6

Accuracy

Kenneth McGwire and Michael Goodchild

6.1. Introduction

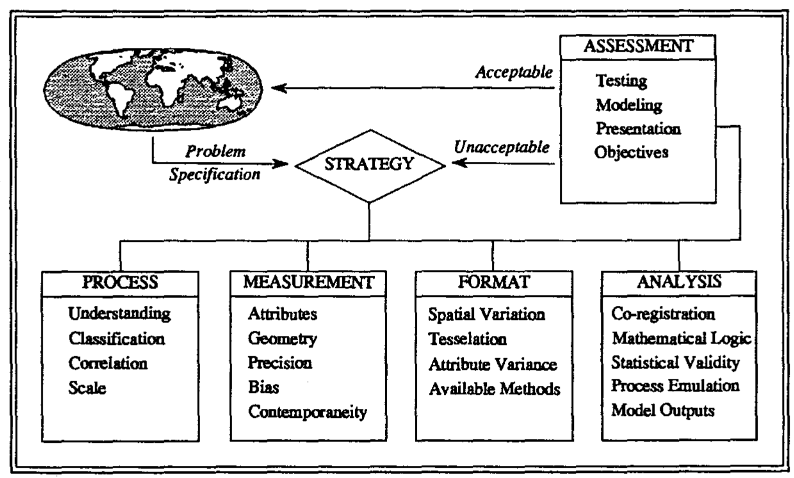
As the fields of remote sensing and GIS mature, these two disciplines are increasingly being looked to for ways of quantifying complex social and environmental events. Decision making regarding such events generally involves uncertainty in understanding, quantification, and prediction. In order to assess alternatives in a reasonable manner, some indication of confidence in various information sources must be available. This chapter identifies aspects of spatial data processing that affect the accuracy of information products derived from remote sensing and GIS data analysis. The discussion is set in the context of a spatial data processing strategy for providing information products of high quality and known accuracy characteristics to decision makers.

Though the integration of remote sensing and GIS has been promoted for some time (Shelton and Estes, 1981, Marble and Peuquet, 1983), and to greater or lesser extents implemented, only recently has there been some recognition of the need to draw from the background of both communities to synthesize an integrated view of accuracy for spatial information (e.g., Chrisman, 1989). Error accumulation in remote sensing and GIS data processing is difficult to track, both in terms of the availability of data for validation and the conceptual understanding of error sources, their propagation, and individual or cumulative effects. In remote sensing, the acquisition of consistent spectral response from the Earth's surface is made difficult by the complexities of sun/target/sensor geometry, bidirectional reflectance distributions, nonuniform atmospheric characteristics, and spectral variability (Duggin, 1985; Asrar, 1989). In GIS, positional and thematic accuracies are a

compromise of scale, contemporaneity, media stability, and compilation standards (Mahling, 1989; Goodchild and Gopal, 1989). As will be shown, many of the uncertainties encountered within these two complementary disciplines are analogous. The challenges of representing social and environmental phenomena accurately with remote sensing and GIS will be considered as falling within a more general heading of spatial data processing. The following chapter sections organize accuracy issues within spatial data processing into five interrelated areas referred to as process, measurement, format, analysis, and assessment. Figure 6.1 presents the framework for this discussion.

Figure 6.1 poses a question regarding a spatially distributed process and then develops a strategy to determine functional relationships between observations. This strategy defines the methods that will be used to transform available data sources into information products for decision making. The ability to provide accurate information regarding the spatial process will be fundamentally dependent on the level of understanding of the system under study. Issues related to the original delineation and description of a system with regards to a particular query will be discussed as PROCESS. Data collected to quantify components of the system under study are subject to errors or uncertainty in MEASUREMENT. Data FORMAT, which is required to establish a consistent and manageable data structure, may also limit accurate representation of natural or social measurements. ANALYSIS of data may use

Figure 6.1. Accuracy issues in spatial data processing



some accepted relationship between process components or may develop new relationships between data sources. The techniques used to synthesize new information products may strongly affect accuracy. Finally, prudent experimental design also requires that direct accuracy ASSESSMENT be performed prior to acceptance of results. If the observed system is not adequately represented by an information product, then the initial strategy must be examined for flaws, and a new strategy may be formulated.

The strategy selected to address a particular question defines the types of data used, the methods of analysis, and the ability to assess results objectively. This strategy will be affected by numerous indirect considerations, including time, cost, and expected short-term versus long-term benefits. Specific strategies will vary by application and circumstance; however, all spatial data processing scenarios will require consideration of some, if not all, of the generic concerns outlined in this chapter.

6.2. Process

The data and operations of a GIS represent a workable model of a physical and/or cultural system. To answer a query regarding a particular system, the relevant observations with which to represent that system must first be identified. In order to obtain relevant data to answer a query, some level of understanding is required regarding the process under study. This knowledge allows adequate delineation of relevant boundaries of a system and determination of the most parsimonious solution to a query. Examples of how inadequate knowledge of process may reduce accuracy include the use of inappropriate generalizations, mistaking causation from among correlated variables, or focusing on an inappropriate scale of observation.

Classification of phenomena within a discrete taxonomic scheme may facilitate identification and communication regarding complex phenomena. Inherent in any classification scheme is the concept that within-class variability is less than between-class variability. In a spatial context, this principle applies not only to the definition of the classes, but also to the regions created when the phenomenon is mapped. Despite this, the way in which features are classified may vary depending on the goals for which the taxonomy was developed. To ensure accurate representation of a phenomenon, it is important to understand the context under which original categorizations were developed prior to their acceptance in subsequent studies. The classification scheme developed for image analysis by Anderson et al. (1976) may be used to demonstrate this situation. This taxonomy was originally developed in the eastern United States for mapping land use and land cover. Though intended to be general and flexible, this scheme may not be optimal for general studies of natural vegetation. For example, chaparral vegetation of Mediterranean climates might be classified "rangeland" at the coarsest level of the taxonomy, even though no grazing is possible, the above-ground biomass may be similar to forests, and the community has very unique compositional and structural properties.

In addition to understanding the generalizing characteristics of a chosen observation, interrelationships in the system under study must also be understood. GIS-based analysis has increased the number of social and environmental variables that may be cross-referenced to answer a query. As with the problem of multicollinearity in statistics (Montgomery and Peck, 1982), there is the potential to mistake the correct source of causation from among a number of correlated variables in GIS-based analysis. Thus, approaching spatial data analysis without sufficient understanding of the subject under study may result in inaccurate inference. The use of correlative relationships may limit the accuracy and extensibility of observations used to address a query. The variable accuracy of automated classifications derived from remotely sensed data exemplifies this general problem of correlative relationships. Spectral reflectance characteristics of the landscape may be correlated with land use/land cover patterns, but specific spectral reflectance patterns are not always inherently tied to classes with informational value. Thus, automated identification of certain classes may be difficult, and extensions of locally generated relationships between spectral reflectance and land use may break down over time and/or distance. Classification schemes that causally link information classes to spectral response have provided some improvement in accuracy (Jensen, 1978; Running, Loveland, and Pierce, 1994), although the information content of these classes may be specific to certain applications.

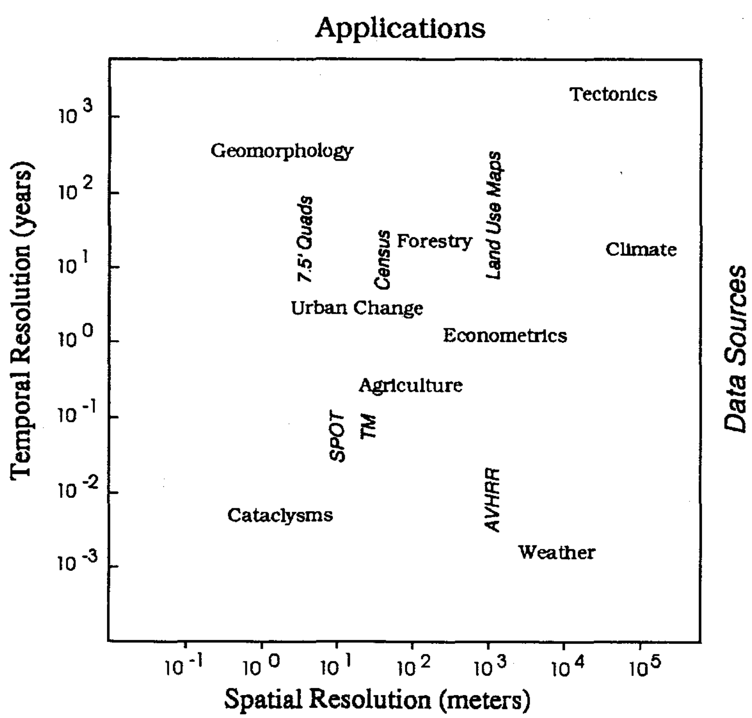
The scale of observation used to characterize a system may also have a profound effect on the accuracy of information derived from spatial data processing. In answering a query, the extent and detail of observation required to delineate a system in time and space must be determined. The resulting spatial and temporal scales of observation establish a limited frame of reference regarding the process under study. Both data and analytical methods may be tied to specific, but potentially conflicting, frames of reference. Observed patterns in a system may vary widely depending on the degree of precision and range of conditions under which observations are made (Getis and Franklin, 1987; Turner et al., 1989a; Moore and Keddy, 1989). This situation has been widely documented from parametrization of evapotranspiration (Jarvis and McNaughton, 1986) to the modifiable unit area problem of spatial econometrics (Openshaw and Taylor, 1981). Several papers have addressed such scale-dependent patterns in both GIS and remote sensing applications

(Turner et al., 1989b; Townshend and Justice, 1990; Stoms, 1992; McGwire, Friedl, and Estes, 1993).

The analytical methods or process models that use spatial data may also be limited to the frame of reference under which they were conceived. This may arise when the driving processes of a system are dependent on the historical state of that system. For example, the presence of a certain forest species in an area might be more dependent on the occurrence of a recent disturbance (fire, landslide, etc.) than it is on temperature and moisture regimes. Though one might be able to characterize the climatic conditions that allow a single tree to grow at the current time, an understanding of the factors affecting the presence of the species in that region would typically require analysis at broader spatial and temporal scales.

The spatial and temporal scales selected to represent a system must be compatible with both the scope of required information and the expected range of system variability. The graph in Figure 6.2 identifies the approximate scales at which selected applications (horizontal entries) and data sources (vertical entries) occur. Many methods have been used to quantitatively describe the scale-dependent behavior of spatial data, including fractal, geostatistical, block variance, and power spectrum techniques (Ludwig, 1979; Lovejoy and Schertzer, 1988; Legendre and Fortin, 1989). These methods may be useful in determining whether the characteristic variability in a dataset is compatible with the temporal and spatial scales of process models and subsequent decision making.

Figure 6.2. Time/space scales of applications and data sources



6.3. Measurement

Once the general types and resolutions of observations that are needed to quantify a system are determined, the specific nature of available measures must be understood. Acquiring and entering data often comprise a large proportion of the total cost of a GIS (Kennedy and Guinn, 1975). Additionally, many environmental and social variables cannot be exhaustively quantified. The technical limitations of data collection, storage, and processing will affect the ability to quantify spatial phenomena accurately. As a result, the data products used in spatial analysis are often surrogates for desired observations, being selected on the basis of availability as much as their functional significance in the system under examination. At a more fundamental level, all measurements are imperfect or incomplete representations of a natural system. The ability to understand the accuracy characteristics of derived information products depends on understanding the characteristics of available data inputs. Although the philosophical basis of measurement and sampling theory is beyond the scope of this paper, accuracy issues arising from the precision and bias of measurements will be discussed in this section.

6.3.1. Precision

Whereas accuracy is defined as a measure of the difference between a measured value and the truth (often defined with reference to a source of assumed higher accuracy), precision is defined as the degree of detail in the reporting of a measurement and is most often determined by the characteristics of the measuring instrument. In principle, measurements should be reported with a precision that matches their accuracy, but that principle is often ignored in data processing applications when results are reported to the maximum precision available to the system. Products with a relatively low degree of thematic or spatial precision may still be useful so long as the resulting uncertainty is matched to the requirements of the original query. For remotely sensed data, spatial resolution will often be identified by pixel size or instantaneous field of view. In cartography, accuracy has typically been addressed through the map scale (in this context, a representative fraction) and the existence of a variety of map compilation standards. As an example, the National Map Accuracy Standard (NMAS) of 1947 historically used by the U.S. Geological Survey requires that 90% of samples selected from well-defined points are within 0.025 inches (0.64 mm) of their correct positions. Thus, a 1:24,000 scale map might be assumed to have a spatial resolution of about 15 m. However, the assumptions that relate map scale to spatial resolution may be misleading, especially since many land cover maps show only those features that exceed a larger threshold size, termed the minimum mapping unit (MMU). Thus, for some maps the positional accuracy of polygon boundaries may be less useful as an indicator of spatial precision than the MMU, particularly with regard to land classification.

In the context of GIS, the predominant form of measurement occurs in the digitizing or scanning of maps. Given the traditional reliance of GIS on inputs from analog maps, much attention has been focused on uncertainty in the creation, digital representation, and registration of feature boundaries. One of the classic monographs in this area by Peucker (1976) addresses the fundamental nature of the cartographic line and relates the accuracy of digitizing to the rate of sampling with respect to the curvature of boundaries. The variability between operators in digitizing selected features has also been studied empirically (Maffini, Arno, and Bitterlich, 1989). As a result of variability in data capture, the spatial resolution of data obtained from a map may be somewhat coarser than that of the original product. However, the various stages of digitizing or scanning rarely introduce positional uncertainties of more than the 0.5 mm already present in most map products and allowable under traditional map accuracy standards.

In remote sensing, the spatial, spectral, and radiometric resolution of a sensor determines the accuracy with which features may be measured. Positional accuracies for the *Landsat* Thematic Mapper and *SPOT* sensors have been likened to those of 1:50,000 and 1:25,000 scale maps, or 32 m and 15 m, respectively (Welch, Jordan, and Ehlers, 1985; Konecny et al., 1987). The geometric fidelity of digital imagery is affected by both sensor and platform characteristics. Though airborne scanning systems are capable of higher spatial resolution than satellite-based platforms, aircraft platforms are subject to greater variability in platform attitude and altitude. Their proximity to the ground also increases topographic distortions and off-nadir scan angles. A number of sophisticated efforts at correction of airborne scanner data are being developed, including real-time GPS measurements and more sophisticated correction algorithms (Fisher, 1991; Ehlers and Fuller, 1991; Fogel and Tinney, 1994). These issues are addressed in more detail in Chapter 2 of this monograph.

The spatial accuracy of satellite-based image data continues to increase with both improved spatial resolutions and wider availability of precision topographic correction. This increase in spatial resolution combined with pointable sensor design will also allow topographic mapping in remote areas at a much higher level of accuracy than may be found in existing map products. Though higher spatial resolution tends to increase the information content of image data, the associated increase in scene complexity may make automated classification of certain surface features more difficult (e.g., Toll, 1984; Williamson, 1989). In contrast to increased spatial resolution, sensor design in the coming era of remote sensing for global change research is focusing on the development of more precise spectral measurements at coarse spatial resolutions, which will facilitate global data coverage (e.g., MODIS-N has 36 spectral channels).

In addition to accuracy limitations in manual data capture or direct measurement, the accuracy of data sources may be altered by spatial processing. As previously discussed, the validity of attribute measurements may be tied to the spatial resolution of original observations. Processing may increase or decrease the spatial precision of derived data products in a way that may lead to inaccurate interpretation. In one sense, features in data products may be represented with spurious precision. Such a situation may occur in the vector domain through compositing of data with differing spatial resolutions into a single data product. In the raster domain, data may be resampled to finer resolutions using interpolation criteria that do not take into account the underlying distribution of the phenomenon being measured. For example, cubic convolution is often used to resample image data because the resulting product is visually more appealing than alternate methods. However, resulting data values may exceed the actual range of reflectance. Similarly, the new generation of digital elevation models being generated by the U.S. Geological Survey are interpolated from hypsometric data in 7.5′ quadrangles. Although the models provide a better behaved surface than the previous Gestalt photomapping methods, this interpolation uses limited information on the nature of real topographic surfaces and its assumptions may be naive. More specifically, the interpolation uses a fixed weighting function rather than adapting to local patterns of variance.

Alternatively, processing may degrade the spatial precision of data products by convolving values found within a neighborhood. Examples of such operations include calculations of slope and aspect from digital elevation models or low-pass filtering of digital image data. Typically, algorithms that calculate slope and aspect do not resample output values to the neighborhood over which they are representative. Using these derived products at the spatial precision of source data may cause inaccurate interpretation, especially if the scale of the original observations was marginal for a particular application. In some cases, the Nyquist sampling theorem may provide guidance in determining the limitations of spatially convolved data products.

6.3.2. *Bias*

A second accuracy issue is whether measurements display bias with respect to attribute or position. Whereas random fluctuations may cause imprecision in measurement, bias refers to systematic differences in measurement characteristics, which may also depend on the location or time of measurement. Bias may range from a lack of completeness in enumeration to an inconsistent relationship between a phenomenon and the measurement technique being used. Information products derived from biased measurements may misrepresent relationships in a system if the bias is not recognized (for further details, see Section 6.4). The simplest case of attribute bias is that of miscalibration – a systematic inconsistency regardless of location and land cover type. For example, spectral measurements in remotely sensed data may be systematically misrepresentative of surface reflectance if sensor calibrations and atmospheric corrections are not properly applied. In GIS analysis, simple bias in positional accuracy may arise when data are digitized from map media with unstable shrink/swell characteristics. Changes of up to 2% in the dimensions of paper map products have been documented with extremes in temperature and humidity (Monkhouse and Wilkinson, 1973). Failure to recognize differing datums in map sources may also cause simple positional bias (e.g., NAD27 vs. NAD83).

When bias in a data product is dependent on the feature, location, or time of measurement, it may become difficult to state in what ways resulting information products are affected. In such a case, the risk of using information products may vary based on the alternatives being considered. Complex biases may affect both the attribute and position of mapped features. An example of feature-dependent, attribute bias occurs in map products when uniformity of detail is sacrificed for selective representation of those features that are useful for orientation (Roth, 1991). Interestingly, the 1947 NMAS does not

address the thematic accuracy of mapped features. Similar feature-dependent, thematic bias is encountered in automated land use/land cover classifications derived from remotely sensed imagery. Because relationships between an object's spectral characteristics and its rank and associations in a land use/land cover classification scheme may be weak, errors in such machine-generated land cover maps are generally not consistent across classes. As a result, the overall map accuracy statistic may not be relevant to a particular application. Though the Anderson classification scheme (Anderson et al., 1976) developed for image interpretation addresses thematic bias by specifying that accuracy "should be about equal" for all map classes, this criteria acts only as a guideline for product compilation and does not provide a basis for conceptual error modeling.

Attribute bias may also be dependent on the size of the feature being measured. In remote sensing, the spectral response of surfaces will be blurred by both the intervening atmosphere and sensor optics. As a result, spectral measurements for smaller features may not be as accurate as for features covering large fields of view (Kaufman and Fraser, 1984). Similarly, in the GIS domain, Turner et al. (1989b) document the preferential deletion of specific land covers with small spatial extent as spatial precision decreases. Attribute bias may be location dependent as well, as in the case of uncorrected atmospheric effects that vary within a remotely sensed image.

Positional bias may also vary in a complex manner depending on location. An example of variable positional bias can be observed in image data that has not been corrected for topographic distortions. These distortions are localized scale changes and occur as a result of varying distance between the imaging system and the land surface (Paine, 1981). Complex positional bias may also be introduced to planimetrically accurate source data through nonlinear coordinate transformation. For example, confusion regarding map projection parameters may create distortions that surpass simple coordinate offset.

6.3.3. *Temporal issues*

Though the temporal dimension is usually fixed in spatial data products, issues of measurement precision and bias are still relevant. Efforts are currently being directed at the compositing of global or regional data sets to support global change research (e.g., IGBP, 1992). Such composite data must capture a consistent view of the landscape despite dynamic surface and atmospheric processes that change during and between the times of image acquisition. Despite such time-dependent variations, the high temporal repeat rate of the Advanced

Very High Resolution Radiometer (AVHRR) sensor allows composites of this data to correspond well with certain environmental parameters (Tucker et al., 1983; Prince and Tucker, 1986). Temporal inaccuracy in GIS-based analysis may also result from a lack of database concurrency, since a GIS database is a static representation of what is often a dynamic system. Not only might the database misrepresent the current status of a natural system, but data acquired from different periods may create a representation of states that never exist contemporaneously. Such concurrency issues suggest that a method of assigning lifetimes to data products should be developed. However, the enforcement of temporal validity may be intractable due to the unpredictable or discontinuous nature of certain processes. The issue of database concurrency is a prime motivation for more sophisticated integration of remote sensing and GIS technologies.

6.4. Format

Digital analysis may allow more rapid and flexible assessment of complex systems than manual methods. To manipulate digital measurements effectively, some common framework for data analysis must be adopted. The selection of a particular database structure, or data model, may significantly affect one's ability to pose queries and derive information (Date, 1986). In GIS, the choice of a data model affects the inherent capacity both to represent spatial phenomena and to characterize product accuracy. Ideally, the choice of a data model should be driven by the need for accurate representation of real spatial variation, in order that decisions based on GIS analysis be as reliable as possible. In reality, the choice of a data model is often driven by the limited capabilities of particular software choices, by the constraints of measurement systems, or by the user's experiences and biases.

Measurements of simple scalar values, such as the distance between two points or the height of a tree, are easily represented as numbers and are readily transferred to the digital environment. However, the digital representation of spatial variation requires much more sophisticated approaches. Spatial data handling systems provide a variety of data models for defining attributes within a two-dimensional field (Goodchild, 1992). In the context of remote sensing–GIS integration the user is frequently limited to two data models, termed here as the raster and polygon models. Each data model has advantages and disadvantages in terms of maximizing and quantifying the accuracy of spatial representations. The raster model quantifies landscape attributes within the explicit control of a systematic sampling frame. The term "polygon model" is used in this chapter in place of the more common term "vector" in order to clarify certain functional requirements of a data model for GIS analysis. We define the polygon model as a piecewise approximation to a two-dimensional field, in which the plane is divided into nonoverlapping and arbitrarily shaped regions based on the attribute being represented. This is contrasted with the limited representational capabilities of vector-based computer aided drafting (CAD) packages, which also build polygons from line segments, and line segments from points, but do not enforce consistent organization of features within a two-dimensional field.

To maximize the accuracy of a spatial representation, the choice of either data structure is dependent on the features being sampled. The polygon data model is generally used to represent various types of area classification, or area class maps, for landscape attributes such as soils, land use, or land cover. Although allowing precise measurement of the positions of polygon boundaries, the geometric precision of the polygon model may have little to do with accuracy. Mapped boundaries are often no more than crude approximations to broad zones of transition, and the polygons they define are often far from homogeneous. The potential of the polygon model for greater positional accuracy is justified for cases where changes in land characteristics occur along well-defined boundaries between relatively homogenous spatial units. Those features which are continuous, or for which the scale of observation makes precise taxonomic determination impossible, might be quantified most effectively within the systematic sampling of the raster model. As an example of the former case, Kumler (1992) documents greater accuracy in representing continuous topographic surfaces with raster digital elevation models than with the vector encoding of a triangulated irregular network (TIN). Certain data sets used in general circulation models (GCMs) represent the latter case by coding land use/land cover as percentages of primary and secondary classes within a raster. The spatial control of the raster data structure makes systematic determination of percent cover for each land cover type far more tractable than would be possible in the vector domain. However, the traditional, rectilinear sampling of the raster model may create distorted representations of spatial variation as well. For example, to reduce problems with directional bias found in the typical raster model, Burroughs (1988) utilized a hexagonal tessellation for a fire simulation. The challenges of global scale studies may also exceed the representational capabilities of simple, raster data structures. In response to this, Goodchild and Yang (1992) have developed a tessellation of the globe based on triangular decomposition of an octahedron for which facets have approximately equal area and shape.

In an integrated remote sensing and GIS environment, both data models must often be used in conjunction. The pixels of a raster data file act as the spatial objects whose numerical attributes form an approximation of the two-dimensional field being measured. In some applications, such as land cover classification of image data, these attribute values may be used to classify each pixel into one of several classes. Contiguous pixels with identical attributes may then be grouped to form zones of uniform class, and the boundary of each zone may be identified as an ordered set of coordinate pairs in some suitable coordinate system. This process produces the polygon model, as spatial variation is now described by a partitioning of the space into irregularly shaped polygons. Although each polygon will be homogeneous with respect to class, there will likely be substantial within-polygon variation in spectral response. This will be especially true if the polygons are subsequently aggregated in order to generalize the representation to coarser spatial or taxonomic resolution. However, if raster-derived polygons are not suitably generalized before converting to the polygon model, topological relationships may become confused along complex polygon boundaries. Alternately, conversion from the polygon to raster model also introduces uncertainty with respect to the position of feature boundaries. Frolov and Maling (1969) use the size distribution of cells bisected by line features to provide an error estimate for this effect. Goodchild (1980) later refined this estimate by introducing the effect of serial correlation in line segments.

The chosen data model will also affect the ability to characterize uncertainty in data products. Perhaps the most commonly cited method for characterizing uncertainty of boundaries in the polygon data model is the epsilon band (Perkal, 1956, 1966; Blakemore, 1984). The epsilon band is a zone around the observed position of a line within which the true position of the line is expected to lie with some measure of confidence. Some sources of error, such as digitizing, may contribute a constant epsilon. However, for many applications of GIS to land classification, the variable width of transition zones between adjacent polygons cannot be adequately represented by a constant epsilon distance. Although the epsilon band concept is not necessarily limited to the polygon data model, its implementation in a raster approach may be inefficient. This inefficiency arises because topological information is not explicitly stored to indicate which neighboring classes might be confused, and the spatial resolution required for characterizing boundary uncertainty may be more detailed than that of the original raster data file. Although the epsilon band provides a useful way of describing uncertainty in a line's position in the polygon model, it has not been possible to formalize it as a parameter of a statistical model of uncertainty or to make rigorous connections between it and those statistical models that have appeared in the literature (Keefer, Smith, and Gregoire, 1988; Goodchild, Sun, and Yang, 1992). Thus, the epsilon band remains a useful, but isolated, concept.

In general, not all points contained within a polygon will be correctly represented by the assigned class. Spatial information on such within-class variance may be more easily represented with the raster data structure. Continuous error estimates within the field of measurement may be possible for both categorical and continuous raster attributes. Field-based error modeling within categorical data products is typified by the per-pixel confidence values derived from maximum likelihood classification of land cover in remotely sensed image data. The kriging interpolation technique provides an example of field-based error modeling for continuous, spatially autocorrelated measures. Kriging uses an empirical model of spatial autocorrelation to create error estimates for every interpolated point (Journel, 1989). As with the polygon data model, attribute heterogeneity occurring within raster cells will generally be unavoidable. The kriging approach may be implemented using a block method to appropriately estimate error variance within a grid cell, rather than for a specific point. Unfortunately, it has not been possible to make analytic connections between the field-based view of uncertainty and that inherent in cartographically based descriptors, as noted previously with the epsilon band.

6.5. Analysis

The topological foundation of GIS permits assessment of spatial relationships beyond the abilities of computer-aided drafting and relational database approaches. GIS permits a holistic approach to system characterization that allows unique information to be synthesized from disparate data sources. However, the ability to accurately integrate multiple data sources is first dependent on the degree to which absolute geometric registration between data sources can be enforced. Inconsistencies in boundary location between data sources result in the creation of spurious, sliver polygons during intercomparisons. Boundary uncertainty between data products is pervasive, as perfect repeatability is not possible in either map compilation or in the digitizing process. Specific problems arising from cartographic overlay have received much attention (Mead, 1982; Newcomer and Szajgin, 1984; Veregin, 1989; Chrisman, 1989). Goodchild and Gopal (1989) ascribe this difficulty in computerized spatial representation to an inability to compensate adequately for differences in map characteristics in a manner that is comparable to manual

Table 6.1. *Mathematical operations associated with measurement types*

Measurement	Characteristics	Examples	Valid Operations
Categorical	Classification into a taxonomy where the ordering of class values is arbitrary.	Soil series Political jurisdiction Acceptable/Unacceptable	Equals
Ordinal	Relative ordering made with unknown or unequal intervals between groupings.	Highest → Lowest Fastest → Slowest Standard of living	Less than Greater than
Interval	Continuous measurement using equal intervals made from an arbitrary zero point.	Sensor DN value Elevation above sea level Degrees Fahrenheit	Addition Subtraction Scaling by a constant
Ratio	Continuous measurement using equal intervals made relative to an absolute value of zero.	Calibrated irradiance Income Degrees Kelvin	Multiplication Division

interpretation. Misregistered features not only change positional or area estimates in resultant information products, but may also generate inappropriate relationships between landscape features. In this latter case, derived information products may inaccurately represent the coexistence of environmental or cultural features of the landscape. As mentioned in the previous section, the epsilon band approach may allow for automated resolution of boundary uncertainty.

The accuracy of spatial data analysis is also dependent on appropriate application of mathematical, statistical, and process-emulating manipulations. An example of limitations in mathematical manipulations may arise in the common GIS analysis of generating a weighted, linear combination of data products. This approach may be used in site suitability scoring or in generating indices such as the universal soil loss equation (USLE). Commercial GIS systems do not currently have the metadata management capabilities to enforce mathematical logic in overlay operations and as a result may allow invalid relationships between data products. Index-based approaches are also common in digital image processing (Kauth and Thomas, 1976; Tucker, 1979; Crist and Cicone, 1984), and the reliability of such transformations depends on an adequate match between data calibration and applied mathematical manipulations. Measurements in a GIS may be categorical, ordinal, interval, or ratio in nature (Harvey, 1969). The characteristics and valid mathematical operations for these data types are summarized in Table 6.1. Those operations indicated as valid in the initial entries of Table 6.1 will also be valid for the data types listed subsequently. The accuracy of index-based methods is also dependent

on weighting schemes that accurately reflect the process under study. Whereas physical processes are often statistically parametrized, the creation of valid weighting schemes for representing social value systems, such as desirability, may be quite difficult. Such weighting schemes must be derived from unbiased, informed consensus and are generally difficult to translate into an interval or ratio measurement scale.

Information derived from spatial data products through statistically based analyses will be constrained by the assumptions of statistical techniques, which may in turn be confounded by the effects of spatial autocorrelation. Although the assumptions of various statistical methods are beyond the scope of this chapter, selected problems that are common with spatial data will be mentioned. Social and environmental data often violate the assumptions of multivariate normality required by classical statistics. Studies have demonstrated how inappropriate assumptions of multivariate normality in spectral data reduce the accuracy of automated land cover classification techniques (Maynard and Strahler, 1981; Skidmore and Turner, 1988). Curran and Hay (1986) demonstrate how measurement error in remotely sensed data may cause biased estimates in regression models for landscape parameters. This problem of error in regressors is generalizable to predictive relationships derived from map data. Multicollinearity – the existence of linear relationships between explanatory variables – is also common and presents a problem in regression modeling (Montgomery and Peck, 1982). In cases of multicollinearity, variance estimates for regression weights derived from ordinary least squares are inflated, resulting in potentially unstable values. As mentioned in Section 6.1, multicollinearity presents the danger of mistaking causation for correlation. Heteroskedasticity – the dependence of error variance on the magnitude of a measurement – frequently occurs in social and environmental data as well. Although there are robust techniques that may be useful in dealing with heteroskedasticity, this situation still provides problems in terms of efficient parameter estimation.

Finally, spatial autocorrelation, the tendency of proximate samples to have similar values, is practically universal in spatial data. This condition may violate the independence of samples required in classical statistics, resulting in underestimated sample variance and inflated confidence estimates. Techniques used to characterize spatial autocorrelation may include summary statistics such as Moran's "I" or graphic approaches such as semi-variogram or block variance analyses. The effects of spatial autocorrelation have been shown to reduce the accuracy of statistical land cover classifications when representative samples are not randomized (Craig, 1979; Campbell, 1981;

Labovitz, 1984). These effects have also been studied as they relate to local image variance (Woodcock and Strahler, 1987; Jupp, Strahler, and Woodcock, 1988, 1989), biophysical parametrization (McGwire et al., 1993), and error assessment (Congalton, 1988ab). Methods based on these findings should be developed to improve digital classifications, drive sampling methodologies, and deflate confidence estimates. The general lack of knowledge of methods for working with spatial data and a lack of integrated statistical tools within existing software packages are major impediments to error assessment in the analysis phase. The development of flexible statistical tools that take into account the particular difficulties of spatial datasets and the organization of these tools into a usable software environment may encourage adequate consideration of statistical assumptions in the development of higher order information products. Work in this area is currently being pursued with software such as the SPACESTAT package developed through the National Center for Geographic Information and Analysis (Anselin, 1992).

Though direct process modeling of social and environmental phenomena involves difficulties with simplification and suitable specification of boundary conditions and forcing functions, this approach may be more theoretically sound and generalizable than empirical methods based on statistical analyses. To explicitly model physical or social processes, GIS data are often exported to specialized, discipline-specific software. Results are then imported again for integrated assessment. Examples of such simulation environments include ties between hydrological data in a GIS and ground water models (Nystrom et al. 1986; Foresman, 1984) or links between crop classification strategies for imagery and econometric modeling (Schultink, 1982). The ability of process models to pass some indication of accuracy for derived products back to the GIS will vary. Simplifying assumptions, such as assuming independence of error between model components, may allow the propagation of error variance through a model to be estimated. Kerekes and Landgrebe (1991) provide such an example in their simulation of remote sensing systems. However, for complex nonlinear models or in cases where simplifying assumptions are not reasonable, the ability to estimate error propagation through a model may be limited to generic sensitivity analyses. In such cases, GIS and remotely sensed data may go beyond their role as a data source and might even be used to calibrate process model outputs. For example, Maas (1988) describes the use of spectral data to keep climatically based agricultural yield models on track with actual field conditions.

Process-oriented relationships between spectral reflectance and surface parameters are being studied in an effort to increase estimation accuracy for land surface parameters. Such efforts include development of directly invertible models of radiative transfer (Goel and Grier, 1986a, b, 1988; Sellars, 1985) and spatial variance (Li and Strahler, 1985; Franklin, 1988). One of the major difficulties with invertible modeling approaches is the requirement for large volumes of data (e.g., multiple look angles, ground surface characteristics, etc.). A possibly more tractable approach that is being tested estimates the physical composition of pixels through absorption features as measured by high-resolution imaging spectrometers. Examples include identification of surface mineral composition and plant canopy chemistry through distinctive spectral absorption features (Huete, 1984; Swanberg and Peterson, 1987; Kruse, Calvin, and Siznec, 1988). Inference of more abstract features in image data, such as land use, requires a complex understanding of natural and cultural systems. Expert systems and contextual classifiers have been tested to identify such features in image data (e.g., McKeown, Harvey, and McDermott, 1984); however, the complexity of the knowledge domain required to identify abstract features limits application of these methods to very specific tasks.

6.6. Assessment

As evidenced in the preceding sections, spatial data processing is an abstraction in which care must be taken to ensure that actual relationships in the system under study are accurately estimated. Therefore, it is critical that the validity of derived information products be tested to provide a reasonable estimation of confidence for use in decision making. Accuracy information is required in a decision-making process in order to understand the risk involved in relying on GIS-based information products. Such information may be of great importance in selecting between alternatives with respect to a particular risk-taking behavior. Part of the challenge of this accuracy assessment lies in direct quantification and visualization of product error (Beard, Buttenfield, and Clapham, 1991). Indirect assessment methods also play a valuable role in ensuring that high-quality information is produced by the spatial data processing flow. These indirect methods include conceptual and empirical models of the sources and propagation of error, as well as its impact on subsequent decision making.

The type of accuracy assessment required may depend on whether the results are relative or absolute measurements. In cases where simple information on distance or area is derived from a single data source, error such as a simple coordinate offset may not be significant. However, as described in the previous section, information derived from multiple spatial data sources will generally require enforcement of absolute positional accuracy. A similar dichotomy applies to thematic accuracy assessment when derived information products are either interval (relative accuracy) or ratio (absolute accuracy) in nature. Some authors suggest dividing accuracy assessment in GIS operations between attribute and locational components (Vitec, Walsh, and Gregory, 1984; Walsh, Lightfoot, and Butler, 1987). Such a division may be naive for GIS and remote sensing representations of continuous variation in fields, since the spatial objects that populate the database are to a large degree artifacts of the process of representation.

There is a large body of literature on direct thematic accuracy assessment in remote sensing (Hord and Brooner, 1976; Card, 1982; Aronoff, 1982a, b; Congalton, Olderwald, and Mead, 1983; Rosenfield and Fitzpatrick-Lins, 1986), and an extensive collection of these articles has been compiled by Fenstermaker (1994). These efforts use contingency matrices to compare database contents with samples derived from ground survey or some other information source in which there is a high degree of confidence. These matrices provide detailed information on the types and magnitudes of error found in original data or derived information products. In remote sensing classifications, the matrix typically relates the class assigned to a pixel in the database with the class determined for the same pixel by ground survey (per-point assessment). In many GIS applications, which use the polygon data model, the error matrix may compare the class assigned to an entire polygon with the class assessed by visiting the polygon in the field. This per-polygon assessment clearly omits within-polygon variability from the definition of accuracy.

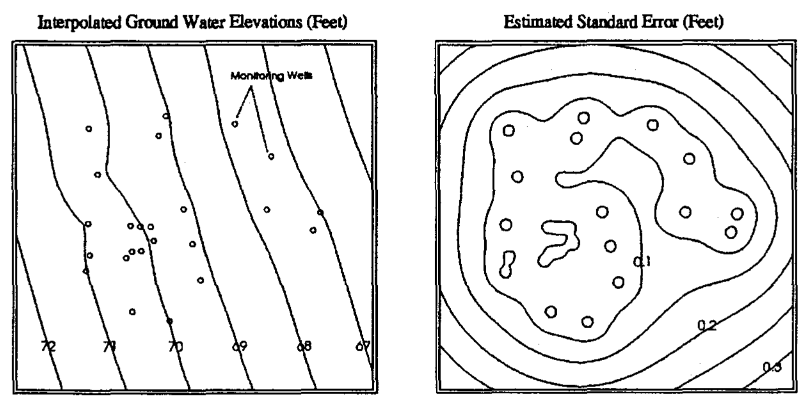
The simplest statistic derived from the contingency matrix is the percent correctly classified (PCC) or the percent of cases falling on the diagonal of the matrix. The matrix may also be examined using row or column aggregates to test the accuracy of the map product with respect to estimated errors of production and subsequent use (Aronoff, 1982a; Story and Congalton, 1986). Row and column statistics may also provide user and producer accuracies for individual classes, but they lack the sensitivity to describe cases where accuracy is strongly dependent on confusion between specific class pairs. These dependencies will be represented by off-diagonal entries in a contingency matrix.

Because some points will be classified correctly by chance even in a random assignment of classes, PCC is often rescaled to discount this effect, yielding the kappa statistic (Congalton et al., 1983; Rosenfield and Fitzpatrick-Linz, 1986). The kappa statistic is sensitive to the off-diagonal entries of a contingency matrix. In addition, the distribution for the kappa statistic is

asymptotically normal; thus, the significance of differences between alternate map products may be tested (Congalton et al., 1983). At present, no analytic connection has been made between kappa and conventional measures of positional map accuracy, such as epsilon. Despite the utility of both the PCC and kappa statistics, these measures reduce the dimensionality of error characterization to a single metric and may never adequately describe products with variable class accuracy. Basically, any reduction from the full contingency matrix to a smaller set of representative statistics reduces information content. Thus, presentation of the full contingency matrix along with thematic data products may be required for proper assessment of product accuracy or suitability.

Methods based on error matrices assess representations for fields of categorical variables, such as land cover class or soil class. Although this accounts for much of the information in GIS databases, it is also important to assess fields measured on continuous scales, such as spectral response or topographic elevation. This class of error estimation has been addressed in both GIS and remote sensing literatures. For example, McGwire and Estes (1987) compare the error assessment capabilities of moving-average and kriging interpolation methods. Using cross validation, a single error statistic may be generated from moving-average interpolations. In contrast, kriging yields a field of error estimates, providing a better understanding of the positional dependence of uncertainty. Figure 6.3, taken from McGwire and Estes, presents an example of the accuracy assessment made possible by kriging. Whereas cross

Figure 6.3. Error assessment using the kriging technique



validation of a moving-average interpolation provided a product with a single standard error estimate of 0.81 ft, the kriging method reduces uncertainty and shows how error would be expected to vary throughout the site. Atkinson (1991) demonstrates the use of a similar technique for determining an accurate estimate of the unbiased mean of a continuous variable for samples within a pixel.

In addition to assessing the accuracy of attribute measurement within a field, it may also be important for the decision maker to understand the positional accuracy of features. Uncertainty in the placement of a point can be characterized as a two-dimensional probability density function centered on the observed location of the point. It is commonly assumed that errors in the x and y directions are uncorrelated and normally distributed with the same variance, so that the distribution is circular normal. A common statistic based on this distribution is the circular map accuracy standard (CMAS), which is defined as the radius of a circle centered on the observed point and containing the true location with 90% probability, or more generally, the 90th percentile of the distribution. CMAS is commonly used as the basis of map accuracy standards, including the National Map Accuracy Standard of 1947. Other statistics, such as the root mean square error (RMSE) and the mean square positional error (MSPE), are also used to describe the probabilistic position of points. The majority of image processing and GIS software packages derive RMSE as a diagnostic of the geometric transformations used to coregister data. However, this diagnostic should not be confused with an independent accuracy assessment because test points are not independent of the transformation parameters. The resulting RMSE value is therefore likely to provide only a best case estimate of positional error. In contrast to point data, no satisfactory models of positional uncertainty currently exist for lines. Uncertainty in the position of linear features must be handled differently from points because adjacent positions along a line are not likely to be independent (Keefer et al., 1988), because the direction in which displacement occurs becomes ambiguous with respect to the x or y coordinate dimensions, and because it is possible to match observed and true locations only in limited circumstances. As mentioned in Section 6.3, uncertainty in boundary location has been described, but not modeled, using the epsilon band approach (Perkal, 1956). Blakemore (1984) demonstrates such use of the epsilon band to return uncertain responses to queries about the containment of a point within a polygon.

In order to track error accumulation effectively, methods are required to assess the generation of error associated with specific processes and to keep

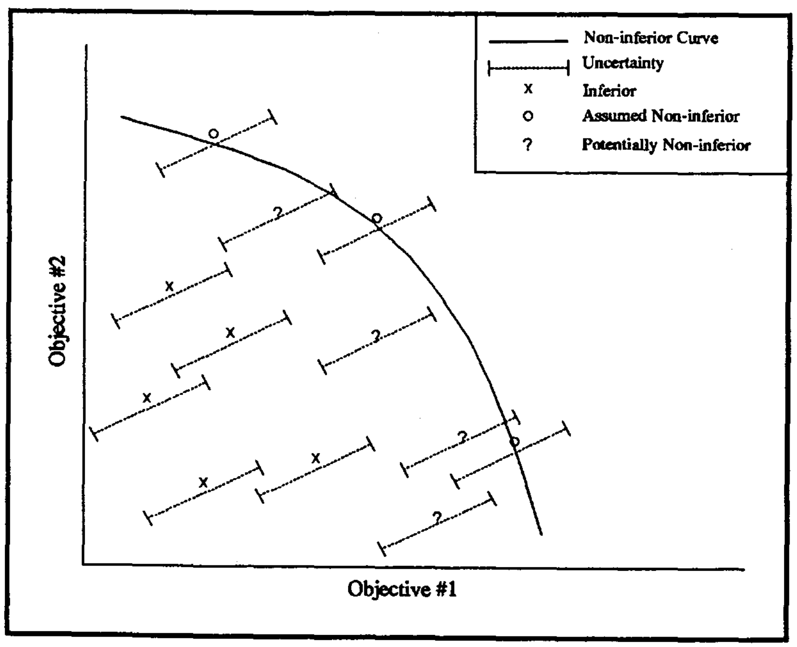
an accounting of the spatial, temporal, and attribute characteristics of this accumulating error. Methods for providing a transcript of data processing histories exist in many commercial remote sensing and GIS packages. However, this capability has generally been unsophisticated and has not allowed for specific inclusion of error characteristics. An integrated solution to tracking the data processing flow, called lineage tracing, is described by Lanter (1989). This approach uses a LISP language shell in which the Arc/Info GIS package (produced by Environmental Systems Research Institute) is run. The described algorithm allows automated backwards and forwards reconstruction of intermediate data products between data inputs and information outputs. Ongoing research is focusing on the application of this technique to modeling error accumulation in GIS information products (Lanter and Veregin, 1990).

Clear communication of spatial data using graphic and text products is also critical for accurate representation of information to decision makers. Much of the work in this area has already been addressed by the long tradition of manual cartography. However, whereas traditional cartography has focused on the legibility and perception of geographic information, recent recommendations suggest that visual representations of uncertainty associated with geographic data be provided as well (Beard et al., 1991; Lunetta et al., 1991). In going beyond these visually oriented approaches, the relationship between map accuracy and specific information requirements of the environmental and policy domains must be evaluated. Accuracy of information products must be evaluated in relationship to the risks of a management decision, whether regarding agricultural production or deforestation policy. This evaluation might be explored in the context of operations research methodologies for multiobjective decision making (e.g., Hobbs and Voelker, 1978). Figure 6.4 plots the attractiveness of possible alternatives with respect to two different objectives. Examples of these objectives might be reducing environmental impact versus cost of operation. In this graph, a curve of noninferior alternatives may be identified that represents the trade-off between objectives. Uncertainty may vary between the information products used to quantify these objectives. As a result, the position of points relative to the noninferior curve is probabilistic and identification of noninferior sites may be less obvious. At present, methods for incorporating uncertainty in multiobjective decision making are not well developed (Solomon and Haynes, 1984). However, such techniques may benefit from the formalization of a comprehensive model for accuracy assessment.

6.7. Conclusion

Our understanding of accuracy issues in spatial data processing has yet to be fully described within an accepted, coherent framework. Several sources provide intensive investigations of error sources and modeling. Kerekis and Landgrebe (1991) have simulated remote sensing systems to the point of predicting the effects of spatial autocorrelation on supervised land cover classifications. Veregin (1989) provides an excellent organization and review of error assessment and modeling techniques. However, the goal of a coherent system that integrates efforts such as these has proven elusive. Such a conceptual framework would allow better understanding of an information product's "fitness for use" (Chrisman, 1991) in a given application. At present, the level of understanding of accuracy issues in the research community is well in advance of the corresponding level of understanding in current practice. Unfortunately, significant improvement in the accuracy of spatial data or in the ways uncertainty is dealt with in practice will only occur at substantial cost. Nevertheless, computational capabilities have been increasing at a dramatic rate, and the additional processing required by methods for modeling and visualizing uncertainty is increasingly manageable. Further research may also develop more efficient methods for characterizing accuracy and resolving uncertainties.

Figure 6.4. Plot showing variable confidence in quantifying objectives



Several areas require further investigation if accuracy is to be improved or error characteristics better understood. In remote sensing, technical developments such as advanced geometric rectification capabilities and feature extraction methods will increase the quality of data that may be brought into GIS-based analysis. Accuracy characteristics will vary between data products, and metadata management capabilities must be developed to make these characteristics accessible to users and error tracking processes. Methods for understanding the effects of spatial data processing on product accuracy and managing these effects in ongoing analyses are required. As an example, a system might alter information on the spatial precision of a secondary data product that is generated by a neighborhood operation. The challenge of statistically characterizing spatial processes also stands, both in developing accurate predictive relationships from map-based variables and in describing the interdependent roles of thematic and positional error. Methods for using accuracy information in spatially oriented decision making must also be better developed in order to understand specific risks associated with using information products.

An alternate approach for developing an accuracy model might focus on decision-making requirements rather than the specific processes that create error. This approach would focus on specifying the minimum set of spatial and aspatial descriptors for information product accuracy that would be required in the decision-making process. The resulting standardization of information required from various GIS software packages might then highlight existing deficiencies and provide impetus for further development of critical components for internal representation and manipulation. This developmental approach emulates that of the relational database model, which in its early stages was specified by its functional appearance to the user rather than its internal structure and manipulation language (Date, 1986). To date, few relational databases are totally compliant with the full conceptual development of the relational model. However, specification of consistent, yet flexible, user interaction promoted wide acceptance and spurred interest in further development.