ABSTRACT

It is suggested that new ways of thinking about spatial analysis are required to optimize the use of digital technologies. “Real-world-to-map” simplifications must be quantified and incorporated into spatial analysis with particular emphasis on boundary and category error. This could also lead to a better way to integrate thematic map-based and remotely-sensed data. However, such developments will require the creation of a new family of spatial operators which work on “error” or “uncertainty” surfaces rather than categorical thematic maps. These operators should work across maps and polygon boundaries rather than by considering absolute location to be the primary characteristic of spatial entities. The use of these concepts in digital spatial technologies may also require the development of an alternative data structure. The Voronoi diagram is presented and discussed as a possible solution.

INTRODUCTION

Thematic maps show the categories for a particular “theme” over a given area. For example, a thematic map might indicate the population density -- persons per unit area -- within each “township” for a region of interest. The density could either be shown in “raw values” -- 2.2 people per hectare -- or categorically -- 1 to 3 people per hectare with the township polygons color-coded accordingly. Either way, such a map would show definite values for distinct areas, and polygons would be assumed to be uniform within. Finally, boundaries between townships would be mapped as (dimensionless) lines meaning that changes in population density between townships would occur instantaneously.

That such geographical representations have existed for a long time is not surprising. Indeed, most people are familiar with such storage formats for spatial data. However, it is this familiarity which makes one forget some of the inherent limitations in such maps. One is simply the constraints which are placed on any cartographer by “pen and paper technology.” That is, using lines as delimiters of “meaningful” polygons is an easy way to represent “the world” on a piece of paper. However, this conventional thematic map representation has also evolved to its present form due to data limitations. In the preceding example, one might actually desire population density by “section of townships.” However, if population records have not been kept for every “section,” one cannot produce such a map. Thus one must group units spatially until one has a datum for each polygon -- townships in this case. This also makes it evident that data availability will determine the size of the minimum mapping unit. Finally, existing thematic map representations are also employed partly because humans have a limited ability to comprehend complex spatial surfaces. Any person who has taught the interpretation of contour maps to first-year students knows the difficulties inherent in enabling students to understand even that familiar, and relatively simple, geographical representation.

Relative to the use of GIS and remote sensing technologies (which together we call Digital Spatial Technologies or DSTs) for natural forests, there are numerous implications of traditional map representations and their inherent limitations. In forest management, a considerable amount of effort is expended to produce “cover type” maps, “site quality” maps, etc. In creating such maps, polygon boundaries are treated as being dimensionless, identifiable, and locatable on the ground. Furthermore, polygons are treated as if they are completely uniform within (though one might attach confidence limits to
the attributes of a given polygon). Overlays of the polygon boundaries of separate feature-maps are done

deterministically using Boolean logic meaning that queries of resulting maps will indicate that a point is
either definitely “in” or “out” of a given polygon. Moreover, confidence limits for totals (for volume over
an area, for example) are based on statistical sampling errors exclusively, and not on potential errors in
boundary location and/or category identification.

It is apparent that in certain endeavors, the topics mentioned in the previous paragraph are not an issue. In
human geography, for example, boundaries may be, in fact, dimensionless, identifiable, and locatable.
However, making the assumption that, in the mapping of natural resources, boundaries can be represented
the same way may lead to a number of gross oversimplifications. Firstly, in any natural system, polygon
boundaries are not definite, nor are they unchanging over time. For example, the difference between a
Birch(Spruce) species mix and a Spruce(Birch) mix is relatively subtle. Thus the identification and location
of a boundary between similar types has a high potential for error. Furthermore, because one of these
types may gradually evolve into the other, the boundary is likely to change over time. Secondly, polygons
of natural features are not uniform within. Thus to treat an area in the center of a mature forest polygon,
for example, as definitely having the same volume as an area closer to the edge of a scrubland may be a
questionable procedure.

We believe that a number of issues must be addressed relative to the use of DSTs in the management of
natural forests. The purpose of this paper is to present these issues for consideration, discuss them in
some detail, and propose possible solutions. Specifically, we will focus on four subjects: new ways to
measure and use spatial errors, an alternative map representation to aid thematic map and remote-sensing
integration, the creation of “probabilistic” instead of deterministic spatial operators, and a suitable data
structure for the incorporation of these ideas into digital spatial analysis.

MEASUREMENT AND USE OF SPATIAL ERRORS

Selected spatial errors have been the subject of considerable DST research. These have focussed largely on
digitizing errors in GIS and/or geometric correction and geographic registration in remote sensing.
However, the study of digitizing errors reflects largely an implicit assumption that features being digitized
are identifiable and locatable. Thus the study of such errors -- including those of geometric correction in
remote sensing -- “make sense” in subjects such as human geography and plantation forestry. But in
natural forest stands, the very existence of a boundary line may be in doubt, not merely its location.
Considering these factors, digitizing and registration errors are likely to comprise a relatively small amount
of the “total error” in the conversion of natural spatial phenomena to a digital map product (Chrisman
1982).

The authors of this paper have identified two types of errors that we believe are more important in natural
forest stands than digitizing and registration errors. We refer to these as “boundary” and “category” errors.
Rather than being thought of as “mistakes” these should be interpreted simply as real-world-to-map
“simplifications.” Generally, one assumes that such errors are merely inherent in map-based data, and,
therefore, ignorable. Or that, because they are difficult to quantify and use, they must be assumed to be
non-existent. The implications of such assumptions can be severe.

For example, suppose that in a forest there is a swampy area directly beside a highly productive forest
type. On a map of the area, the boundary between the two would be drawn as a line thus suggesting that
each type can actually be precisely delimited. In “the real world,” however, the boundary between the two
would actually be a broad transition zone. If the cartographer consistently estimated that there is less
swamp than actually exists by placing the boundary in the wrong place, forest planning efforts could be
unreliable. Specifically, one would overestimate the amount of volume on the forest because the
cartographer had overestimated the amount of high-productivity land.

This is one issue that we are trying to address in DST research. It can be understood intuitively that some
natural boundaries will be relatively easy to identify and locate. One example is the boundary between a
clearcut and a mature forest. Similarly, some individual categories of stands are easy to identify such as
pure, even-aged stands of jack pine (Pinus banksiana Lamb.) in Québec. However, even with pure forest
stands, boundaries become difficult to identify if these stands are surrounded by a highly variable natural type such as “scrubland.” Recognition of this factor suggests that there may be a relationship between boundary and category errors.

Two projects are underway to attempt to identify and quantify both category and boundary errors and to establish a relationship between the two. In one project, within a boreal forest in Quebec 64 sample plots have been established on an 8-by-8 grid with plots spaced 50 metres apart. The dbh of all trees which are within 5.64 m of plot center and which are 10 cm dbh and greater have been measured. Similarly, the dbh of all trees within 3.57 m of plot center which are 1 to 10 cm dbh have been measured. Finally, the heights and dbh’s of three dominant or co-dominant trees have been measured at each plot. This grid of sample plots is in the process of being located on a photo-interpreted map of forest cover types. From this, we intend to determine if there is a relationship between humanly-interpreted aerial photographs and the ground-based forest inventory information. Because of the relatively tight spacing of plots, we should be able to identify and quantify points at which both boundary and category errors occur. These data will also be used with remotely sensed digital data collected from a sensor mounted on a fixed-wing aircraft. With this information we intend to evaluate the ability of high-resolution spectral data to detect forest type boundaries relative to the characteristics of each forest type.

The second project involves synthetic remote-sensing data generated from a computerized image generator. We believe the magnitude of boundary and category errors will be found to be related to the difference in the means and variances between two adjacent forest types. For example, the boundary between two forest types with similar means but widely different variances should be relatively easy to detect. Similarly, the boundary between two types with the same variance but different means should be easy to identify. With an image generator, one can control both the mean and variance of the simulated forest types. Images will be generated by establishing a template -- a thematic map -- of known types, assigning a mean and variance to each type for image generation, and generating an image from this information. By having individuals with similar training in the interpretation of aerial photographs examine the images and draw boundaries among types that they identify, we should be able to assess differences in boundary identification and location for various pairs of types relative to their means and variances. From this information we should be able to relate boundary and category errors to the differences in the means and variances of each type.

ALTERNATIVE MAP REPRESENTATION FOR THE INTEGRATION OF THEMATIC MAP AND REMOTELY-SENSED DATA

Although not readily apparent, the topics discussed in the previous section provide a conceptual methodology for linking information from satellite images and thematic maps through a common format. In this, the information contained in thematic maps which are stored in a GIS may be considered “vector information” and information stored in satellite images as “raster information.” For the purposes of this section, the fashion in which the data is stored internally in the computers is not the issue. That is, it is not our intent to discuss computer algorithms to smooth the “stairstepping” which can occur when raster information is converted to a vector representation. Rather it is our desire to focus on the differences between data which effectively represent a sample grid of points over a region (a “raster”) and that which draws lines around areas which are considered to be homogeneous at a given map scale (“vector polygons”).

Currently, in the analysis of remotely-sensed images, considerable activity is devoted to classification procedures with the goal being to produce “better” thematic maps. Although the answer might appear to be obvious, one can ask the question “Why do we strive to do this?” One reason is that when the first remotely-sensed data became available in the early 1970s, it was “understandable” that initial research efforts would be devoted to producing maps that humans could understand. hus a somewhat implicit goal of image classification was/is to produce the same type of map which is produced from human interpretation of a geographical landscape. As such, the classified image has often been viewed as the final result of a particular analysis. As GIS as a discipline has developed, however, a second reason to produce thematic maps from remotely-sensed data became evident. Because the value of a link between GIS and remote sensing was (and continues to be) recognized, a “common format” was required. Since the goal of
many image scientists already was to produce thematic maps (i.e., classified images), it may have seemed “natural” to force the satellite product into a form that would fit into a GIS -- i.e., a thematic map.

The problem with this latter activity is that by employing only a classified satellite image in a GIS, one has discarded considerable amounts of information. While the goal of a common form of representation remains necessary for suitable data integration, forcing satellite imagery to conform to highly simplified categorical thematic maps means that, potentially, one is operating at an inappropriate level of real-world categorization to perform optimal analysis. If one can move to a less simplified level of data representation, subsequent spatial analysis may prove to be more useful. The most immediate problem is to find a way to be able to represent “GIS thematic map information” at a less simplified, categorized level. That is, one would like to quantify and utilize the raw information that went into making the map.

This is where the error analysis that was discussed in the previous section may be useful. As envisioned, the establishment of a link between boundary and category error should allow one to produce a series of “error” or “certainty” surfaces for a map. One such surface would be required for each class. The surface for a given class would show the certainty or “likelihood” of finding that class at a given location, rather than merely showing a deterministic map having fixed boundaries for all classes.

Such a representation would also be consistent with the classification of remotely-sensed images. All classification techniques are probability-based and the final thematic map that is produced is merely a categorization of probabilities. Usually this is done by assigning a pixel to its “most likely” class with the class to which a pixel is ultimately assigned is merely that class having the highest probability. This may be acceptable if one class’s probability is high relative to the probabilities for all other classes. But if the probabilities are nearly equal, and/or if all probabilities are relatively low, considerable information will be lost if a given pixel is merely classed as “Type B”.

This indicates the potential power of viewing the digital representation of maps as “certainty” surfaces. In deterministic remote sensing analysis, for example, one would classify a pixel as “Type B” if there are 10 possible classes, and the probability of Type B is 0.10 (10%) while the probability for the other 9 classes is equal at 0.09 (9%). (Type B is the “most likely” by virtue of having a higher probability than any of the other classes.) Intuitively, however, this is considerably different from a pixel being classified as Type B because it had a 0.99 (99%) chance of being Type B. Similar dynamics operate on thematic maps. The center of a shallow lake with a highly irregular shoreline has a much higher probability of being “Lake” than those locations at the edge. The work described previously concerning boundary and category errors may allow this and similar concepts to be quantified and incorporated into future spatial analysis. Such a step would provide a valid way for treating “raster” (remotely-sensed) data in the same fashion as “vector” (GIS) data.

**CREATION OF PROBABILISTIC SPATIAL OPERATORS**

Assuming that one can develop a methodology to create the “certainty” surfaces described and store them in a suitable data structure, it remains to understand how to analyze them. This is particularly difficult to ponder considering the way in which spatial analysis is conducted presently. Current techniques are largely deterministic and based on Boolean logic. A classic example of this type of spatial (and GIS) analysis is that of “factory location.” In such a scenario, one wants to determine the “optimal place” at which to locate a factory. In doing so, one must consider the stability of the soil, severity of the terrain, drainage pattern of the region, distance from population centers,... etc., etc. Thus one digitizes a map of each factor, overlays these and identifies in the final step those resulting polygons which have the desired soil-drainage-terrain—... etc. combination. Almost inevitably when this is actually done, those polygons having the desired characteristics are extremely dispersed over the surface and highly irregular in shape and size.

Intuitively, humans understand that one probably cannot actually go into the “real world” and locate the polygons and their exact boundaries. Nor is it the case that these polygons are truly the only “ideal” areas on the map. One must remember that maps of soil, drainage, etc. have variable levels of precision and thus map boundaries are not necessarily located or represented correctly. For example, the real world boundary
between two soil types is a “fuzzy” transition zone, not an abrupt boundary as it must be mapped with pen and paper technology. Thus, conceptually, a “more honest” result from the analysis cited would be a map showing the likelihood or “certainty” of having all (or some) of the desired characteristics at a given location.

Thus spatial operators are needed that somehow are based on “certainty” and give answers in terms of the likelihood of finding areas having the desired characteristics. One possibility for analyzing these surfaces would be to use the probabilities stored on the certainty surfaces as weights for each factor. For each location, one might simply sum (or average) the probabilities for each desired class to obtain a score which could be mapped for human interpretation (Figure 1). One potential problem with this is that such scores would mask the fact that the “sufficiency” for one factor might be 1.00, but for another 0.00. Thus an area with a relatively high average score could actually be completely unsuitable. However, techniques could be developed to remove from consideration those areas which have a certainty below a certain threshold for any specified factor. Another possibility for avoiding this problem would be to sum the probabilities only if the desired class on a particular feature map was the most likely.

A conceptually different methodology which would require considerably more rigor and quantitative expertise would be the incorporation of “fuzzy set theory” into GIS analysis. Fuzzy set theory is a mathematical discipline which attempts to account for the idea that not all objects can be classified into one,
and only one group. Thus it uses “fuzzy logic” instead of Boolean logic to reach a solution. Of interest here is that one may be able to use “fuzzy information” -- i.e., that having some uncertainty associated with it -- rather than deterministic data in trying to produce spatial results using fuzzy set theory. This manner of thinking is a significant departure from the way most spatial analyses have been viewed in natural resources. In forestry, for example, it is assumed that a variable of interest can be measured “exactly” except for sampling error (as measured by the statistical standard error). However, the information contained in the standard error is generally not used until “the end” of the analysis. Even at that point, its use is limited largely to the placing of confidence intervals around numerical point estimates. Conceptually, fuzzy set theory differs from this by trying to incorporate “uncertainty” or “error” at every step of analysis.

Regardless of the methodology which is developed ultimately, the next generation of spatial operators also requires an additional change in philosophy. Currently, spatial analysis is done largely by examining the factor at a given location in isolation of objects and factors in the surrounding area. However, the greatest value of maps (arguably) is their ability to show relative location. This unfortunate paradox is evident in most commercial GIS software packages. Though vector systems must internally maintain topology -- i.e., information about neighbouring polygons -- very few systems allow queries of “adjacency” or “touching.” (Fortunately, an increasing number of systems are beginning to be able to do this.) The problem stems partly from the fact that theoretical conceptual models for “neighbourhood analysis” are generally not very well-developed. That is, human understanding of how to define a neighbourhood and what to do with “neighbourhood information” is extremely limited. Yet it seems intuitive that the “value” at any given location will often be strongly related to the “values” at locations “around” it. Current spatial operators work largely using a “pin through the maps” philosophy and examine the characteristics of each location in isolation. Future spatial operators must incorporate the concept of looking across a map surface to evaluate spatial relations rather than down on a location in isolation.

DATA STRUCTURES

The ideas discussed previously may eventually be found to be extremely difficult to incorporate into existing raster or vector spatial data structures. Both of these structures have a fixed unit which is indivisible -- cells for rasters and polygons for vectors. Yet many of the ideas discussed herein require a data structure which has the potential to produce a different “value” for a factor if one moves an “infintessimal” distance across a surface. Furthermore, to be able to develop “across surface” spatial operators, a data structure is required for which an integral part of the data structure is information on the adjacency of objects.

One promising possibility is the Voronoi Diagram which is composed of Theissen polygons. (Terms which are often used instead of, or in relation to, Voronoi Diagrams are “Dirichlait Tesselations” and “Delaunay triangles.”) The boundaries of the Theissen polygons for a given map object are defined by the perpendicular bisectors of the object’s immediate neighbours (Figure 2). Although one generally constructs Theissen polygons from points, recent work by one of the authors (Gold 1990) has extended the idea to line segments. Each Theissen polygon contains only the area that is closer to a given object than to any other. In doing so, the Voronoi Diagram maintains information about the “neighbours” of every object on the map. Since this information is central to the Voronoi diagram as a data structure, the requirement that adjacency be known and available for all objects is satisfied.
Another requirement that Voronoi diagrams fulfill is that they have the potential to produce a unique value for each point on a map. Gold (1989) has used Voronoi diagrams with "area-stealing" techniques for successful interpolation of topographic surfaces. In "area stealing," one has an existing Voronoi diagram containing Thiessen polygons based on (in this case) points of elevation. To obtain the elevation for any given point on the map, one places a "query" point at the desired location and reconstructs the Voronoi diagram as if the point were a "real" point. One then examines the area of the new Thiessen polygon of this imaginary point relative to the Thiessen polygons on the previous "real" Voronoi diagram. An area-weighted average of the elevations of each Thiessen polygon from which area was "stolen" for the new Thiessen polygon is then calculated and assigned to the point. The imaginary query point is then removed so that subsequent interpolation will utilize only the original points of known elevation. It can be seen that if one changes the location of the query point even slightly, different areas would be "stolen" from polygons on the "real" Voronoi diagram thus producing a different elevation.

Though continuous (elevation) data were used as an example in the preceding paragraph, a similar concept might be developed for categorical thematic data in order to integrate the topics discussed. One might be able to digitize a thematic map and construct its associated Voronoi diagram. Area stealing techniques might then be used to produce the error surfaces described using the area of each stolen polygon as the certainty. For example, if the insertion of a given point created a polygon which "stole" 60% of its area from a "Type A" Thiessen polygon and 40% from a "Type B" Thiessen polygon, the certainty of that point being Type A would be 60% with a 40% likelihood of being Type B.

It is stressed here that such an operation would force one to make a large number of implicit assumptions whose implications have yet to be studied. For example, consider a case where a clearcut is next to a mature forest. Both types have relatively low variances and the boundary between them is fairly readily identifiable. In such a case, one should weight the certainty of a given point being one of the two types by the variance (and possibly mean) of each type. However, an appropriate measure and/or weighting formula for such a concept, and its associated underlying statistical properties have yet to be explored. Furthermore, while area-stealing techniques have proven their value for point data, they have yet to be extended to lines. As noted, the concepts of the Voronoi Diagram have been extended for lines which also allows for the creation of polygons. A similar extension has yet to occur for the area-stealing techniques. Indeed, it is not yet clear that this can be done. By definition, any imaginary "query point" inserted into a polygon-based -- as opposed to point-based -- Voronoi diagram will steal area only from a single polygon. Thus developing a certainty surface from a polygon-based Voronoi diagram may not be possible, although this is under study.

Nonetheless, the Voronoi diagram as a data structure appears to have the potential to integrate all of the concepts discussed. It has one additional potential advantage that will be mentioned only briefly in closing. Though Voronoi diagrams were initially conceived for point data only, one of the authors has extended these concepts for lines. Thus at its present stage of development, the Voronoi diagram can accommodate both point and line data. If a raster containing, for example, remotely-sensed data can be viewed as a point sample of a surface, it is evident that the Voronoi diagram can readily accommodate this information.
Similarly, if one has a thematic soils map with linear boundaries, one can also incorporate this into a Voronoi diagram. Thus the ability to handle both point and line data in the same fashion may provide additional benefits from the Voronoi diagram as a data structure.

CONCLUSIONS

As DSTs evolve and progress, four issues must be addressed. First, there must be better ways to quantify and utilize spatial errors. In particular, “real-world-to-map” simplifications must be identified, measured, and incorporated into spatial analysis. Second, the use of these errors should provide a natural way to integrate remotely-sensed and thematic map-based data. Third, the integration of the two data types into a single structure will require a new family of spatial operators for spatial analysis. Rather than operating on the classified thematic maps in the GIS, these operators should utilize the underlying error surfaces. They should also work “across” maps rather than employing a “pin through the map” concept. Finally, the embodiment of these concepts into usable software algorithms will probably require the development of a new spatial data structure. Such a structure will require that each point on the map has the potential to be assigned a different value, and that information about adjacent map objects be an integral part of the structure. The Voronoi diagram has already shown potential to meet both of these needs.

LITERATURE CITED


