Implications of error and uncertainty for an environmental planning scenario: A sensitivity analysis of GIS-based variables in a reserve design exercise

Charlene Rae, Kristina Rothley, Suzana Dragicevic

Abstract

The creation of a protected area network is used as an example to demonstrate the potential effects of uncertainty and error in geographic information systems (GIS) data on our ability to make reliable land planning decisions. A graph-based model of the landscape was employed where nodes represent habitat patches, edges represent inter-node paths, and inter-node distance is measured as the least-cost path according to a resistance surface. Optimal reserve networks were generated by considering the trade-offs between area and connectivity, as measured according to correlation length. A “base case” network was first created and then varied according to the land cover categorical definitions used to extract the patch map, the resistance surface, and the cell size used to represent the raster patch and resistance surfaces. The largest recommended network was over 101% larger than the smallest recommended network. The identity of the small patches predicted to be critical stepping stones and the projected effectiveness of the alternative network configurations was widely disparate. Our results support the need for a precautionary approach to compensate for data uncertainties. Through sensitivity analysis, alternative scenarios can be created for decision-makers that highlight the most conservative options and emphasize the potential uncertainties surrounding the outputs.

Keywords: GIS; Sensitivity analysis; Correlation length; Protected area planning; Connectivity; Reserve design

1. Introduction

Land planners and resource managers increasingly rely on spatially explicit information that is stored, analysed, and displayed with geographic information systems (GIS; Crosetto et al., 2002; Crosetto and Tarantola, 2001; Edwards and Lowell, 1996; Norton and Williams, 1992). Crosetto et al. (2002) note that an increasing amount of politically sensitive decisions are being made based upon the information derived from digital spatial models. Often these decision makers have limited experience with the technical issues related to the use of GIS and, in particular, are unaware of the uncertainty, when the difference between produced information and reality is unknown, or error, when there is a measurable known difference between the produced information and reality, associated with resultant information (Heuvelink, 2002; Hunter et al., 1995; Stoms et al., 1992; Zhang and Goodchild, 2002). While its visual output is compelling, GIS does not mitigate against uncertainty as new information can be produced without any measure of its reliability (Lanter and Veregin, 1992). Ultimately, uncertainty and error in GIS output could lead to the misinterpretation of results or to inappropriately high confidence levels for particular management actions (Foody and Atkinson, 2002). Poor land planning decisions related to environmental management and conservation can be exceptionally costly due to their irreversibility (Norton and Williams, 1992).

The primary introduction of uncertainty and error can occur as the result of measurement imprecision and inaccuracy during the initial collection of spatially explicit data. For example, one of the most common methods of creating digital spatial data is to “digitize” features, such as cadastral boundaries or contour lines, represented on paper maps by tracing their outline with a cursor encoding their coordinates. Errors may occur during the initial trace or during any operation which acts to join the coordinates or generalize the resultant features (Burrough and McDonnell,
Remotely sensed data, including air photos or satellite images, which are increasingly employed in land planning may also contain uncertainty and error related to the sensor (hardware) systems or with image pre-processing such geometric rectification, data conversion, classification, and data generalization (Lunetta et al., 1991).

Secondary addition of error and uncertainty enters during subsequent data processing. For example, when changing between vector and raster formats of spatial data, conversion quality and boundary representational accuracy depends highly on the cell size of the resulting digital raster map (Congalton, 1997). If the cell size used is too large, then features (especially narrow features such as roads or rivers) are “lost.” The querying of existing digital spatial data to create new data layers can be problematic if some features exhibit partial fuzzy membership to one or more categories (Morris, 2003). Other processing errors may include querying information from data layers and buffering features (Congalton, 1997; Morris, 2003; Veregin, 1989).

In this study, we use an example environmental planning scenario to demonstrate the potential effects of uncertainty and error on our ability to make reliable land planning decisions. In particular, we focus on the secondary addition of error and uncertainty, where pre-existing spatially explicit is geoprocessed to provide information to support planning. Our example scenario (Rothley and Rae, 2005) is the creation of a protected area network, which is a currently popular strategy undertaken by land planners to protect natural systems in a rapidly expanding human dominated landscape (Moilanen (2005) and Rothley (1999) have more discussion on reserve design strategies). We generated trade-offs between connectedness of protected area network alternatives and total network size and used these curves to decide how big and where the network should be. Our strategy was to first generate these trade-off curves according to a specific set of technical decisions and assumptions. These curves served as our independent measure of truth (Stoms et al., 1992) or “base case” which, for purposes of this exercise, we assume to be the error-free case. We then vary our geoprocessing strategy, including changing the ways in which queries were performed or data are transformed, to introduce error/uncertainty into the decision making process, recreated these curves, and observed the effects on the management recommendations had our analysis been limited to the base case. Our example protected area planning scenario involves the application of graph-based metrics for landscape analysis. Our methods can be generally applied to any GIS-based predictive modelling that employs spatial data.

2. Methods

2.1. The example study system

The Resort Municipality of Whistler in British Columbia, Canada is an approximately 13,000 ha mountainous region comprised of the coastal western hemlock (62%), mountain hemlock (19%), alpine tundra (15%), and Engelmann spruce sub-alpine fir (4%) biogeoclimatic zones (Fig. 1). The municipality is bounded on the east and west by the Coast Mountains and bisected in a southwest to northeast direction by a highway and by a valley containing a series of lakes connected by creeks. The population of the Resort Municipality of Whistler has increased dramatically from 531 in 1976 to 8896 in 2001 (Statistics Canada, 2001) and the landscape has become increasingly fragmented as a result of logging, the introduction trails for an internationally-known skiing resort, and development to support a high volume of summer and winter tourism.

In 2002, The Resort Municipality of Whistler assembled a steering committee of stakeholders to design a municipal-wide network of protected areas. As part of the discussions for this land planning exercise, it was decided that second-growth forest patches (containing 21–140-year-old trees) would be included in the protected area network if they contributed to connectivity. According to the British Columbia Ministry of Forests
Biodiversity Guidebook (British Columbia Ministry of Forests, 1995), the coastal western hemlock biogeoclimatic zone is categorized as “natural disturbance type 2” for which the frequency of cross-elevational connectivity is high for natural ecosystems. Therefore, we wanted to know how the cross-elevational connectivity provided by alternative protected area networks varied as a function of the total amount of second-growth forest patches included in the network. Two GIS-based datasets were available to explore this question. The first was a polygon map of land cover for the entire municipality (smallest mapped polygon 0.05 ha; B.A. Blackwell and Associates Ltd., 2004). Each polygon could bound an area containing as many as three different land cover types so, for example, one polygon might contain 80% second-growth forest, 10% old-growth forest, and 10% rock. The second dataset was a 20 m elevation contour map (from 1:20,000 mapsheets; Government of British Columbia, 1996). All inputs used in our analysis were derived from these two vector-based datasets.

2.2. Graph-based connectivity metrics

We used connectivity metrics calculated from a graph model of the Whistler landscape to describe the connectivity of alternative protected area networks (Bunn et al., 2000; Cantwell and Forman, 1993; Keitt et al., 1997; Urban and Keitt, 2001). Graph-based landscape models have recently become a popular tool for evaluating connectivity because their nodes and edges are compelling analogues of the patches and corridors that comprise real landscapes (Urban and Keitt, 2001). Potential barriers to dispersal in the matrix surrounding habitat patches, such as roads, can be incorporated into the calculations. It is also possible to adjust the models according to the dispersal capabilities of particular focal species. This is important because a landscape’s connectivity is a species-specific quality that depends on the organism or process in question and its ability to cross-barriers, perceptivity, and risk averseness in non-habitat (With et al., 1997).

A graph model of a landscape consists of “nodes” at the centroid of each habitat patch (in our case, these were the second-growth forest patches) and “edges” representing the path between each pair of nodes (Fig. 2). A raster “resistance surface”, where each cell’s value is the relative cost for an organism to move across the cell, is used to calculate the least-cost path between the closest edges of each patches pair. The length of this path, measured in cost units, represents the distance between each patch pair (the “edge length”). A metric of the landscape’s connectivity is then computed according to this model (see Rothley and Rae, 2005; Bender and Fahrig, 2005; Winfree et al., 2005, for discussion on efficacy of connectivity metrics to describe animal movement). The metric we used, called “correlation length” (Keitt et al., 1997), is calculated in four steps. Starting with the full graph model of including the complete set of nodes and edges, all edges with lengths longer than the maximum dispersal distance, d (measured in cost units), for some particular organism are removed. The second step is to define the “clusters” within the graph, which are the groups of nodes joined by edges less than d. The third step is to calculate the radius of gyration for each cluster, which is a measure of cluster size for irregular, sinuous shapes (Keitt et al., 1997) as follows:

\[ R_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \sqrt{(x_j - \bar{x_i})^2 + (y_j - \bar{y}_i)^2}, \]  

(1)

where \( R_i \) is the radius of gyration of cluster i, \( \bar{x}_i \) and \( \bar{y}_i \) the mean x and y coordinates of the cells in cluster i, \( x_j \) and \( y_j \) the coordinates of the \( j \)th cell in cluster i, and \( n_i \) is the total number of cells in the cluster. Finally, the correlation length (\( C_d \)) of the node and edge set is calculated as:

\[ C_d = \frac{\sum_{i=1}^{m} n_i R_i}{\sum_{i=1}^{m} n_i}, \]  

(2)

where \( n_i \) is the number of clusters in the landscape (see Rothley and Rae, 2005).

The importance (\( I_{dk} \)) with respect to connectivity of each patch k in the landscape can also be calculated as:

\[ I_{dk} = \frac{C_d - C_{dk}}{C_d}, \]  

(3)

where \( C_{dk} \) is the new correlation length when all edges to the patch k’s node have been removed. Patches with high importance values are considered critical stepping stones, since removal of these patches results in a relatively high connectivity loss for the landscape. We used the patch importance calculation to generate the trade-off curves been landscape connectivity and total protected area network area (as described in the following section).

2.3. Analysis

Our strategy to explore the effects of uncertainty and error on the protected area planning exercise was to perform our GIS-based analysis according to a particular set of technical decisions and assumptions (referred to as the “base case”), and then recreate this analysis based on three variations in these decisions and assumptions. The variations were adjustments in the way that the two input datasets, the land cover map and the contour map, were processed and were chosen to specifically investigate the effects on our results of the following: (1) the land cover categorical definitions used to extract the patch map, (2) the cell size used to represent the raster patch and resistance surfaces, and (3) errors in the resistance surface. All pre-analysis of the spatial data (queries, vector to raster conversions, assembly of the resistance surface) was performed using ArcView 3.2 GIS software.
For the base case (Fig. 3), the second-growth patch map was created by querying the land cover map for all polygons that contained any second-growth forest (i.e., second-growth forest could be the largest, second largest or smallest component of the land cover bounded by the polygon). Ten high elevation “target” patch polygons were added evenly across the ridgelines that bound Whistler to anchor the origin and destination points for organisms conducting cross-elevation movement (Malcolm and ReVelle, 2002).

All polygons were then converted to raster GRID format with 50 m cell resolution and then exported to the ASCII format used by SELES. To create the resistance surface, we assumed that the cost incurred by an organism moving over any given location was a function of the slope at that location. Ten high elevation “target” patch polygons were added evenly across the ridgelines that bound Whistler to anchor the origin and destination points for organisms conducting cross-elevation movement (Malcolm and ReVelle, 2002).

All polygons were then converted to raster GRID format with 50 m cell resolution and then exported to the ASCII format used by SELES. To create the resistance surface, we assumed that the cost incurred by an organism moving over any given location was a function of the slope at that location. Ten high elevation “target” patch polygons were added evenly across the ridgelines that bound Whistler to anchor the origin and destination points for organisms conducting cross-elevation movement (Malcolm and ReVelle, 2002).

For the base case (Fig. 3), the second-growth patch map was created by querying the land cover map for all polygons that contained any second-growth forest (i.e., second-growth forest could be the largest, second largest or smallest component of the land cover bounded by the polygon). Ten high elevation “target” patch polygons were added evenly across the ridgelines that bound Whistler to anchor the origin and destination points for organisms conducting cross-elevation movement (Malcolm and ReVelle, 2002). All polygons were then converted to raster GRID format with 50 m cell resolution and then exported to the ASCII format used by SELES. To create the resistance surface, we assumed that the cost incurred by an organism moving over any given location was a function of the slope at that location. Ten high elevation “target” patch polygons were added evenly across the ridgelines that bound Whistler to anchor the origin and destination points for organisms conducting cross-elevation movement (Malcolm and ReVelle, 2002).

For the base case, we calculated the correlation length of the complete landscape where all second growth patches were included in the landscape, varying \( d \) by 500 cost unit increments between 0 and 38,500 (the \( d \) at which all patches were joined in a single cluster). The importance index \( I \) for each patch was also calculated for each value of \( d \). The patch node, excluding the 10 target patch nodes which were added to anchor the origin and destination points for organisms conducting cross-elevation movement, with the minimax \( I_{dk} \) (the smallest \( I_{dk} \) of the set of maximum \( I_{dk} \) for each \( d \)) was removed (the edges to the node were dropped). This removed patch is least likely to be a stepping stone for any \( d \). Correlation length, \( C_d \), and \( I_{dk} \) were then recalculated for the resultant patch node set for each \( d \), and again the patch with the minimax \( I_{dk} \) was removed. This “pruning strategy” was repeated until only a single second growth patch node and the 10 added target patch nodes remained. The result of this analysis is a set of curves, one for each value of \( d \) between 0 and 38,500 cost units, showing the trade-off between connectivity \( C_d \) and protected area network area. For clarity, we only report the results for five \( d \) values: the “best disperser” with \( d = 38,500 \) cost units, an “intermediate disperser a” with \( d = 20,000 \) cost units, an “intermediate disperser b” with \( d = 10,000 \) cost units, an “intermediate disperser c” with \( d = 6000 \) cost units, and the “poor disperser” with \( d = 1000 \) cost units. If for a given reserve design scenario there was a particular focal species of importance, then consideration could be restricted to the results derived from parameters consistent with the behaviour and life history of that species.

Our first variation on the base case was to change the land cover categorial definition used to extract the patch map. For the base analysis, second-growth patches were defined as any polygon that contained any second-growth forest. We then recreated the patch map twice: (1) only selecting polygons where second-growth forest was either the most prevalent or second-most prevalent land cover type in the polygon, and (2) only selecting polygons where second-growth forest was the most prevalent land cover type in the polygon. The identical pruning strategy was then repeated starting with the complete landscape where all second growth patches were included in the landscape and iteratively removing patches until only a single second growth patch node and the 10 target patch nodes remained.

Our second variation on the base case was to change the cell size used to represent the raster patch and resistance surfaces. The original patch query (any polygon that contained any second-growth forest) was reinstated, but this time the resultant polygons were converted to raster GRID maps with cell sizes of 10, 20, 30, and 40 m. Four new resistance maps were also created according to the same steps as for the base case but with these new cell sizes. Each map pair, e.g., the 10 m patch map and the 10 m resistance map, was then processed using the same pruning strategy. Note that for this analysis we are considering the original vector polygons from which the patch and resistance grids are derived to be error-free, and that error level is
then determined by how precisely the raster representation of the polygons corresponds to the vector representation.

Our final variation on the base case was to introduce errors to the resistance surface. In the base case, highways, urban areas, and water surfaces were all assumed to very costly, and cells containing these land cover types were given the maximum resistance score. We also filled the cracks in the highway to prevent erroneous least-cost path calculations. This time we pretended to “forget” some or all of these adjustments. Five new resistance surfaces with 50 m cell size were created as follows: (1) same as base case except do not consider water as high cost, (2) same as base case but do not consider water or development as high cost, (3) same as base case but do not consider water or development as high cost and do not fill in the highway cracks, (4) only consider slope, and (5) assume the resistance in all cells is equal (so that the least cost path between the nodes is the straight line Euclidean distance). Each new resistance map was processed with the original patch map using the same pruning strategy.

3. Results

The characteristics of the patch map varied depending on the query used to extract the second-growth polygons and cell size. Total patch number ranged from 47 to 57 (Table 1). Total patch area ranged from 1153 to 1242 ha. The maximum least-cost path varied between 103,724 and 182,981 cost units.

For the base case, connectivity decreased as total protected area was reduced for all five disperser types (Fig. 4A). For the best disperser this relationship was relatively linear between 1236 and 380 ha. For intermediate dispersers a, b, and c there is a sharp drop in connectivity at 790 ha (identified visually relative to the general slope of the trade-off curve for larger protected area networks). This drop or “transition zone” represents the removal of a stepping stone patch. Our recommended protected area network, then, according to this round of analysis would be the configuration corresponding to 796 ha (Table 1; for each round the recommended network is that with area just larger than the first transition zone area for any disperser). This is represents 64.4% of the total second-growth forest area.

The trade-off curves were nearly identical for the first variation on the base case when we changed the definition by which second-growth polygons were extracted from the land cover map. For example, the trade-off curves between connectivity and area when we only selected polygons where second-growth forest was the most or second-most prevalent land cover type show the same linear relationship between connectivity and area for the best disperser and the loss of a stepping stone patch between 756 and 750 ha for intermediate dispersers a, b, and c (Fig. 4B; the results depicted in this figure are characteristic for all those obtained in the first variation on the base case). When second-growth patches were defined as polygons where second-growth forest is the most prevalent land cover type, this stepping stone patch is lost between 828 and 820 ha. The recommended protected area network, then, according to this round of analysis would be the configuration corresponding to 756 ha for the first new definition of second-growth patches and 828 ha for the second new definition of second-growth patches (Table 1). These configurations represent 65.6% and 67.7% of the total second-growth forest, respectively, which is similar to the recommendation derived from the base case.

In contrast, the trade-off curves changed considerably for the second variation on the base case when we the changed cell size for the raster maps. For example, when the 30 m cell size was used an important stepping stone patch is lost for intermediate disperser a between protected area network configurations corresponding to 1115 and 1110 ha (Fig. 4C; these curves were chosen for display because they are most dissimilar to the base case). When the 20 m cell size was used, intermediate disperser c does not lose a critical stepping stone until the total area of the network is reduced to 640 ha (in contrast to the base case where this stepping stone is lost at 790 ha). According to this round

<table>
<thead>
<tr>
<th>Version</th>
<th>Number of patches</th>
<th>Total patch area (ha)</th>
<th>Maximum edge length (cost units)</th>
<th>Recommended management:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reserve area (ha)</td>
</tr>
<tr>
<td>Baseline</td>
<td>56</td>
<td>1236</td>
<td>182981</td>
<td>796</td>
</tr>
<tr>
<td>Variations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patch query</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>52</td>
<td>1153</td>
<td>182981</td>
<td>756</td>
</tr>
<tr>
<td>2</td>
<td>57</td>
<td>1224</td>
<td>182981</td>
<td>828</td>
</tr>
<tr>
<td>Cell size (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>47</td>
<td>1240</td>
<td>182620</td>
<td>797</td>
</tr>
<tr>
<td>20</td>
<td>49</td>
<td>1236</td>
<td>179111</td>
<td>885</td>
</tr>
<tr>
<td>30</td>
<td>53</td>
<td>1242</td>
<td>183620</td>
<td>1115</td>
</tr>
<tr>
<td>40</td>
<td>53</td>
<td>1234</td>
<td>175479</td>
<td>694</td>
</tr>
<tr>
<td>Resistance surface</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>56</td>
<td>1236</td>
<td>103724</td>
<td>800</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>1236</td>
<td>106736</td>
<td>554</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>1236</td>
<td>106951</td>
<td>860</td>
</tr>
<tr>
<td>4</td>
<td>56</td>
<td>1236</td>
<td>114456</td>
<td>714</td>
</tr>
<tr>
<td>5</td>
<td>56</td>
<td>1236</td>
<td>131288</td>
<td>688</td>
</tr>
</tbody>
</table>
Fig. 4. Trade-offs between landscape connectivity, as measured by correlation length, and total area of second-growth patches for organisms with varying dispersal capabilities (♦ poor; — intermediate “c”; — intermediate “b”; — intermediate “a”; — best). (A) Generated by removing patches one at a time in reverse order of their importance as stepping stones for cross-elevation connectivity (least important patches were removed first). (B) Second-growth patches were defined as polygons where second-growth forest was the most or second-most prevalent land cover type. (C) Definition of the second-growth patches and resistance surface were identical to the base case but the cell size for the raster maps was 30 m instead of 50 m. (D) The definition of the second-growth patches was identical to the base case but the resistance surface was based on slope only.

of analysis, the total recommended protected area network size varies from 1115 ha if the 30 m cell size is used to 695 ha if the 40 m cell size is used (Table 1). These configurations represent 89.7% and 56.3% of the total second-growth forest, respectively.

The largest change in the trade-off curves occurred when we introduced error into the resistance surface. When only slope is considered, there are no critical stepping stones for either the best disperser or the intermediate a disperser (Fig. 4D; these curves were chosen for display because they are most dissimilar to the base case). Intermediate dispersers b and c do not lose a critical stepping stone until the protected area network is reduced from 580 to 570 ha. When a uniform resistance surface is used, there are no critical stepping stones for either the best disperser or the intermediate a disperser and the intermediate b disperser loses a critical stepping stone between 590 and 580 ha and the intermediate c disperser loses two critical stepping stones between 780 and 700 ha. The total amount of second-growth forest recommended for the protected area network for this round of analysis ranged between 554 and 860 ha (Table 1).

4. Discussion

A land planner charged with the design of a protected area network would like to know the answers to three questions: how big should it be, where should it be, and how well is it predicted to perform. Given the potentially high monetary value associated with competing land uses, the planner would also like to know how reliably these questions can be answered so that decisions can be made defensibly. Error and uncertainty in the spatial data used to address these questions could translate into inefficient, unreliable, or erroneous land management beyond what might be expected from the complications of dealing with complex, real-world ecosystems.

The error and uncertainty we introduced into our spatial data in the form of variations in the technical decisions and assumptions for our GIS-based analysis had substantial effects on the recommended protected area network alternatives that would be offered to decision-makers. The largest recommended network was over 101% larger than the size of the smallest recommended network (Table 1). This represents an enormous change in the amount of land that would need to be committed to the network and, most likely, a sizeable increase in the funds that would need to be expended. There was some spatial overlap in the alternatives and, in general, small patches were predicted to be less important for maintaining the landscape’s connectivity because they were removed earlier during the pruning. But for each round of analysis, the identity of the small patches that were predicted to be critical stepping stones was highly varied. The loss of just one of these patches could translate to a shift in the landscape’s character from being highly connected to being highly discon-
nected for the organisms that used that patch. The capacity of the alternative network configurations to maintain connectivity for organisms with varying dispersal capabilities, as predicted by the trade-off curves, was also widely disparate (Fig. 4A–D). According to the base case (Fig. 4A), a reserve network as small as 796 ha would support a passable level of connectivity for all organisms and would not exclude any critical stepping stones. But according to the trade-off curves generated with 40 m raster cells (Fig. 4C), the intermediate c disperser will experience a sharp decline in landscape connectivity for any reserve network smaller than 1115 ha.

According to our analysis, the cell size used when the vector maps were translated to raster maps had the largest effect on the results. It could be argued that our form of connectivity analysis was particularly sensitive to raster resolution because of the dependence of the least-cost inter-patch distance calculations on the presence or absence of costly linear barriers to movement (Adriaensen et al., 2003). But the profound importance of cell size is likely to hold true any time vector data is converted to raster. Square-pixels can never perfectly represent the curved boundaries of polygons. During the translation from vector to raster, a considerable amount of information can be lost (Burrough and McDonnell, 1998). This information loss represents error that can only be identified through a direct comparison between the raster and vector data sets, and its effect on subsequent calculations can only be detected through a sensitivity analysis such as that performed here. One strategy to mitigate these errors is to always perform spatial analysis at the finest resolution possible (Stokes and Morrison, 2003). The drawbacks to this strategy are that the required computing power increases with decreasing cell size, and that it may be difficult to judge how fine is fine enough (De Genst et al., 2001). For example, the life history traits of different focal species may require analyses at varying resolutions (Araujo, 2004; Aspinall, 2001).

While there were no consistent trends between the level that scenarios were “pushed” away from the base case and the characteristics of the implied best reserve networks, there are three useful products from this type of analysis for land planners. First, the sensitivity of the recommended reserve network to data uncertainties is readily apparent. This justifies the time and expense devoted to an extended analysis period before decisions are made. Second, the relative importance of uncertainty and error sources can be ranked, allowing scarce resources that could be directed to reduce these sources to be allocated most efficiently. In our scenario, raster cell size was the most important source of uncertainty and, therefore, the obvious target for further analysis. Third, the range of results generated by uncertainty analysis forms the bounds for what could be expected in terms of management options and outcomes. For example, as cell size was varied, the recommended reserve network area ranged between 694 and 1115 ha. The relationship between cell size and reserve size was discontinuous because the effect of narrow costly landscape features became exaggerated with increasing cell size, but abruptly disappeared when the features were too fine relative to the cell size to be represented. But planners would now have a specific range of network sizes that should be considered and an understanding of their most and least conservative options.

No matter how visually convincing digital spatial data can be, it should always be noted that these data are in fact a model of the real world. A precautionary approach should be adopted to compensate for data uncertainties. Through sensitivity analysis, alternative scenarios can be created and the most conservative options can be highlighted for decision-makers with emphasis on the potential uncertainties surrounding the outputs. An important challenge to this approach is finding the balance between transparency and information overload, which could confuse decision-makers and substantially lengthen deliberation times.

Stoms et al. (1992) note that it is important not to tend towards the extremes of managing uncertainty, which are to disregard it or to look so critically at outputs that information is discarded due to potential uncertainties. Uncertainty will always be present in spatial modelling despite our best efforts to eliminate it. It is therefore important to acknowledge, study, and effectively communicate issues of uncertainty surrounding any output results of spatial modelling to the decision maker. More informed decisions can be made, by helping those with decision-making power to understand the potential uncertainty in the information they are being provided.

Finally, our analysis dealt specifically with the use of ecological data for land planning. While there has been an increasing interest in error and uncertainty analysis within the GIS literature, little relatively research has been conducted on the impacts of error and uncertainty on decision-making in ecology (Hunter et al., 1995; Ruckelshaus et al., 1997; Stine and Hunsaker, 2001; Verbeylen et al., 2003). Indeed, GIS has been applied to many diverse ecological problems such as habitat modelling, home range analysis, and reserve design (Pereira and Itami, 1991; Schadt et al., 2002; Selkirk and Bishop, 2002). As for our case study, much GIS-based research is used to support policy development for the management of both habitat and species (McComb et al., 2002; Woolf et al., 2002). Further research on the impacts of error and uncertainty on decision-making in ecology is warranted, as GIS is well suited to deal with inherently spatial ecological datasets.

Acknowledgements

We would like to thank the Resort Municipality of Whistler and Andrew Fall. This study was partially funded with NSERC Discovery Grants awarded to the second and third authors.

References


