge effects, visual indication of the relative accuracy of contour locations, downgrading of values that are in error and minimization of error propagation effects. [**Isometric mapping**]

174. Nichols, J. D. (1975). Characteristics of computerized soil maps. *Soil Science Society of America Proceedings*, 39(5), 927-932.
Interpretive soil maps are constructed by computer by mapping the dominant soil type within each grid cell. This study examines the level of agreement between interpretive and detailed soil maps by comparing the acreage devoted to different soil types. Empirical results show that agreement tends to decline as cell size and the level of cartographic detail increase. [**Soil properties**]

175. Openshaw, S. (1988). Learning to live with errors in spatial databases. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Coping with error in spatial data bases involves the development of methods that account for the effects of error on spatial operations. Formal models of error propagation are not always required, since it may not be necessary to attain high levels of accuracy and since many applications require only that errors are not unreasonably large. The approach put forward in this paper is based on stochastic simulation. This approach is general, rests on few assumptions and is computationally tractable. It offers an indication of the possible level of error in output pro-ducts given a plausible set of assumptions about errors in the input data. Examples are presented in which this approach is incorporated into spatial analytic models. [**Geographic modeling**]

176. Otawa, T. (1987). Accuracy for digitizing: Overlooked factor in GIS operations. *GIS '87*, 295-299.
This paper examines error in polygon boundaries introduced by manual digitizing. A set of polygons on a soil map were digitized by subjects with no prior digitizing experience. The area of most polygons was within 7 percent of the mean area and errors rose in magnitude as polygon size increased. [**Digitizing error**]

177. Pearson, R. (1988). Approaches to detecting and correcting areas of unreliability in geographic databases. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
In assessing positional accuracy, the criterion used should be minimum absolute error rather than minimum squared error. Using coefficients from the smallest principal component yields an average error comparable to regression and adds flexibility to the research design since new coefficients may be obtained without major re-computations. [**Planimetric accuracy**]

178. Peucker, T. K. (1976). A theory of the cartographic line. *International Yearbook of Cartography*, 16, 134-143.
This paper focuses on the concept of a bounding rectangle encompassing a digital cartographic line. This rectangle has numerous implications and applications in the field of computer cartography. For example, it may be applied in map overlay, where different representations of the same cartographic line may exist on successive data layers due to errors associated with digitizing or line generalization. If the error variance is known, an appropriately scaled rectangle can be constructed such that any pair of lines may be considered to be equivalent if their rectangles overlap. [**Digitizing error; Generalization error; Map overlay**]

179. Peucker, T. K. (1980). The impact of different mathematical approaches to contouring. *Cartographica*, 17(2), 73-95.
Accuracy plays a role in the selection and implementation of a contouring algorithm. If the horizontal accuracy of the control points is low, few points should be selected. If the vertical accuracy is low, the approximation should be global and low-order. Simple algorithms are often appropriate when the data are inaccurate and the nature of the underlying surface is unknown. [**Isometric mapping**]

180. Prisley, S. P. & Smith, J. L. (1987). Using classification error matrices to improve the accuracy of weighted land-cover models. *Photogrammetric Engineering and Remote Sensing*, 5-9(9), 1259-1263.
Remotely sensed data often serve as inputs to geographic models. Suitability analysis, for example, relies on estimates of the area of uniform- cover types. Inaccuracies in area estimates can reduce model reliability. Reliability can be improved by adjusting area estimates to account for classification error. This paper illustrates this procedure using information from a classification error matrix. [**Area estimation; Classification accuracy**]

181. Rattray, M., Jr. (1962). Interpolation errors and oceanographic sampling. *Deep-Sea Research*, 9, 25-37.
Formulae are derived which give interpolated values for an oceanographic variable and an estimate of the accuracy of interpolation. The error in interpolated values is composed of measurement error and error arising from interpolation. Interpolation error is calculated as the difference between two Lagrange interpolated polynomials. Measurement error is expressed as the ratio of the standard deviation of the interpolated values to the standard deviation of the individual measurements. The derived formulae may also be used to devise sampling schemes that yield an interpolation result of desired accuracy. [**isometric mapping**]

182. Rhind, D. W. (1971). Automated contouring: An empirical evaluation of some differing techniques. *The Cartographic Journal*, 8(2), 145-158.
Interpolation accuracy may be affected by the accuracy of control point values, the spatial distribution of control points, the sampling methodology, the complexity of the surface and the nature of the contouring algorithm. Accuracy is typically defined in terms of the deviations between actual and estimated values for a set of points. However, a high degree of accuracy does not necessarily imply that the interpolated surface will have a reasonable visual appearance. Moreover, quantification of accuracy depends on the availability of adequate reference data. [**Isometric mapping**]

183. Rhind, D. W. (1975). A skeletal overview of spatial interpolation techniques. *Computer Applications*, 2(3/4), 293-309.
This paper reviews methods of spatial interpolation and briefly discusses some potential sources of error. Interpolation accuracy may be estimated with reference to more accurate data, by incorporating reliability information prior to interpolation, or by using methods such as least squares or kriging which produce error estimates. However, the validity of least squares or kriging rests on assumptions that are often not met in an operational environment. Moreover, statistical tests of accuracy typically focus on the accuracy of individual points rather than the whole surface. A high degree of interpolation accuracy does not necessarily imply that the interpolated surface will have a reasonable visual appearance. [**Isometric mapping**]

184. Robinson, J. E. (1973). Frequency analysis, sampling, and errors in spatial data. *In* J. C. Davis & M. J. McCullagh (Eds.), *Display and Analysis of Spatial Data* (pp. 78-95). London: Wiley.
Error in contour maps may be attributed to positional error in control points and measurement error in the values associated with these points. The effects of both sources of error can be estimated statistically and displayed as an amplitude spectrum using Fourier analysis of maps or map cross-sections. This approach facilitates assessment of the effects of random measurement error and aliasing error as a function of the distance between control points. Positional error may be assessed from the phase spectrum. This approach may also be used to evaluate the effects of trend analysis, whether by polynomial surface, Fourier series or spatial filtering. [**Topographic maps**]

185. Robinson, J. W., Gunther, F. J. & Campbell, W. J. (1983). *Ground truth sampling and Landsat accuracy assessment* (Document N83-26161). National Technical Information Services.
Classification error matrices may be used to assess classification accuracy for thematic maps. The matrices may be used to estimate the probability that a pixel is correctly classified, the probability that a pixel belonging to cover type i is assigned to type i and the probability that a pixel assigned to cover type i is actually a member of type i. The k statistic may be used to estimate the percentage improvement in classification accuracy over a purely random assignment of cover types. Cluster sampling may be inappropriate for acquiring a sample of pixels for accuracy assessment, since pixel values are often spatially autocorrelated. Stratified sampling procedures are an alternative but require a priori knowledge about the presence of different cover types. [**Classification accuracy**]

186. Robinson, V. B. (1984). Modeling inexactness in spatial information systems. *Modeling and Simulation*, 15, 157-161.
This paper reports on some preliminary efforts to develop a linguistic approximation of distance as a natural language concept. Since linguistic concepts can accommodate the vagueness inherent in human language, the acquisition of these concepts by a

spatial information system (SIS) may facilitate natural language queries and provide insights into the relationships between natural language and the results of an SIS query. As an example, the paper outlines a learning process whereby an SIS acquires an approximation of the concept of 'near' by a question and answer approach. In contrast to many previous natural language applications, this approach need not assume that the SIS user has a highly developed ability to translate inexact natural language concepts into exact concepts for the SIS. [**Modeling uncertainty**]

187. Robinson, V. B. & Strahler, A. H. (1984). Issues in designing geographic information systems under conditions of inexactness. *1984 Machine Processing of Remotely Sensed Data Symposium*, 198-204.
Fuzzy logic is introduced as a basis for representing and manipulating inexactness in a geo-graphic information systems. In modeling with fuzzy logic, the basic concept applied is a measure of relative accuracy (i.e., a membership value) that may be assigned to each record in a relational data base or flat file. Four specific cases are examined. Non-fuzzy schema with non-fuzzy data is equivalent to the standard relational data base or flat file. Fuzzy schema with non-fuzzy data represents a case in which the data model can cope with inexactness but the data can be captured exactly. Fuzzy schema with fuzzy data represents the ultimate case of inexactness. The ruling paradigm is non-fuzzy schema with fuzzy data, where the data model is expressed precisely but the data cannot be captured exactly. Some examples of modeling approaches for this last case are discussed. [**Modeling uncertainty**]

188. Roller, N. E. G. (1984). Effective integration of data sources for optimizing resource inventories. *Eighteenth International Symposium on Remote Sensing of Environment*, 1577-1603.
A general methodology is presented whereby ancillary data, in this case provided by remote sensing, can be used to improve the accuracy of polygon boundary delineation. The methodology is based on the identification of credible polygons within the polygons defined in the original survey. [**Map overlay**]

189. Rosenfield, G. H. (1971). On map accuracy specifications: Part IL Horizontal accuracy of topographic maps. *Surveying and Mapping*, 31(l), 60-64.
This paper presents a practical technique for assessing the horizontal accuracy of topographic maps. After transferring the geographic coordinates of the graticule onto a plane-rectangular coordinate system, an affine transformation is applied to fit the coordinates to the plane-rectangular system. The residuals represent the horizontal error of the graticule points and the circular standard error may thus be computed. The same approach is then applied to a sample of planimetric features. The total horizontal map error is determined from the circular standard errors of both the graticule and planimetric features. [**Topographic maps**]

190. Rosenfield, G. H. (1981). Analysis of variance of thematic mapping experiment data. *Photogrammetric Engineering and Remote Sensing*, 47(12), 1685-1692.
This study explores the applicability of analysis of variance for assessing thematic map classification accuracy. Probabilities from a classification error matrix may be transformed using the normality assumption, the arcsine transformation or the logit transformation. Test criteria include mean squared error, aresine mean squared error, harmonic mean squared error and multiple range tests. An empirical example using two-way analysis of variance demonstrates how inferences may be made about differences in accuracy among cover types and the effects on map scale on overall classification accuracy. [**Classification accuracy**]

191. Rosenfield, G. H. (1986). Analysis of thematic map classification error matrices. *Photogrammetric Engineering and Remote Sensing*, 52(5), 681-686.
The log-linear model may be used to test hypotheses about thematic map classification accuracy. This model ensures that the sums of the predicted probabilities for cover types fall within the allowable range of zero to one. It allows for comparison of different thematic classifications and the accuracy with which different cover types are assigned. [**Classification accuracy**]

192. Rosenfield, G. H. & Fitzpatrick-Lins, K. (1986). A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 52(2), 223-227.
Thematic map accuracy may be assessed with the k coefficient, while the conditional k coefficient may be used to assess the accuracy of individual cover types. These coefficients use information from all cells of the classification error matrix and account for errors of omission and commission. The k coefficient is compared to other coefficients that have been proposed in the literature. [**Classification accuracy**]

193. Rosenfield, G. H., Fitzpatrick-Lins, K. & Ling, H. S. (1982). Sampling for thematic map accuracy testing. *Photogrammetric Engineering and Remote Sensing*, 48(l), 131-137.
A sampling method for assessing thematic map accuracy is described that uses stratified systematic unaligned sampling with an additional random sample for under-represented cover types. The minimum sample size for each type may be determined from the specified accuracy standard and the required level of confidence. Statistical tests are developed for assessing the accuracy of individual cover types and the overall accuracy of the map. [**Classification accuracy**]

194. Rosenfield, G. H. & Melley, M. L. (1980). Applications of statistics to thematic mapping. *Photogrammetric Engineering and Remote Sensing*, 46(10), 1287-1294.
For large samples, thematic map classification accuracy may be assessed with the normal approximation to the binomial distribution. Confidence limits may be computed for the accuracy estimate. Sample size should be selected as a function of the required level of confidence. The accuracy of different maps may be compared with t-tests, non-parametric tests and analysis of variance. [**Classification accuracy**]

195. Saalfield, A. (1988). Census Bureau research concerns for accuracy of spatial data. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Topological constructs must be consistently applied during updates of digital cartographic data bases. Rational or finite precision geometry may provide a means of producing consistent topology. In conducting large surveys, many operational concessions must be made that affect the reliability of the resulting sample. Current survey practices are reviewed and the limitations and new opportunities afforded by GIS technology are identified. [**Data quality issues**]

196. Shepard, D. S. (1984). Computer mapping: The SYMAP interpolation algorithm. *In* G. L. Gaile & C. J. Willmott (Eds.), *Spatial statistics and models* (pp. 133-145). Dordrecht: Reidel.
Interpolation accuracy depends on the nature of the interpolation algorithm, the complexity of the underlying surface and the distance between control points relative to the frequency of spa-tial variation. For the SYMAP interpolation algorithm, the relative root mean squared error between actual and interpolated values is shown to be closely related to the number of control points. The functional form of this relationship is presented. [**Isometric mapping**]

197. Smith, G. R. & Honeycutt, D. M. (1987). Geographic data uncertainty, decision making and the value of information. *GIS '87*, 300-312.
This paper demonstrates how the value of information concept can be used to estimate the value of reducing uncertainty in spatial data. The methodology employed is a decision tree, a graphic representation of a decision problem in which decisions have both alternatives and out-comes. Decision nodes in the tree represent points in the decision process where a choice must be made between alternatives. Chance nodes represent points where uncertainties are resolved. The value of information represents the value of reducing or eliminating uncertainties associated with chance nodes. The expected value of perfect information represents the increase in the value of a decision node associated with completely resolving uncertainty associated with a chance node. A geological example is presented to demonstrate how these parameters can be computed from the decision-tree. [**Modeling uncertainty**]

198. Smyth, S. (1988). Implicit positional accuracy in a large online spatial database: Data model translation project. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Utility companies often query information systems about the location of utility features under-ground. These systems need an accuracy statement to formalize the occurrence of error in accordance with liability concerns. The error tracking system must operate at the same speed as the query system. A system of nested tessellations may be an appropriate model for the error tracking system. [**Planimetric accuracy**]

199. Star, J. (1988). Sources of error in thematic classification of remotely sensed data. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Classification of remotely sensed data is often performed without examining the underlying frequency distribution of pixel values. It is simply assumed that the data are normally distributed. It is not clear how deviations from normality affect the reliability of the classification process. Simulation studies are needed that examine the effects of random noise in input data, training field selection criteria and different classification distance metrics. [**Classification accuracy**]

200. Stearns, F. (1968). A method for estimating the quantitative reliability of isoline maps. *Annals of the Association of American Geographers*, 58(3), 590-600.
A method is presented for estimating the theoretical accuracy of an isoline map based on known or estimated parameter values and characteristics of the mapped variable. Reliability equations are developed for measurement error (including observational error, time error and positional error), synopticity error and interpolation error (including truncation error and propagation of

error). Errors are combined mathematically by transforming them into errors in the mapped variable. Overall map reliability is expressed in terms of total variance. It is possible to design surveys to meet a specific level of variance. [**Isometric mapping**]

201. Stoms, D. (1987). Reasoning with uncertainty in intelligent geographic information systems. *GIS '87*, 693-700.
This paper focuses on methods for coping with uncertainty in spatial data associated with random error, vagueness and incompleteness of evidence. These methods include Bayesian probability, Schafer's theory of evidence, fuzzy set theory and non-monotonic logic. Bayesian probability is shown to be best for random error, fuzzy set theory for vagueness, and Schafer's theory of evidence and non-monotonic logic for incompleteness. [**Modeling uncertainty**]

202. Story, M. & Congalton, R. G. (1986). Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing*, 52(3), 397-399.
This paper argues that thematic map classification accuracy should not be summarized with a single index, such as the percent of sample points correctly classified. It is also important to consider errors of omission and commission for individual cover types, which are reflected in the off-diagonal elements of the classification error matrix. [**Classification accuracy**]

203. Switzer, P. (1975). Estimation of the accuracy of qualitative maps. *In* J. C. Davis & M. J. McCullagh (Eds.), *Display and analysis of spatial data* (pp. 1-13). London: Wiley.
This paper develops a model for estimating the effects of vector to raster conversion on the reliability of qualitative maps. Reliability is defined in terms of the area of disagreement in cover type assignments on the actual and raster maps. The model is based on the probability that two points separated by a certain distance are correctly assigned to different cover types. This probability varies as a function of map complexity. The model is based on the geometry of the raster elements and does not require knowledge of the actual map. [**Vector to raster conversion**]

204. Switzer, P. (1988). Precision of interpolated maps. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis,* Montecito, California.
Issues of importance in assessing the accuracy and precision of interpolated maps include sample design, non-stationarity, optimality and sensitivity of interpolated values, and spatial-temporal and qualitative data. Interpolation accuracy may be assessed using model-based sampling, Bayesian analysis, kriging, cross-validation and re-sampling. Covariance modeling may be applied to sampling design to determine the optimal locations of sample points. [**Isometric mapping**]

205. Theobald, D. M. (1988). Accuracy and bias issues in surface representation. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Two factors must be considered in assessing the accuracy of digital elevation models (DEMs). The first is the ability of the source document to capture surface variations. The second is the reliability of information derived from the DEM, which is affected by aspects of sampling, interpolation and spatial resolution. DEM accuracy refers to the spatial distribution of error as well as the deviation between the actual and estimated surfaces. Accuracy- is not equivalent to acceptability, since accuracy requirements are application-specific. Spatial resolution requirements for DEMs should be viewed in terms of the interaction between terrain variability and sampling frequency. [**DEMs**]

206. Thomas, 1. L. & Allcock, G. McK. (1984). Determining the confidence level for a classification. *Photogrammetric Engineering and Remote Sensing*, 50(10), 1491-1496.
A method is described for computing the* confidence interval for thematic map accuracy statements. The method is based on the size of the sample and the number of correct and incorrect classifications observed in a sample. Allowance is made for the error associated with assigning a heterogeneous pixel to the dominant cover type. [**Classification accuracy**]

207. Thompson, M. M. (1973). Standards for modern maps. *Proceedings of the American Congress on Surveying and Mapping Fall Convention*, 140-147.
New technology and new ways of presenting maps demand revisions of existing map accuracy standards. This paper reviews certain aspects of such standards, including the geometric reference system, map symbology and aesthetics, and accuracy and quality standards. Suggestions are made as to how these aspects might be revised and updated. With regard to accuracy and quality standards, the author advocates the use of statistical terminology, the cataloging of standards for different kinds of maps and the formulation of quality control standards for map content, readability and aesthetics. [**Data quality issues**]

208. Thompson, M. M. (1971). On map accuracy specifications: Part 1. Practical interpretation of map-accuracy specifications. *Surveying and Mapping*, 31(l), 57-60.
Map accuracy standards are not so mathematically incontrovertible that they yield the only interpretation of accuracy. Statistical expressions of accuracy, for example, typically assume that horizontal error is normally distributed. Tests of vertical accuracy are often inapplicable in areas of non-uniform terrain. Qualitative aspects of map accuracy are also important but often ignored in map accuracy standards. [**Topographic maps**]

209. Tikunov, V. S. (1986). *Some issues in the modeling approach to cartography*. Unpublished manuscript.
The author distinguishes between the notions of technical map accuracy and the reliability of map content. The latter refers to the degree to which a geographic phenomenon corresponds to its cartographic model. The author demonstrates how the reliability of modeling may be enhanced by seeking independent, confirmatory evidence. [**Geographic modeling**]

210. Tobler, W. (1988). Frame independent spatial analysis. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
The results of any analysis of geographical data should be frame independent, in that they should not depend on the spatial coordinates employed. Previous work indicates that some methods of analysis do not yield the same results under alternate areal aggregations, but methods of analysis do exist that seem to yield frame independent results. In some cases, the effects of a spatial aggregation can be determined with linear spatial models or by considering the aggregation to be a spatial filter with an estimable response function. It is proposed that only those methods that exhibit frame independence be considered to be appropriate for geo-graphical analysis. [**Geographic modeling**]

211. Tryon, T. C., Hale, G. A. & Young, H. E. (1955). Dot gridding air photos and maps. *Photogrammetric Engineering*, 21(5), 737.
The accuracy of area estimates based on dot planimetry depends on the density of the dot grid. Empirical results show that for low-density grids, a decrease in the area of the region being measured results in a loss of accuracy. Satisfactory results can be obtained with low-density grids only when large regions are involved. [**Area estimation**]

212. Turk, G. (1979). The GT index. *Remote Sensing of Environment*, 8(l), 65-75.
The percentage of sample pixels correctly classified can be a misleading index of thematic map accuracy since it equates apparent accuracy to the efficacy of the classification procedure. The GT index is proposed as an alternate accuracy index. This index measures the deviation between the actual assignment of cover types and a purely random assignment by accounting for correct assignments that occur due to chance alone. [**Classification accuracy**]

213. Veregin, H. (1987). *Error modeling in geographic information systems: A review*. Unpublished manuscript.
This paper reviews previous research on error modeling for geographic information systems with particular emphasis on models for vector to raster conversion and map overlay. A conceptual framework is posited for identifying the sources and types of errors in these operations. The framework distinguishes between categorical and numerical input data, error propagation and error production, and operational, cartographic and thematic errors in output products. Shortcomings of current error models for vector to raster conversion and map overlay are elucidated and some themes pertinent to error modeling for other operations are identified. [**Map overlay; Vector to raster conversion**]

214. Veregin, H. (1988). Error modeling in geographic information systems. Paper presented at the *Canadian Association of Geographers Annual Meeting*, Halifax, Nova Scotia.
This paper reviews previous research in error modeling for map overlay with particular emphasis on variables that affect the types of errors present in output products. A conceptual model for categorizing these variables is presented. Errors associated with map overlay are shown to depend on the nature of the input data, the spatial data structure employed, the types of errors present in the input data and the mode of error accumulation. Interactions among these variables often make it difficult to isolate individual effects in output products. The relative importance of individual variables determines the strategies that might be adopted for measuring, managing and reducing error. [**Map overlay**]

215. Veregin, H. (1988). Error modeling for the map overlay operation. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Informa-tion and Analysis*, Montecito, California.
This paper critically examines some current approaches to error modeling for the map overlay operation. In applying error models for map overlay it is important to consider the sources of error present in the input data and the methods used to detect and measure these errors. The ability to successfully manage and reduce error in output products is in turn dependent on the nature of the error model applied. [**Map overlay**]

216. Veregin, H. (1989). A review of error models for vector to raster conversion. *The Operational Geographer*, 7(1), 11-15.
This paper posits a simple taxonomy of error models for the vector to raster conversion opera-tion based on the source of error and the characteristics of the input data. Error models for this operation may emphasize cartographic or thematic error sources and, in the latter case, focus on either numerical or categorical data. Existing error models for vector to raster conversion are reviewed with the aim of elucidating how error sources and input data characteristics determine the types of models that may appropriately be applied. [**Vector to raster conversion**]

217. Vitek, J. D. & Richards, D. G. (1978). Incorporating inherent map error into flood-hazard analysis. *Professional Geographer*, 30(2), 168-173.
This paper is concerned with the implications of inherent error in topographic maps for identifying flood-prone areas. Inherent error includes error associated with coordinate transformations, map construction practices and map symbology. Vertical and horizontal error may also be introduced during map compilation and may be magnified during map use. In assessing flood hazard, error reduces the reliability of the line identified as the flood limit. A gradational symbol should replace the line and the limit of flooding should appear as a zone. [**Geographic modeling**]

218. Vitek, J. D., Walsh, S. J. & Gregory, M. S. (1984). Accuracy in geographic information systems: An assessment of inherent and operational errors. *Proceedings, PECORA IX Symposium,* 296-302.
This paper evaluates the status of error assessment for spatial data bases used in geographic information systems. Examples of inherent error are presented to indicate the need to develop an error statement for output products. The amount of inherent error is a function of the map projection, construction techniques and map symbology. Examples of operational error are described to illustrate the types of errors that may be present in output products. Operational error may be introduced during data entry, manipulation, extraction and comparison. Producers of output should specify the types and levels of error present based on interactions between inherent and operational errors. [**Data quality issues**]

219. Vonderohe, A. (1988). An experimental field lab for testing accuracies of spatial data bases. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
This study addresses the problem of positional errors in cadastral maps and legal descriptions of land title. Experiments were conducted in which the positions of a set of permanently monumented, second-order, class I points were compared with their positions as obtained from exist-ing cadastral maps and legal descriptions. The cadastral maps showed a high degree of positional error, while legal descriptions exhibited gross errors such as gaps and overlaps. In constructing maps from legal descriptions, ancillary. data should be used and high accuracy levels should not be expected. [**Planimetric accuracy**]

220. Walsh, S. J. (1988). Inherent and operational error within a GIS: Evaluation from a user perspective. Paper presented at *First Specialist Meeting on Accuracy of Spatial Data Bases, National Center for Geographic Information and Analysis*, Montecito, California.
Inherent and operational errors contribute to a reduction in the accuracy of data contained in a geographic information system. Source documents contain inherent error as a function of the methods followed in their creation. Operational error results from data manipulation. Error assessment is the responsibility of both the producer and the user of data. Producers may document data base quality but users must be able to evaluate the appropriateness of the data for a given application. User sophistication in evaluating fitness of use will have a positive impact on the validity and reliability of output products. [**Data quality issues**]

221. Walsh, S. J., Lightfoot, D. R. & Butler, D. R. (1987). Recognition and assessment of error in geographic information systems. *Photogrammetric Engineering and Remote Sensing*, 53(10), 1423-1430.
[See also: Walsh, S. J., Lightfoot, D. R. & Butler, D. R. (1987). Assessment of inherent and operational errors in geographic information systems. *Technical Papers, American Society of Photogrammetry and Remote Sensing-American Congress on Surveying and Mapping* Annual Convention, 5, 24-35.1
Errors in composite maps produced by map overlay result from inherent error in data layers and operational error associated with the map overlay operation itself. An empirical test was performed in which map overlay was applied to three raster data layers. For each data layer, the amount of inherent error was assessed with a classification error matrix. Operational error resulted from the lack of coincidence in misclassifed cells on different data layers. Inherent and operational errors combined to render composite maps highly inaccurate, especially when cell size was large and more than two data layers were superimposed. [**Map overlay**]

222. Webster, R. (1968). Fundamental objections to the 7th approximation. *Journal of Soil Science*, 19(2), 354-366.
The 7th approximation, a US Department of Agriculture soil classification system, suffers from three fundamental problems. First, it demands an unattainable level of precision and its mutually exclusive classes lead to inconsistencies when applied. Second, it employs a hierarchical structure which does not account for the dispersed nature of soil distributions. Third, although genetically significant properties are the basis for classification, these can usually be identified only after soils have been classified. Soil mapping units should be defined to reflect local conditions and the presence of uncertainty in class membership. However, this will result in class overlap at the series level and nonconformity in hierarchical relationships. [**Soil properties**]

223. Webster, R. & Beckett, P. H. T. (1968). Quality and usefulness of soil maps. *Nature*, 219, 680-682.
This study examines the effects of spatial variations in soil properties on the accuracy and usefulness of soil maps. Soil map accuracy is a function of variation within soil mapping units. Improvements in accuracy can thus be measured by the intraclass correlation coefficient. Users of soil maps might also wish to know that the within class variance does not exceed a certain threshold. The mean, variance and other descriptive statistics for each mapping unit might also provide useful information to

users. It must be recognized that a difference exists between a soil map and a purely taxonomic classification of survey points, since the former must account for spatial relationships. Map scale must also be sufficiently large to facilitate portrayal of spatial variation at the required level of detail. [**Soil properties**]

224. Welide, M. (1979). Spatial quantification of maps or images: Cell size or pixel size implications. *Joint Proceedings of the American Society of Photogrammetry and American Congress on Surveying and Mapping Fall Meeting*, 45-51.
This study examines the effects of grid cell size on the error associated with vector to raster conversion. The effects of cell size are modeled in terms of the span distribution, or the frequency distribution of distances between cells located on polygon boundaries. These distances are integer multiples of cell size. The distribution is posited to characterize a given map and specifies the form of the relationship between error and cell size. [**Vector to raster conversion**]

225. Welide, M. (1982). Grid cell size in relation to errors in maps and inventories produced by computerized map processing. *Photogrammetric Engineering and Remote Sensing*, 48(8), 1289-1298.
This study develops a model for assessing the errors associated with vector to raster conversion. The frequency distribution of distances between cells located on polygon boundaries, or the span distribution, is used to characterize a given map. This distribution specifies the relation-ship between map error and cell size. In general, map accuracy declines as cell size increases. [**Vector to raster conversion**]

226. White, D. (1978). A design for polygon overlay. *In* G. Dutton (Ed.), *Harvard Papers on Geo-graphic Information Systems*, 6. Cambridge, MA: Laboratory for Computer Graphics and Spa-tial Analysis, Harvard University.
The WHIRLPOOL module of the ODYSSEY system performs map overlay using a measure of tolerance for defining line intersections. This can be used to handle different versions of the same cartographic feature on successive data layers, thereby reducing the incidence of spurious polygons on the composite map. [**Generalization error; Map overlay**]

227. White, M. (1983). Tribulations of automated cartography and how mathematics helps. *Cartographica*, 21(2/3), 148-159.
[See also: White, M. (1983). Tribulations of automated cartography and how mathematics helps. *Proceedings, Auto Carto 6*, 408-418.1
Many errors in automated cartography are attributable to the lack of appropriately applied mathematical concepts. Topological principles are shown to be relevant for encoding polygon boundaries, defining the accuracy of coordinate values and performing map overlay, vector to raster conversion and triangulation for digital elevation models. In most cases the solution to the problem of error lies in the identification of the appropriate ontological issue. Mathematical concepts relevant to automated cartography are presented in a syllabus which emphasizes topological, graph and model theory. [**Data quality issues**]

228. Willmott, C. J., Rowe, C. M. & Philpot, W. D. (1985). Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *The American Cartographer*, 12(l), 5-16.
Algorithms are developed for contouring on a spherical surface in order to investigate error in small-scale climate maps associated with the practice of interpolating in Cartesian two-space. Experimental results reveal significant errors in two-space interpolations, especially in areas containing few data points or having large projection-related distortions. Gross errors were also found to be more apparent for two-space interpolation. [**Isometric mapping**]

229. Yoeli, P. (1984). Error bands of topographical contours with computer and plotter (Pro-ram KOPPE). *Geo-Processing*, 2(3), 287-297.
This paper describes an algorithm for generating error bands on topographical contour maps based on Koppe's formula. This formula emphasizes geometrical errors, or discrepancies between actual and mapped elevation values, rather than form errors, or inadequacies in the overall representation of the surface. Koppe's formula stipulates that the mean squared height error of any point is directly proportional to the tangent of the angle of slope of the terrain at that point. On maps, this error is manifest as horizontal displacement, and the mean squared planimetric error of a contour may thus be derived and plotted. [**Topographic maps**]

KEYWORD DESCRIPTIONS

Measurement of error in spatial databases:
- **Classification accuracy:** Classification accuracy for thematic maps.
- **Planimetric accuracy:** Planimetric accuracy for cadastral and planimetric line maps.
- **GBF standards:** Data quality standards for geographic base files (GBFs).
- **Soil properties:** Spatial variations in soil properties.
- **Topographic maps:** Horizontal and vertical accuracy for topographic maps.
- **DEMs:** Errors in digital elevation models (DEMs).

Accuracy of cartometric estimates:
- **Length estimation:** Accuracy of estimates of line length.
- **Area estimation:** Accuracy of estimates of area.

Errors associated with data compilation:
- **Digitizing error:** Errors associated with map digitizing procedures.
- **Generalization error:** Errors associated with cartographic line generalization.
- **Map projections:** Distortions associated with map projections.
- **Choroplethic mapping:** Errors associated with choropleth mapping procedures.
- **Isometric mapping:** Errors associated with isometric mapping procedures.
- **Isoplethic mapping:** Errors associated with isoplethic mapping procedures.
- **Cartographic communication:** Errors associated with the cartographic communication process.

Error models for specific operations:
- **Areal interpolation:** Error models for areal interpolation.
- **Vector to raster conversion:** Error models for vector to raster conversion.
- **Map overlay:** Error models for map overlay.

Miscellaneous issues:
- **Data quality issues:** General issues pertaining to spatial data quality.
- **Geographic modeling:** Error tracking in geographic modeling.
- **Modeling uncertainty:** Models of uncertainty for spatial data.