INT. J. GEOGRAPHICAL INFORMATION SCIENCE, 2001, VOL. 15, NO. 2, 101–127

Research Article

Map-centred exploratory approach to multiple criteria spatial decision making

PIOTR JANKOWSKI*, NATALIA ANDRIENKO** and GENNADY ANDRIENKO**

*Department of Geography, University of Idaho, Moscow, Idaho 83843
**GMD—German National Research Center for Information Technology, Schloss Birlinghoven, Sankt-Augustin, D-53754 Germany; e-mail: piotrj@uidaho.edu

(*Received 10 December 1999; accepted 4 May 2000*)

Abstract. Spatial decision support is one of the central functions ascribed to Geographical Information Systems (GIS). One of the foci of developing decision support capabilities of GIS has been the integration of maps with multiple criteria decision models. Progress in this area has been slow due to a limited role played by maps as decision support tools. In this paper we present new prototype spatial decision support tools emphasising the role of maps as a source of structure in multiple criteria spatial decision problems. In these tools the role of map goes beyond the mere display of geographic decision space and multicriterion evaluation results. Maps becomes a 'visual index' through which the user orders decision options, assigns priorities to decision criteria, and augments the criterion outcome space by map-derived heuristic knowledge. As the additional means of structuring multicriterion spatial decision problems we present an experimental use of data mining, integrated with dynamic maps and multiple criteria decision models, in order to reduce a problem's dimensionality. We conclude the paper with future research directions emphasising map-based support for group decision making.

1. Introduction

Spatial decision making is an everyday activity, common to individuals and groups. People take into account the realities of spatial organization when selecting a locale to live, choosing land development strategy, allocating resources, or managing infrastructure. Most of individual spatial decisions are made ad hoc, without any formal analysis. These decisions are often based on heuristics and internalized preferences supported by mental maps and 'mental geocoding' of information about decision options. This analytically simple approach to spatial decision making can be explained by a relatively small 'decision equity' at stake in daily pursuits, such as the selection of a place to shop or an entertainment venue. The cost of making a poor choice (decision) can be a smaller selection of goods, higher prices than elsewhere or a boring evening spent at a nightclub. In contrast with these everyday decision situations faced by individuals, the decision equity of organizations is often quite

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online © 2001 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/13658810010005525

high. The organizations are hence more likely to use more sophisticated analytical approaches to support spatial decision making.

One such approach relying on Multiple Criteria Decision Models (MCDM) has been the active area of research and applications in spatial decision analysis (Malczewski 1999a). The objective of using MCDM is to help find solutions to decision problems characterized by multiple choice alternatives, which can be evaluated by means of performance characteristics called decision criteria. Because much of modelling done with Geographic Information Systems (GIS) deals with evaluating locational choice alternatives on the basis of suitability criteria considerable attention in the last decade was devoted to integrating MCDM with GIS software. These efforts resulted in a number of prototypes (Janssen and Rietveld 1990, Carver 1991, Church *et al.* 1992, Faber *et al.* 1995, Jankowski and Ewart 1996, Lotov *et al.* 1997, Wu 1998) and a few commercial software packages (Eastman *et al.* 1995, GeoChoice Inc. 1999). The state of the art in this area has been based on the *tight* integration strategy (Jankowski 1995). Depending on specific implementation solutions, either MCDM tools are implemented within a GIS package (for example, IDRISI, Eastman 1997) or GIS functions (such as spatial data query and mapping) are implemented in MCDM software (Fisher *et al.* 1996). The third option is to integrate both modules (GIS and MCDM) at the operating system level. For example, Jankowski *et al.* (1997) describe the use of the Dynamic Data Exchange (DDE) protocol to integrate a GIS-based map visualization module with MCDM and group voting modules.

The mapping and visualization capabilities of current GIS-MCDM software range from maps representing solutions obtained with simple non-compensatory decision rules to graphical representations of criterion and decision spaces integrated with more sophisticated compensatory decision techniques (Malczewski 1999a). Still, the current software tools provide only a limited map-based decision support for problem structuring. Although some innovative decision support maps were proposed for multiple criteria-based location/allocation analysis (Armstrong *et al.* 1992, Malczewski 1999b), maps have been used predominantly as presentation media either to display the results of spatial decision analysis or to inform about the location of decision options. The use of maps as analytical tools in spatial decision analysis has been little explored.

In order to test the use of maps as analytical decision tools Jankowski and Nyerges (in press) conducted an experimental study of human-computer-human interaction with 109 volunteer participants formed into 22 groups, each group representing multiple (organizational) stakeholder perspectives. The experiment involved the use of GIS integrated with multiple criteria decision models to support group-based decision making concerned with the selection of habitat restoration sites in the Duwamish Waterway of Seattle, Washington. Maps used by groups included the variety of general thematic maps and images representing situational/contextual characteristics of restoration sites and some specialized maps depicting the solutions of multiple criteria decision models and the measure of group consensus regarding the selection of restoration sites. These specialized maps are described in Jankowski *et al.* (1997).

Findings of the study (Jankowski and Nyerges in press) showed that maps played only a limited support role in various decision stages of the experiment. While the use of multiple criteria decision models by groups remained steady throughout different phases of the decision process, the use of maps was much lower during the initial (deliberative-structuring) phase, than during the later (analytical) phase. The participants used maps predominantly to visualize the evaluation results and much less to structure/design the decision problem.

In a different application of combined MCDM-GIS tools supporting the design of water quality improvement strategies the use of maps and decision models proceeded in a cyclic way (Lotov *et al.* 1999). Presentation of MCDM results on GIS-generated maps prompted water management experts to recurrently reformulate the decision problem, more specifically to impose additional constraints in order to account for previously overlooked aspects of the decision situation. Although maps seemed to play a more important role than in the Duwamish experiment, it is questionable whether they might be regarded as problem structuring tools. For experts, a map was mainly a convenient tool used to detect discrepancies between the results obtained from a model and experts' expectations based on some implicit preferences.

The question arises then: what are the effective means of using maps in order to support decision problem exploration and structuring? Casner (1991) clearly demonstrated that different graphical presentations of the same data are required to support different information-seeking tasks. Accordingly, the task of multiple criteria spatial decision analysis needs to be supported by task-specific map displays. What kind of maps are productive for spatial decision making and which direction should be taken in developing more effective tools for multiple criteria spatial decision analysis are still open research questions. In this paper we address them by demonstrating new ideas for integrating maps with MCDM using highly interactive, exploratory map displays coupled with multiple criteria decision models and data mining algorithms.

To ground the presentation of theoretical ideas in a substantive domain we present prototype software called DECADE (**D**ynamic, **E**xploratory **Ca**rtography for **De**cision Support) and demonstrate its application to a decision problem of funding allocation for primary health care services. More specifically, in §2 we outline research ideas concerning the use of maps in spatial decision support. In §3 we introduce new tools for map-based multiple criteria spatial decision analysis provided by DECADE. In order to present the tools in a realistic context, we use a real-world decision problem of ranking Idaho counties according to need for funding in primary health care services. We conclude the paper with a summary of suggested decision support maps and with future research prospects concerning the development of collaborative spatial decision support systems.

2. Maps as decision support tools in multiple criteria spatial decision analysis

The results of the Duwamish experiment showed that maps not only played a limited support role in the initial, exploratory stage of the decision process but also that there was a noticeable difference in the use of maps between facilitated (supported by a human facilitator) and non-facilitated group decision sessions. The facilitated sessions were characterized by a more frequent and intentional use of maps then non-facilitated sessions, where the use of maps was sporadic and haphazard. The role of facilitator was to guide the groups through the problem exploration and resolution and to assist in the use of maps and MCDM models. The outcomes of the facilitated sessions showed better quality and higher satisfaction with the decision making process, which is consistent with findings from other studies with group decision support systems (Chun and Park 1998). Since the presence of group facilitator was a source of structure simplifying and synchronizing the use of maps and decision models, we speculate that better integration of maps and MCDM tools through data visualization can also become a new structure improving the understanding of decision situations, and consequently leading to better outcomes of the decision making process. We propose that such integration be achieved through the interactive and dynamic visualization of both, criterion and decision spaces.

2.1. *Integrated visualization of criterion and decision spaces*

It is recognized that multiple criteria spatial decision making requires simultaneous visualization of criterion and decision spaces (Malczewski 1999b). This type of dualistic view opens up an opportunity to study basic relationships between the data (criterion outcomes) and their spatial patterns (arrangements of decision options in geographical space) providing the basis for understanding the structure of a decision problem. Malczewski (1999b) suggests that the main objective of using maps in multiple criteria spatial decision analysis should be the consideration of geographical locations in the process of exploring tradeoffs among the decision criteria and the search for best (compromise) solutions to the decision problem. Tradeoffs among two decision criteria can be depicted in a bi-criterion outcome space by means of a scatterplot so that each point on the plot represents the performance of a decision option on the respective two criteria. Locations of decision options along with the underlying spatial relationships constitute a geographic decision space, which can be depicted on a map. Malczewski (1999b) recognizes that the visualization techniques he suggested have limitations. These limitations are mostly caused by the static nature of the displays. We posit that the dynamic, highly interactive depiction of both criterion and decision spaces can be more productive for understanding the structure of a decision situation.

The idea of dynamic and interactive representation of criterion outcomes and spatial decision options comes from exploratory data analysis (EDA). The concept of EDA emerged in statistics. John Tukey (1977) introduced it as a counterbalance to the traditional statistical techniques of checking *a priori* selected hypotheses. The goal of EDA is to survey previously unknown data and to arrive at plausible hypotheses concerning relationships, patterns, and trends hidden inside the data volumes. Techniques of EDA are mostly based on data visualization, that is the graphical presentation of data that can help reveal important traits and relationships. The taxonomy of currently existing tools for interactive manipulation of graphical displays can be found in (Buja *et al.* 1996).

Over the past decade the notion of EDA has spread from statistics to cartography. Cartographers have recognized the demand for new software supporting the use of interactive, dynamically alterable thematic maps and facilitating 'visual thinking' about spatially referenced data (MacEachren 1994, MacEachren and Kraak 1997).

Particular attention in the area of data visualization and EDA is paid to dynamic linking between multiple data displays where changes in one display are immediately propagated to other displays (Buja *et al.* 1991). The most common method of linking is the identical marking of corresponding parts of multiple displays, for example, with the same colour or some other form of highlighting. Usually highlighting is applied to objects interactively selected by the user in one of the displays. This technique is known as 'brushing'. Linking of a map display with various kinds of statistical graphs was suggested as a primary instrument of the exploratory analysis of spatially referenced data (Monmonier 1989, MacDougall 1992). Later, other kinds of interactive cartographic displays were suggested in addition to linking and brushing. Thus, a user may interactively change visual properties of maps to make them more expressive (Dykes 1997, Unwin and Hofmann 1998, Andrienko and Andrienko 1999a).

One can expect that interactive, dynamically changeable, linked views can also be productive in multiple criteria spatial decision making. Interactive maps may provide the means of decision problem exploration and structuring and compensate for some weaknesses of traditional, static maps. For example, it may be revealing to locate on an interactive map any decision alternative selected in a scatterplot or to evaluate an alternative selected in a map by viewing its multicriterion characteristics on a scatterplot or a 'value path' plot.

2.2. *Capturing implicit geographical preferences of a decision maker*

Simultaneous representation of criterion and decision spaces and the support of interactive map exploration opens a possibility of eliciting a decision maker's preferences for decision criteria not only on the basis of attribute data but also geography. Thus, the decision maker (DM) can be given an opportunity to explicitly choose options considered as good candidates for the solution by selecting them directly on a map or in an accompanying scatterplot. In making the selection, the DM can account for spatial information gained from the map. This kind of input makes it possible, in particular, to derive DM's target values, or aspiration levels for the decision criteria. The aspiration levels thus obtained can be used for evaluating the available options and selecting the best of them, for example, by means of the nearest goal algorithm described in Lotfi et al. (1992).

It is quite clear which capabilities maps and complementary graphic displays should posses in order to make the above outlined approach to decision making possible: they should allow a direct selection of objects both through a map and through supplemental graphics representing criterion outcome space. The multicriterion decision procedures need to be integrated with the visualization. The DM should have the opportunity to view on the map the results of her/his tradeoff selection as well as to be able to change the selection and observe the impact of such a change on the evaluation results. It is desirable that switching between the analysis of criteria and options and the visualization of computation results be done in an easy, seamless way.

Another possible use of interactive maps is the encoding of geographical information in a form suitable for MCDM procedures. Through a direct interaction with a map display the DM can classify the options on the basis of their spatial positions. In this classification s/he can account for various aspects related to locations of options: a relative position with respect to other options (north-south, centre-periphery etc.), spatial relationships with other kinds of geographical objects (closeness to sea, mountains, roads etc.), or a geographical distribution of values of attributes referring to geographical objects (population density, unemployment, etc.). The classes may represent a combination of all relevant aspects. Alternatively, several classifications may be defined each reflecting one or a few of the spatial aspects significant for a decision problem. Then it is important that the DM arranges the classes thus obtained in the order of preference. A preference-based order makes possible treating classes as values of a new decision criterion, capturing the DM's heuristics and map-induced knowledge about spatial aspects of the decision space. Again, in the course of class ordering the DM expresses his/her implicit geographical preferences. For example, selecting a house to live, a person would assign the highest preference to the group of houses located close to a lake but having good public transport connections. The next category might be houses within the walking distance to the city centre, and so on.

2.3. *Reducing the cognitive complexity of multiple criteria spatial decision problem*

Cognitive complexity of multiple criteria spatial decision problem can be considered in both, the criterion outcome space and the geographical decision space. The number of decision (evaluation) criteria represents cognitive complexity in the criterion outcome space. In most of the compensatory MCDM techniques the decision maker is asked to explicitly express his preferences by assigning numeric weights to decision criteria. The higher the number of criteria, the more cognitively difficult it becomes to consistently assign weights that reflect the decision maker's perception of the relative importance of criteria. Assigning weights becomes even more confounded through the fact that weights account for changes in the range of values for each evaluation criterion and additionally they account for varying degrees of importance attached by decision makers to these ranges (Kirkwood 1997). Hence, some compensatory MCDM techniques use the implicit representation of preferences through criterion tradeoffs and aspiration levels (Lotfi *et al* 1992, Jankowski *et al.* 1999).

Another problem is the implicit assumption made in MCDM techniques that evaluation criteria are unrelated, although in practice this is not always the case. Consequently, it would be desirable if the decision maker knew about existing relationships among the criteria and accounted for them in assigning weights: for example, did not assign high weights to correlated criteria.

The reduction of the cognitive complexity of criterion outcome space and the revealing of relationships among the criteria can be achieved by means of standard statistical procedures (e.g. multiple regression, discriminant analysis). These procedures, however, require a user well-versed in statistics to interpret their results, hence they are not very attractive for a participatory spatial decision support system involving users with various levels of expertise in statistical techniques. Instead, we propose the use of more exploratory techniques from the domain of Knowledge Discovery in Databases (KDD).

KDD (Fayyad *et al.* 1996) is a research area of developing various computational methods, known as *data mining* techniques, for probing data with the aim to detect regularities, dependencies, or trends, and to draw generalized descriptions of data features and relationships. The methods of data mining can be grouped according to the purpose of their application into (Fayyad *et al.* 1996):

- Classification: mapping of a data item into one of several predefined categorical classes.
- Regression: mapping of a data item to a real-value prediction variable.
- Clustering: grouping of similar data items.
- Summarization: derivation of a compact description for a subset of data.
- Dependency modelling: detection of significant dependencies among variables.
- Link analysis: revealing of relationships between fields in the database, for example, frequent associations of particular values.
- Sequence analysis: modelling of sequential patterns, in particular, in data with time dependence.

Certainly, not all of the existing data mining methods can be recommended for non-expert users. An example of a data mining technique potentially useful, for a wider audience of users, is the C4.5 Classification Tree derivation algorithm (Quinlan 1993) belonging to the classification group. The C4.5 algorithm is one of the best known, efficient, and widely used techniques in KDD. Its results are easy to understand and interpret. The goal of the algorithm is to discriminate among the classes of objects and produce their collective descriptions on the basis of values of attributes associated with class members. In order to achieve this goal, the algorithm tries to select from the available attributes the most discriminating ones. In many cases a small subset of attributes is sufficient to 'explain' a given classification. These attributes may be regarded as the most important for characterizing the classes. The capability to select a smaller subset of representative attributes can be exploited for the purpose of complexity reduction.

In contrast with the complexity of criterion outcome space, complexity in the decision space is usually expressed by the number of feasible decision options. In the case of spatial decision making, spatial relationships and spatial patterns formed by locations of decision options produce additional complexity. Strict numerical representation of some of these relationships may be hard to achieve, and/or the resulting values may be difficult to interpret by decision makers. Yet the decision makers may be capable of eliciting patterns and relationships from maps and integrating them with their heuristic knowledge about the decision situation. Hence, it is desirable to provide the decision makers with explicit means to express heuristic knowledge and to incorporate it into multiple criteria spatial analysis, for example, as suggested in the previous subsection.

Cognitive complexity of the decision space can be reduced by applying the *Pareto-dominance* principle (Cohon 1978). A feasible decision option is called non-dominated, according to Pareto-dominance, if there is no other feasible option that surpasses this option on any of the criteria without reducing performance on another criterion. More formally, let X denote the set of all feasible decision options, and $C=\{c_1, c_2, \ldots, c_m\}$ is the set of evaluation criteria under consideration. Any element $x \in X$ is characterized by a vector of criterion outcomes $(y_1, y_2, \ldots, y_m)$, where y_i is a value of the criterion c_i, $i \in I=\{1, 2, \ldots, m\}$. Then we say that decision option $x' \in X$ dominates $x'' \in X$ if $y'_i \geqslant y''_i$ for all $i \in I$, and there is at least one $j \in I$ that $y'_j > y''_j$. Here, the symbols $\geqslant$ and $>$ are used to denote the relationship of preference: $\geqslant$ stands for 'no worse than' and $>$ for 'better than'. It is clear that in the Pareto sense option x' is better than x'' if x' dominates x''.

Application of the Pareto-dominance rule results in subdividing the set of decision options into dominated and non-dominated, hence simplifying the structure of the decision space. A decision support tool may display only non-dominated options and hide dominated ones. However, if the evaluation criteria do not represent the geographical aspects of the decision space, it is necessary to show the decision maker both non-dominated and dominated options; there may be dominated options that surpass any non-dominated one in respect to a geographical position. The decision maker can regard such options as good candidates for a solution. Therefore we suggest that all alternatives be shown to the decision maker, but the not-dominated options be graphically distinguished from the dominated ones, for example, by using larger symbols or a different hue value.

3. New tools for integrated map-MCDM spatial decision analysis

In this section we present the system DECADE which offers new prototype tools for multiple criteria spatial decision analysis implementing the ideas expressed in the previous section. DECADE has been developed on the basis of dynamic mapping software, Descartes (Andrienko and Andrienko 1999a). We also describe an experiment with the use of data mining software Kepler (Wrobel *et al.* 1996).

In order to cast the proposed tools in the context of a realistic decision situation, we first describe a decision scenario. We then show how a decision maker could analyse the decision situation using the facilities of DECADE. We describe three possible models of a decision maker's activity in the order of the increasing role played by a map. In the final part of the section we demonstrate the use of data mining in order to reduce the complexity of criterion outcome space.

3.1. *Decision scenario overview*

Idaho is a predominantly rural state with numerous geographic and access barriers. Sixteen of the state's 44 counties are considered frontier (less than six persons per square mile). Using the definition of urban to be a county with a population centre greater than 20 000 persons, there are only 7 urban counties in Idaho. Twenty counties have no population centre of 20 000 or more, but average six or more persons per square mile. Sparsely populated areas, vast desert and numerous mountain ranges complicate the delivery of primary health care services throughout the state.

Idaho has historically claimed one of the lowest physician-to-population ratios in the nation. Currently, Idaho ranks 48th in the nation with 125 physicians in patient care per 100 000. Recruiting health professionals to Idaho's remote areas and to regions in the state with substantial portions of low-income/undeserved persons is a constant challenge. The Idaho Department of Health and Welfare (IDHW) completed a survey of Idaho's rural community facilities (excluding Boise, Nampa, Caldwell, Idaho Falls, Lewiston and Coeur d'Alene) in October 1995, which indicated that 28 of the 47 facilities interviewed were actively recruiting primary care physicians or midlevel providers. The top two reasons for recruitment were increased demand and to relieve the workload of other primary care providers. The analysis completed by IDHW in 1994 indicated a maldistribution of primary care providers with an adequate concentration of physicians and midlevel providers in the more populated areas of the state and inadequate numbers in the more rural and frontier areas. The state Health Professional Loan Repayment Program and the Health Professional Clearinghouse are working to alleviate the challenges that many rural/underserved Idaho communities face when recruiting health care providers.

The decision problem is how to distribute limited funds ($250 000) in order to help the state counties attract health care professionals through repaying their education loans. The distribution of funds among the state counties should be equitable, i.e. based on demonstrated needs, and provide a high likelihood of successful hiring and longer-term retention of health care professionals. The strategy for arriving at a decision on how to allocate the funds is to develop a ranking of Idaho counties based on thirteen evaluation criteria, measured or estimated criterion outcomes, and relative importance weights attached to the criteria (Jankowski and Ewart 1996).

3.2. *Analysis of decision space using dynamically linked displays*

Malczewski (1999b) suggested that an appropriate technique supporting the analysis of decision space is the combination of maps with a 'value path' plot. However, the static, non-interactive display considered by Malczewski has its limitations. If the number of options is large, numerous intersections of lines make it very difficult to follow any particular value path. It is easy to see which option optimizes any particular criterion taken separately, but it is rather difficult to estimate its position in respect to the remaining criteria. Therefore Malczewski (1999b) warns that this technique can be recommended only if the number of options is relatively small. Another problem is that a spatial position of each 'partial optimum' is shown on a separate map. There are natural limitations on the number and size of maps to be put simultaneously on the computer screen. Therefore the number of criteria that can be considered simultaneously is also limited.

The problems with the value path/map display can be effectively overcome by adding interactivity and dynamics. Some highly interactive implementations of the value path plot have been developed in the EDA area (Inselberg 1998). In EDA it is customary to use the term 'parallel coordinates plot' rather than 'value path plot'. DECADE offers the user a single map display dynamically linked with the parallel coordinate plot. When the user points at some option represented in the map, the respective object becomes highlighted in the map, and the corresponding value path becomes highlighted in the parallel coordinate plot. And vice versa, the user may point at some line segment in the plot, and the whole value path becomes highlighted as well as the position of the corresponding option in the map. Such value path/map integration makes it easy to evaluate any option with regard to all decision criteria. Additionally, the user may 'fix' highlighting by clicking on an object in the map or on a line in the plot. The selected object remains highlighted when the mouse cursor moves out of the display or points at other objects. This enables the comparison of the value paths of two or more decision options. The fixed highlighting differs in colour from transient highlighting (white and yellow, respectively, see figure 1).

DECADE allows the user to combine visual analysis of input information, i.e. criterion and decision spaces, with running MCDM procedures and visualizing their results. In particular, one of the implemented MCDM procedures is based on the Ideal Point technique (Hwang and Yoon 1981). The Ideal Point technique involves three aspects: the Ideal (the best criterion scores), the Nadir (the worst criterion scores), and the distances from each option to the Ideal and the Nadir. Consequently, in the Ideal Point method each of the decision options is compared to the Ideal and the Nadir. The criterion outcome distances measured to the Ideal and Nadir, calculated for each decision option, are then aggregated into relative closeness measures. The closeness measure expresses how close each option is to the Ideal and conversely how far it is from the Nadir. The options are then ordered/ranked beginning with the one closest to the Ideal and furthest from the Nadir and ending with the one furthest from the Ideal and closest to the Nadir. The maximum possible evaluation score (the relative closeness measure) is 100, and the minimum possible is 0.

In DECADE the user can start the Ideal Point method from the window called 'Decision options outcome map' (see figure 1). This window is organized around a map depicting the geographical extent of the decision situation—in our case the map of Idaho by counties. Left to the map depicting the decision options the user finds a criterion weight panel providing the listing of decision criteria (ten criteria were selected and, hence, are listed in the panel). Above each criterion name there is a slider allowing the user to select/adjust the criterion weight within the value range (0, 1). The adjustment of one weight causes all other weights to automatically change values proportionally to their values before the adjustment. The north-east pointing arrow, to the left of each weight slider, indicates a benefit criterion (that is, its values

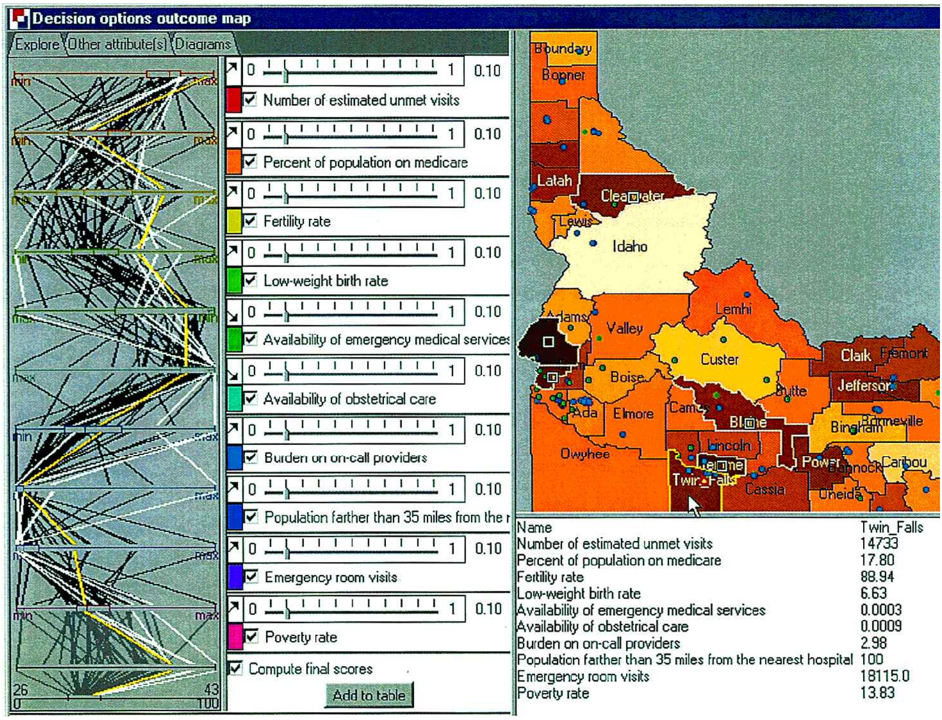

Figure 1. The performance of decision options presented on the map (Idaho counties) can be viewed on the parallel coordinates graph.

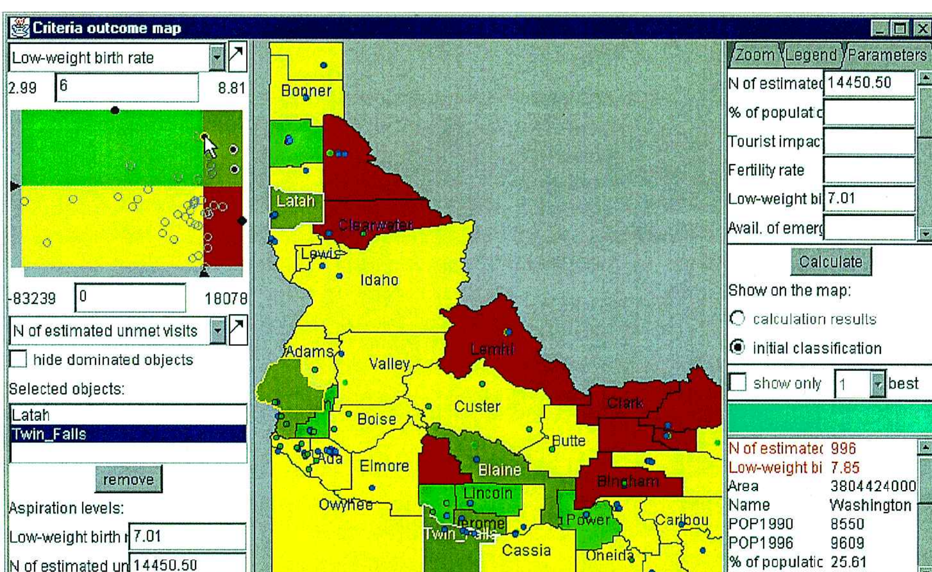

Figure 3. The interface for elicitation of criterion aspiration levels.

are maximized), the south-east pointing arrow corresponds to a cost criterion with values to be minimized. The user can easily change the directionality of criteria (from maximize to minimize) with a mouse click. The panel to the left of criterion values displays the outcomes for each Idaho county (in this case—decision option) on the plot of parallel coordinates. The lines in the parallel coordinate plot can be considered as trajectories of decision option outcomes.

When the user selects the checkbox 'Compute final scores' (see the bottom of figure 1), the system computes the evaluation of options by means of the Ideal Point algorithm and presents the results on the map. The degree of darkness assigned to objects (counties) in the map represents the scores: the higher the score, the darker is the hue. At the same time additional two axes appear in the parallel coordinates plot (see the bottom-left corner of figure 1). On these axes the final scores are represented: on the lower one the scores are mapped onto the scale from 0 to 100, and the upper axis represents the minimum–maximum score interval.

All three panels—map, criterion weights, and parallel coordinates plot are dynamically linked in the sense that any action (e.g. option selection, change of weight) in one panel is immediately reflected on the other two panels. Such a level of interactivity allows the simultaneous exploration of options in the geographical decision space and the decision option outcome space.

On the screen copy presented in figure 1, the user selected several counties with top final evaluation scores. The mouse was clicked when the pointer was located on the value paths that intersect the final score axis at the rightmost positions. White-segmented lines highlight the selected counties (*fixed highlighting*). Their distribution can be easily studied along each criterion axis and provide a sense of the influence of that criterion on the final evaluation outcome. The selected counties are also visible on the map of Idaho. Their borders are highlighted with white line. In figure 1 the cursor points at one of the top ranked counties—Twin Falls. As the map is sensitive to the cursor position, the boundary of this county is highlighted in yellow (*transient highlighting*), and its criterion outcomes are displayed below the map. The corresponding trajectory representing criterion outcomes is also displayed in yellow.

Any change of weights causes the final evaluation score to be immediately recomputed and the order of counties changed along the final score axis. The map is also dynamically updated in response to changes in criterion weights and final scores. This facilitates a sensitivity analysis where the user may select a group of trajectories representing specific counties and test if a slight change in criterion weights causes a change in the final score of selected counties.

Another type of map supporting spatial decision analysis and available in DECADE is a map for classifying decision options based on the distribution of final evaluation scores (see figure 2). It allows the decision maker to interactively break the range of evaluation scores into intervals and experiment with the number of intervals and their respective ranges. Accordingly, the objects in the map are divided into classes based on their evaluation scores. A supplementary display (the dot plot in the upper section of the map window) shows the distribution of the scores. DECADE makes it easy to form classes containing any desired number of options achieving satisfactory final evaluation scores. In the presented decision scenario the user created a class of fundable counties containing the top eight scoring counties (dark grey in figure 2) and a 'near fundable' class with two counties (medium grey in figure 2). The remaining 34 counties were classified as non-fundable.

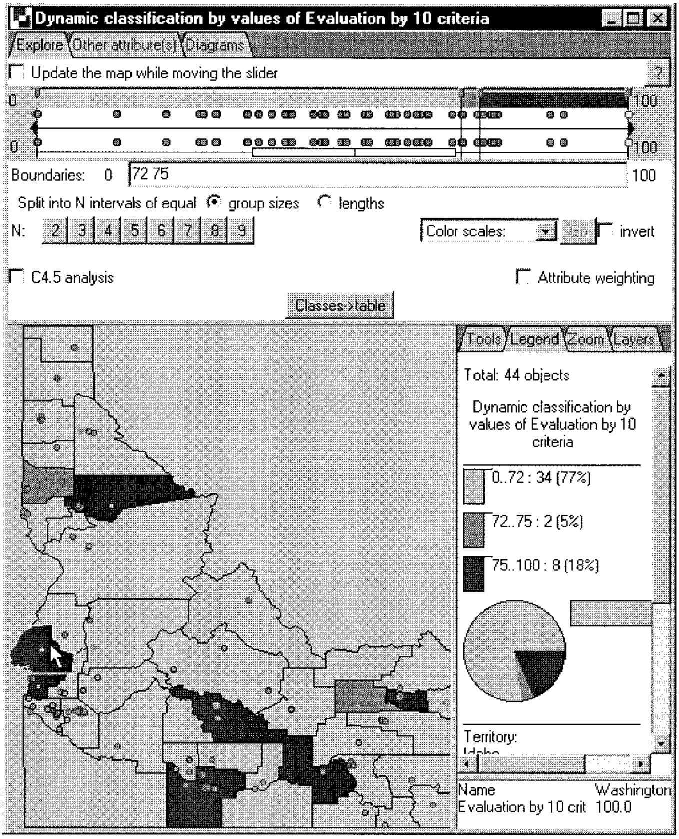

Figure 2. Classification of counties according to the evaluation scores computed using the Ideal Point method.

3.3. *Elicitation of aspiration levels for decision criteria*

The pairwise representation of criterion outcome and decision spaces opens up a possibility of eliciting the decision maker's preferences for criteria not only on the basis of data but also geography. In the previous section we discussed how preferences could be represented by criterion weights. We called this approach *explicit* representation of preferences. Besides the explicit approach, DECADE also offers an *implicit* approach to expressing preferences. In this approach the user, instead of specifying numeric weights corresponding to relative importance of criteria, selects *aspiration levels*. The concept of aspiration levels stems from Herbert Simon's ideas of bounded rationality and satisficing, and states that decision makers may be more interested in finding decision options that come close to certain target values (aspiration levels) rather than optimizing one objective at the expense of others (Simon 1960).

DECADE offers the capability of simultaneously visualizing criterion and decision spaces by linking the map with a scatterplot (see figure 3). As before the interface is organized around a map of the study area. The scatterplot shows the distribution of values of two criteria. A pair of criteria is selected by the user from two criteria lists—above and below the scatterplot, in the upper left of the map interface. Any pair of criteria can be selected, and their outcomes for all options (in our case Idaho counties) can be visualized in a bi-criterion outcome space. Additionally, the distribution of outcomes may be subdivided into classes. The spatial distribution of classes is depicted on the map by distinct hues representing each class.

In figure 3, the range of each of the selected criteria is divided into two intervals, below and above user-specified break-point values. Thus, for the criterion 'Number of estimated unmet visits' the user set the break point at zero to separate counties with positive values from those with negative values. Four groups of counties resulting from this classification are displayed on the map. Counties characterized by low values of both criteria (in our example low-weight birth rate and number of estimated unmet visits) are painted in yellow and those with high values of both criteria in chartreuse (green mixed with yellow). The green colour represents counties with high rates of low-weight births and negative values of unmet visits, and red corresponds to counties with low values of low-weight births and positive values of unmet visits. The scatterplot and the map are dynamically linked through the simultaneous highlighting of objects in the map and the corresponding dots in the scatterplot. These interactive features provide a convenient environment for the analysis of decision problem data.

The map interface provides access to two MCDM techniques: (1) Pareto-dominance based division of the decision space into non-dominated and dominated decision options, and (2) aspiration-levels based evaluation of decision options. The former technique is implemented through a graphical representation of dominated and non-dominated options in the scatterplot: all dominated options are hollow and only non-dominated are displayed as filled, black dots. Selection of the control 'hide dominated objects' below the scatterplot area results in the removal of dominated objects from both the scatterplot and the map.

In the example shown in figure 3, there are five non-dominated counties for the pair of criteria 'low-weight birth rate' and 'number of estimated unmet visits'. The user examined all non-dominated counties and decided to select two counties—Latah and Twin Falls as the geographical objects representing aspiration levels for low-weight birth rate and number of estimated unmet visits. The criterion outcomes of the selected options (i.e. counties) are accepted by the system as aspiration levels for the two criteria. If more than one option is selected, an aspiration level is computed as the median of the corresponding values. In the given example the levels are low-weight birth rate = 7.01%, number of estimated unmet visits = 14 450.50. The user can change aspiration levels for any decision criterion by selecting another point from the scatterplot, or an object from the map, or simply typing in a desired value into the appropriate field of the top-right panel. It is expected, however, that the decision maker will express her/his preferences mainly by a direct manipulation of the display, i.e. through the selection of objects in the map or in the scatterplot. In practice, there is no difference between the two ways of selection since an object becomes simultaneously highlighted both in the map and in the scatterplot as soon as the mouse cursor approaches it. The scatterplot informs the DM about the trade-offs between the criteria while the map depicts option locations, so the DM can base her/his selections on both sources of information (criterion outcomes and geographical decision space). The presented interface is primarily designed for the elicitation of user's implicit, geography-based preferences.

The aspiration levels expressed in this way become the basis for evaluating decision options and finding the best variants in terms of their closeness to the decision goal (given by the vector of aspiration levels). The search for best variants is based on the nearest solution algorithm (Lotfi *et al.* 1992). The nearest solution to goal is represented by the decision option that minimizes the maximum weighted distance between aspiration levels and each decision criterion. The algorithm always identifies solutions that are efficient (non-dominated). In DECADE the user can start the algorithm by pressing the button 'Calculate' in the centre of the right panel (see figure 3). The system calculates the scores on the basis of the available aspiration levels and immediately displays the results. The map in figure 4 depicts ten best counties in the sense of their closeness to the aspiration level-based goal.

DECADE offers the user an opportunity to compare the results of evaluating decision options with both MCDM techniques: Ideal Point with criterion weights and the nearest solution with aspiration levels. For the decision scenario results of the comparison are presented on the map in figure 5.

Out of ten top-scoring counties seven were selected with both MCDM techniques. From the methodological point of view, this level of agreement is to be expected provided that the user's preferences remain stable; both techniques are based on a similar concept of distance between decision options and multicriterion goal but use different mechanisms for expressing preferences (weights versus aspiration levels). From the user's perspective, an agreement between two evaluations made independently helps to raise the confidence in the obtained results.

The availability of two MCDM techniques in DECADE provides a flexibility to accommodate different styles of preference specification. Additional flexibility is offered by the interface to the aspiration level-based method; the DM can either enter aspiration levels explicitly or select the decision options to represent his/her target through direct interaction with the map.

The combination of map and scatterplot was suggested by Malczewski (1999b)

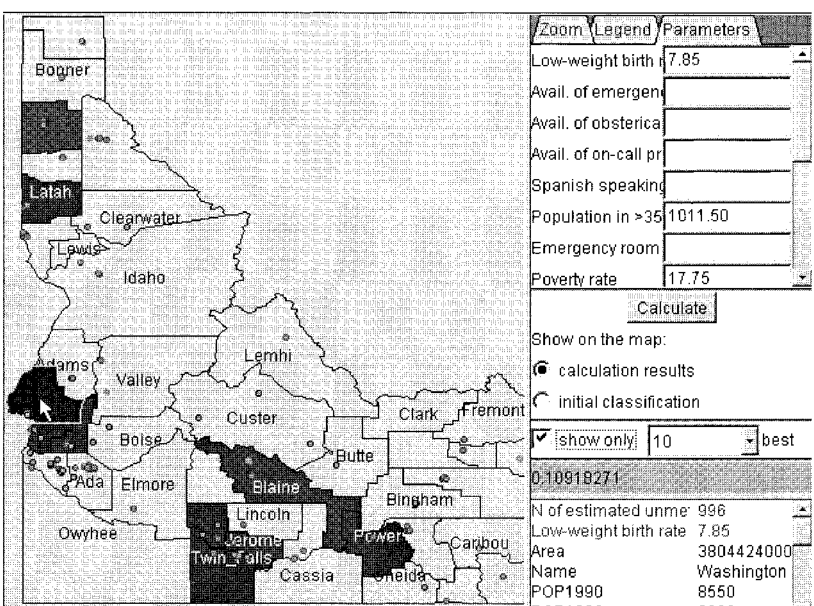

Figure 4. Ten Idaho counties that are nearest to aspiration levels selected for three criteria: low-weight birth rate, poverty rate, and population farther than 35 miles from the nearest hospital.

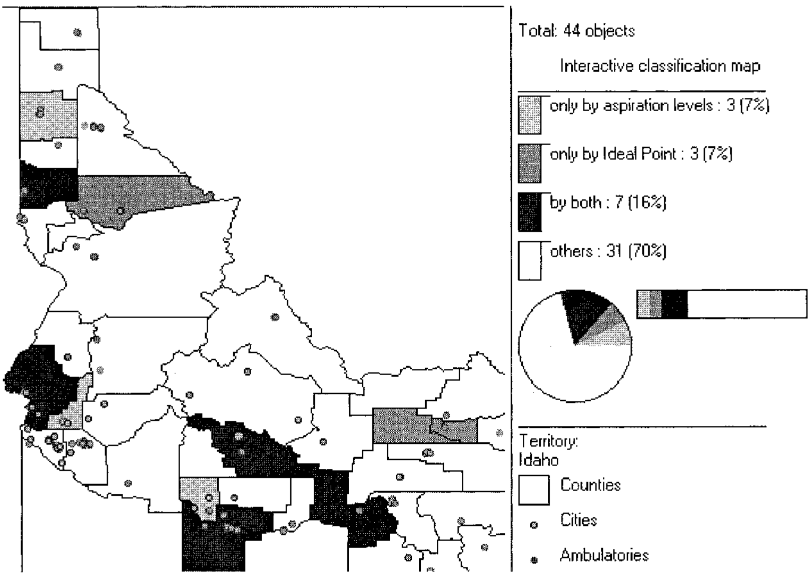

Figure 5. Comparison of evaluation results (ten top scoring counties) between Ideal Point and aspiration level-based nearest solution algorithms.

as an appropriate tool for the analysis of criteria and decision spaces. However, Malczewski considered a static presentation with multiple scatterplots representing all possible pairs of criteria and the same number of maps each showing the 'best' option for the corresponding pair of criteria. The shortcomings of this form of presentation are a limited number of criteria that can be considered and a low legibility of maps due to their small sizes. Interactivity and dynamic links between map and scatterplot can substantially reduce these limitations. In DECADE the user can select any pair of decision criteria. Consequently, one map/scatterplot pair is usually sufficient for the analysis of decision problem data. If, however, the user needs to directly compare two or more maps or scatterplots, this may be achieved by opening several windows.

3.4. *Transformation of geography-based preferences into decision criteria*

As already mentioned in $2.2, an interactive map can be used for direct classification of spatially distributed options on the basis of their geographical properties. DECADE provides a tool, called the interactive classification map, which supports this activity. Within our decision scenario, we demonstrate how the DM can use the interactive classification map to express his/her preferences and heuristics related to the geographical positions of options (i.e. counties). In the example considered there is one criterion that explicitly reflects the geographical aspect of the problem, namely, the number of people living farther than 35 miles from the nearest hospital. In other applications, however, it may be quite difficult to find a numeric measure that adequately represents geographical characteristics of options. In such cases the ability to extract spatial information from a map and encode it in a form suitable for further processing becomes important. We shall now demonstrate how this can be accomplished with the facilities available in DECADE.

Suppose that a DM is convinced that in order to allocate funds to counties it is necessary to consider the distribution of medical facilities (hospitals, ambulatories, primary care sites etc.) together with the network of roads providing access to these facilities. In the considered decision situation the respective criteria (distribution of medical facilities and road network) can be shown on the map. It is also relevant to take into account the population of counties. Unfortunately, exact data about the distribution of the population are not available. The DM decides to use the available information about the total population in each county. Additionally, using the population figures and county areas s/he computes county population densities using the spreadsheet calculation tool of DECADE.

On the DM's request, the system generates three maps: two maps showing the distribution of the total population and population density, respectively, and the interactive classification map. Every map shows locations of medical facilities and roads. All displays in DECADE are linked, which means the DM can select a county on any one of the maps and simultaneously visualize the values of attributes associated with this county. The maps presenting population data are supplemented with dot plot displays, as in the map shown in figure 2. On these displays the DM can observe a relative position of the selected county in regard to the population number and population density.

The interactive classification map (see figure 6) allows the DM to group geographic objects into classes through the direct manipulation of the map. Clicking on an object in the map results in appending it to an active class (thus, in figure 6 'good' is the active class). Using these facilities, the DM classifies counties into those with good, average, and poor connections to medical facilities.

In the next step, the classification can be transformed into a new attribute to be added to the existing table of decision attributes characterizing the counties. This can be accomplished by pressing the button 'Classes→table' in the upper left part of the window (see figure 6). In the example, the DM is interested in using the new attribute as an additional evaluation criterion. This requires the attribute to be numeric. Therefore the DM assigns numeric codes from 1 to 3 to categories (see figure 7). The new attribute now represents the DM-judged quality of intra-county connections to medical facilities: 1—means good connection, 2—average and 3—poor. The new attribute can be treated as an evaluation criterion in MCDM techniques. Thus, the DM can apply, for example, the Ideal Point method to evaluate counties taking into account the new, map-derived criterion.

It is interesting to compare the evaluation scores obtained with the addition of the new, ordinal-scale criterion with the scores computed using the ratio-scale criterion 'number of population farther than 35 miles from the nearest hospital'. Although the map-derived criterion is less precise and more subjective than the ratio-scale criterion (as the former is the result of DM's interpretation of thematic maps), the evaluation results demonstrate a rather high level of agreement.

For the comparison we used DECADE to create a map shown in figure 8. The map is supplemented with a scatterplot used to classify attribute value ranges. We broke each value range into two subintervals in such a way that the ten top-scoring counties in each evaluation became separated from the rest. Looking at the scatterplot, the reader will notice that seven counties received topmost scores in both evaluations. They are painted chartreuse in the map and can be found in the upper right quadrant of the scatterplot. Three counties that were evaluated among the top ten with the criterion 'number of population farther than 35 miles from the nearest hospital' are painted red in the map and can be found in the lower right quadrant

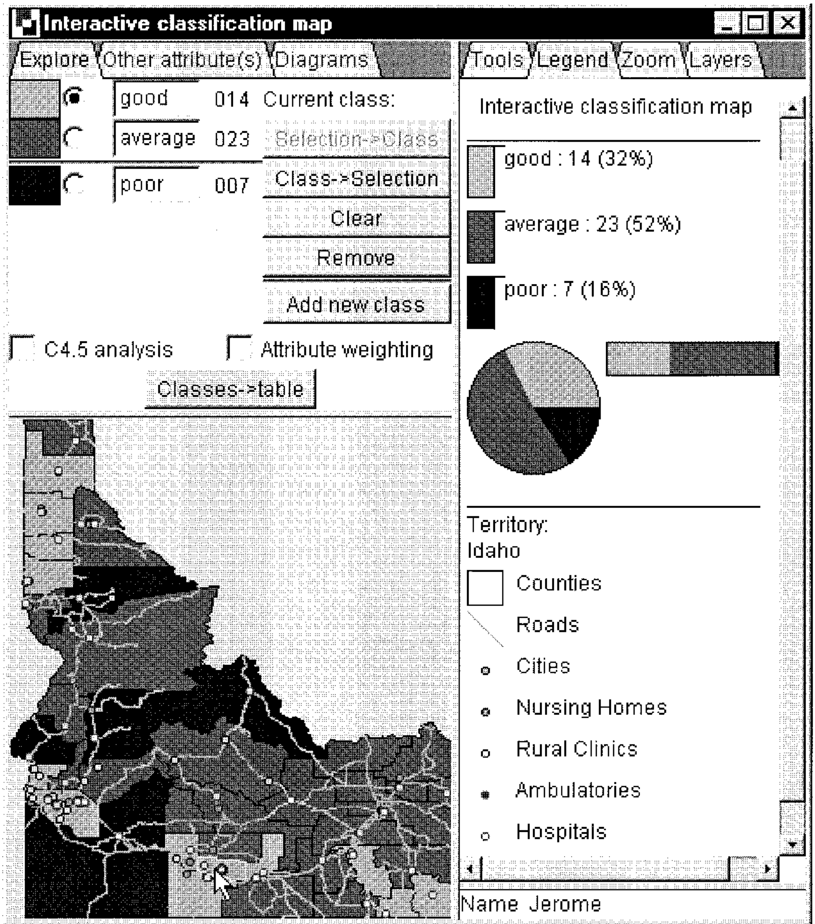

Figure 6. Results of interactive classification of counties accounting for the distribution of medical facilities.

of the scatterplot. There are also three counties classified in the top ten solely on the basis of the evaluation with the map-derived criterion. They are painted green in the map and are visible in the upper left of the scatterplot.

In our example we demonstrated that interactive map display makes it possible to capture map-based observations of spatial properties and relationships. By classifying geographical objects, one may transform map-derived observations into categorical knowledge. On the basis of this knowledge an attribute with nominal or ordinal scale properties can be derived.

3.5. *Reducing the dimensionality of the multicriteria decision problem using data mining*

Up to this point we have focussed our discussion on how interactive maps dynamically linked with other types of graphical displays can support spatial option evaluation. We described three examples of option evaluation based on different models of map use. These models were implemented with tools available in the

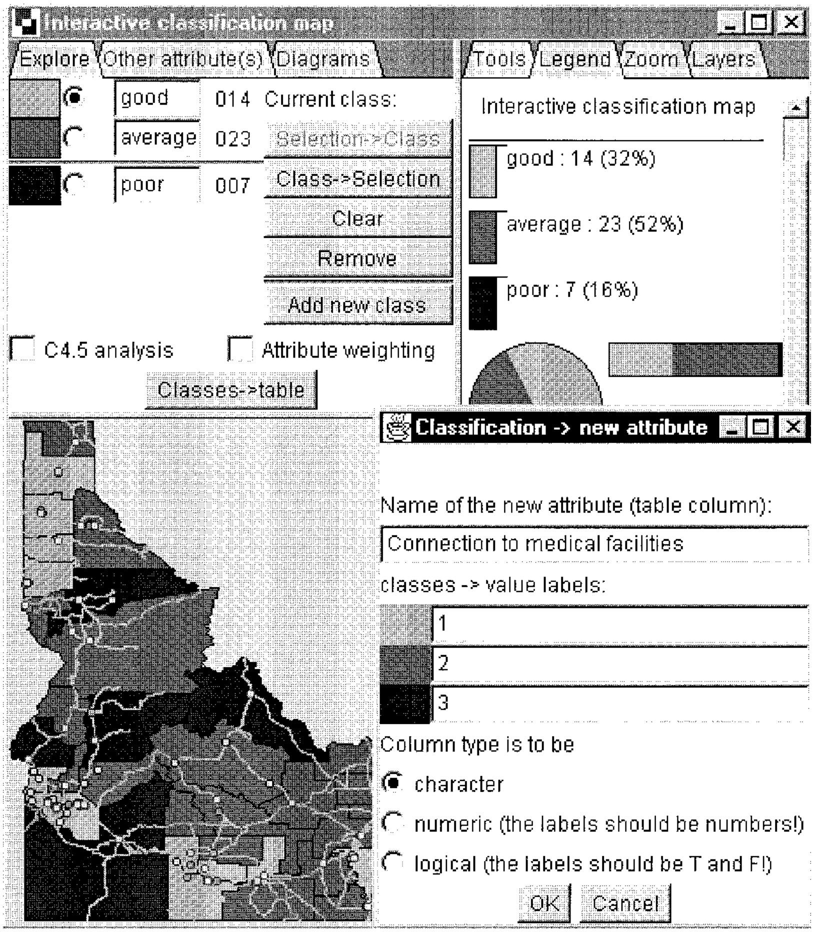

Figure 7. The map-derived classification of decision options can be converted to a numeric attribute.

spatial decision support system DECADE developed on the basis of the mapping software Descartes (Andrienko and Andrienko 1999a).

One of the recent developments in Descartes was its integration with data mining software called Kepler (Wrobel *et al.* 1996). Currently, the link between the systems Descartes and Kepler, allows the user to proceed from a map to data mining procedures applied to data depicted on the map. The results of data mining can be displayed in Kepler, and this display becomes dynamically linked with maps and other graphics in Descartes through simultaneous highlighting of corresponding elements in both systems. The form of Kepler's display depends on the data mining algorithm used. For example, results of some procedures can be represented in the form of a classification tree. The architecture of the link between Descartes and Kepler is described in Andrienko and Andrienko (1999b).

We considered as an interesting idea the use of the link between the two systems for investigating whether data mining algorithms might help reduce the cognitive complexity of spatial decision problem. Specifically, we were interested in the possibility of reducing the dimensionality of the criterion outcome space. If a subset of

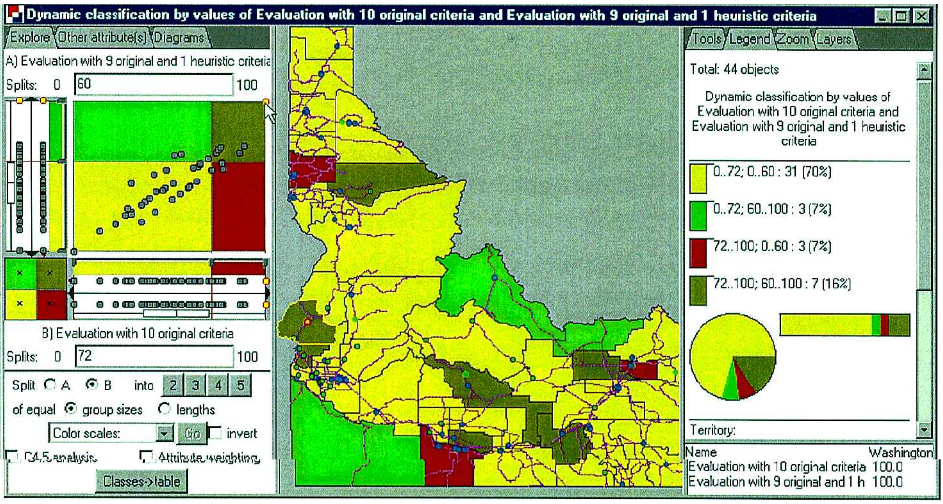

Figure 8. Visual comparison of evaluation results obtained with the criterion 'number of population farther than 35 miles from the nearest hospital' and with the new, map-derived criterion 'connection to medical facilities'.

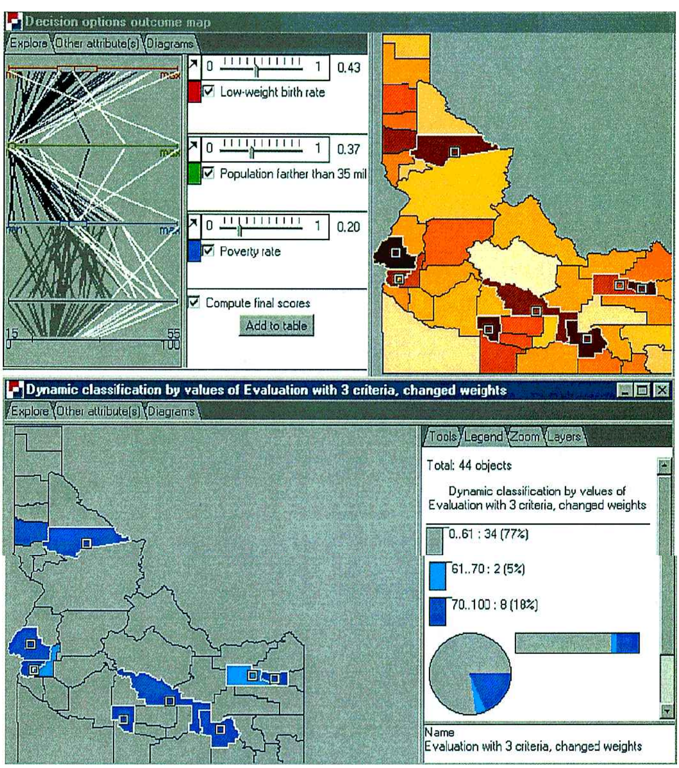

Figure 13. Noticeable change in the criterion weights resulted in only minor changes in the order of top scoring counties signifying the low sensitivity of evaluation criteria.

'strong' evaluation criteria could be found such that they would result in the same or near-same order of decision options as the full set of criteria, then the reduction of cognitive complexity might be achieved. Clearly, the smaller the number of evaluation criteria, the cognitively easier it becomes for the decision maker to express his/her preferences in regard to the criteria and, consequently, to classify, order, or choose the most satisfying decision option.

We carried out an experiment using one of the data mining methods, the C4.5 decision tree derivation algorithm (Quinlan 1993). We used the C4.5 algorithm because of its capability to find a minimum sufficient subset of criteria needed to explain a classification of decision options. The experiment yielded promising results. It demonstrated that data mining deserves further study as a potentially useful methodology for reducing the cognitive complexity of multicriteria spatial decision problems.

The C4.5 algorithm can be applied to any set of objects described by attributes and grouped into a finite number of classes. The nature of the classes is irrelevant for the algorithm. Whenever the user of Descartes has on the screen a map with any classification of geographical objects, she/he can submit this classification to the C4.5 algorithm with the purpose of finding classification-specific attributes (Andrienko and Andrienko 1999c). When the user activates this function, the assignment of geographic objects to classes is retrieved from the map and submitted to Kepler together with other available data about the objects (values of attributes).

The C4.5 algorithm tries to discriminate between the given classes on the basis of the values of available attributes and to produce a decision tree that divides the whole set of objects into groups resembling the specified classes. Each tree node represents a step in the division based on values of one attribute. For example, the counties of Idaho may be divided into those with poverty rate below and above 13.76%. It is important to note that the C4.5 algorithm tries to find the most discriminative attributes, that is the most relevant in regard to the given classification.

Our particular interest was in applying the algorithm to data about decision options classified into acceptable, near acceptable and un-acceptable. Such a classification can be based, for example, on the results of one of the MCDM techniques available in DECADE. Specifically, in our experiment we used the results of the evaluation with the Ideal Point method described in $3.2. The evaluation was done with equal weights assigned to ten criteria. According to the final evaluation scores, the counties of Idaho were classified into fundable (8 counties), near fundable (2 counties), and unfundable (the remaining 34 counties). Our goal was to check whether the same or similar classification could be obtained with a fewer number of criteria. We expected that the number of criteria representing the nodes of the decision tree would be less then the initial ten, and, hence, the method would reveal the most significant criteria. If so, then the cognitive complexity of the decision problem would be reduced by dropping off the irrelevant criteria.

The results of submitting the classification of Idaho counties (three groups: fundable, near-fundable, and non-fundable) to the C4.5 algorithm are presented in figure 9.

Each node in the tree specifies the test of some evaluation criterion, and each branch descending from that node corresponds to one of the possible value intervals for this criterion. The decision tree shows that there are only three relevant criteria for the classification (presented earlier in figure 2): (1) Low-weight birth rate, (2) Population further than 35 miles from the nearest hospital, and (3) Poverty rate

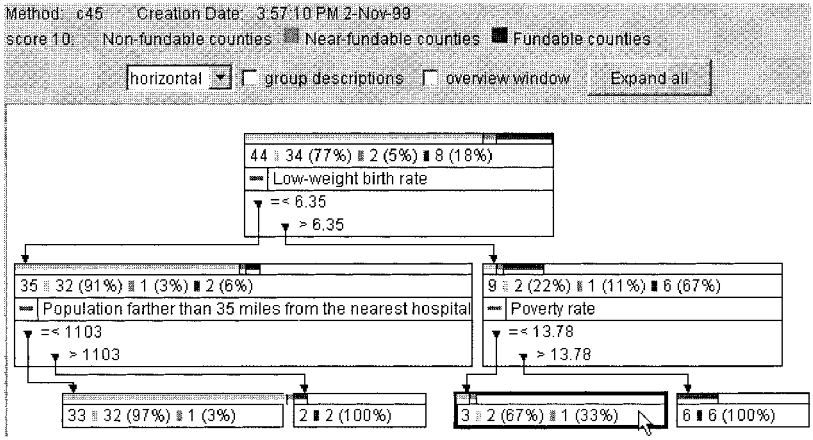

Figure 9. Decision tree for classification of Idaho counties into three subgroups: fundable, near fundable, and non-fundable.

(see figure 9). In other words, the classification of Idaho counties, based on the results of evaluation with ten criteria, can be 'explained' by three criteria. The top node criterion in figure 9—Low-weight birth rate, splits 44 Idaho counties along two branches: 35 counties for which Low-weight birth rate was less than or equal to 6.35% of all births and 9 counties that had Low-weight birth rate exceeding 6.35%. The former branch leads into the node 'Population farther than 35 miles from the nearest hospital' and the latter into the node 'Poverty rate'. Test conditions at these two nodes result then in four branches leading into leaf-nodes. For example, the node 'Poverty rate' along with 'Low-weight birth rate greater than 6.35%' classifies nine counties. These counties are then subdivided into a subgroup of six counties belonging to the 'fundable' class and three counties, from which one belongs to the 'near-fundable' class and two to the 'unfundable' class (figure 9).

As already mentioned, graphical displays of data mining results in Kepler are linked with all maps and supplementary graphics in Descartes. Thus, the distribution of decision options at any node of the tree can be displayed on the decision options outcome map (see figure 10). In the example, three counties of the third leaf node in figure 9 (counting from left to right) are highlighted in figure 10. It is also possible to select an object (county) or a set of objects in the map or other graphics in Descartes and see its/their positions in the nodes of the decision tree.

Since the C4.5 algorithm found that only three criteria were important for the given classification, one would expect to obtain a similar classification using instead of ten only these three criteria—a significant reduction of the dimensionality problem. In order to test this, we re-evaluated the need of Idaho counties for funding using three criteria contained in the decision tree (see figure 11). Next, we compared the result of re-evaluation with the initial evaluation obtained using ten evaluation criteria (see figures 10 and 11). The comparison results are presented in figure 12.

In both maps in figure 12 the darkest shade of grey marks counties chosen as fundable according to each of the evaluations. These are the counties with the top eight scores. Additionally, the counties classified as fundable by means of ten evaluation criteria are simultaneously highlighted in both maps with white boundaries and small square symbols.

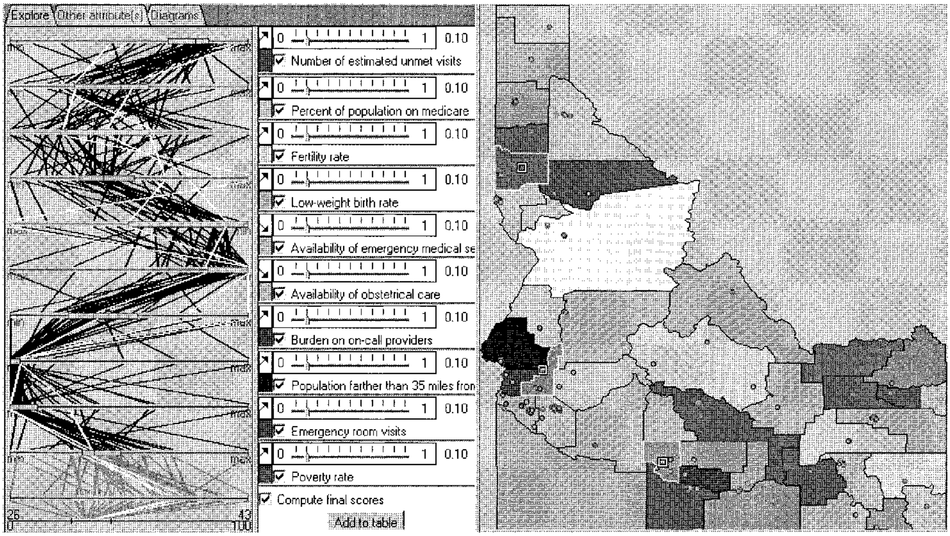

Figure 10. Distribution of decision options at any node in the decision tree can be displayed in the decision options outcome map.

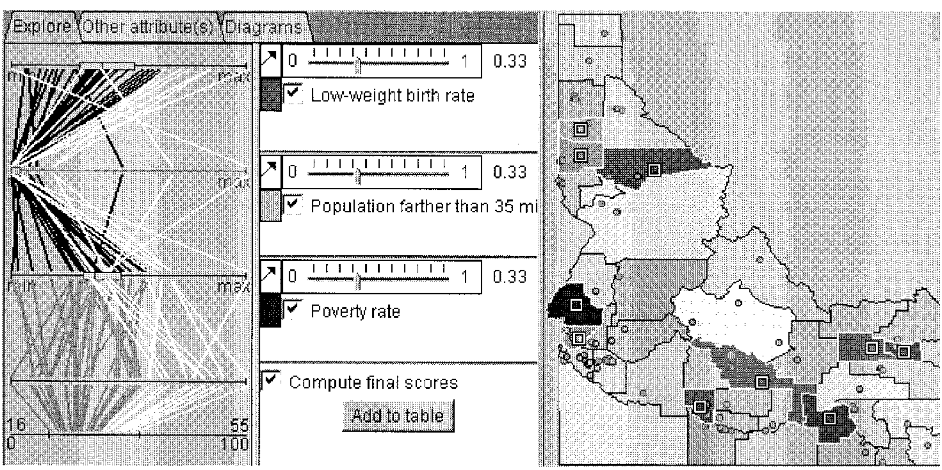

Figure 11. Top ten scoring counties are highlighted simultaneously in the decision option outcome panel and on the map. The evaluation was computed with three criteria.

Comparison of the maps in figure 12 shows a high degree of agreement between the two evaluations: out of eight counties that received top scores in the first evaluation (with ten criteria) six counties (75%) were also among the top eight in the second evaluation (with only three criteria). From the two counties classified as near fundable in the first evaluation one was among the top eight and the other among the top ten in the second evaluation. We repeated both evaluations using different sets of weights and obtained very similar results providing the basis for reducing the number of evaluation criteria and hence, simplifying the problem complexity.

In the next step of this exploratory multiple criteria decision analysis the user could engage in selecting criterion weights and testing the sensitivity of weight

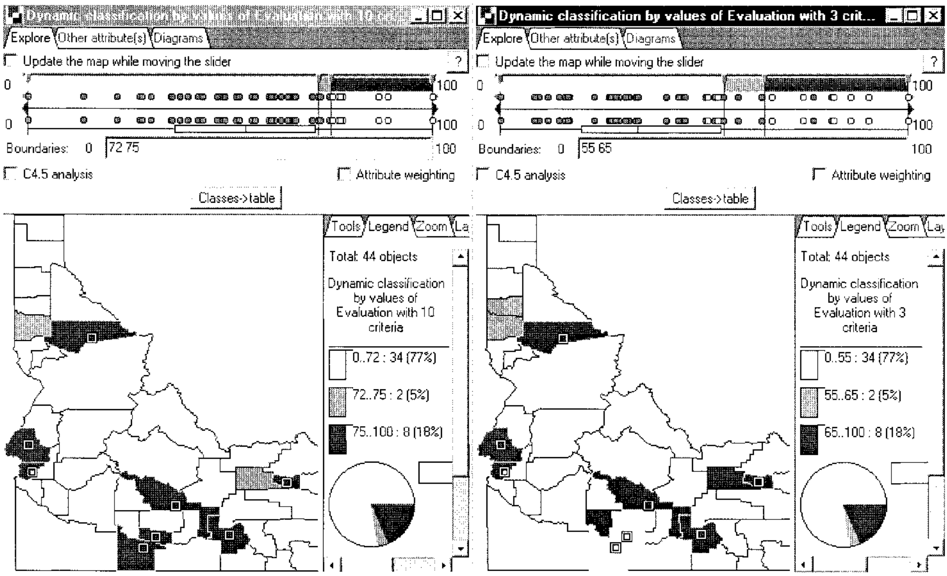

Figure 12. Comparison of two multiple criteria evaluations using first ten criteria and subsequently only three criteria.

changes on final decision option outcomes. The bottom map in figure 13 shows the results of evaluation after changing criterion weights (new weights depicted in the top map in figure 13. Compare with weights in figure 11). The reader can compare the map at the bottom of figure 13 with the map at the right of figure 12. The top eight counties resulting from the evaluation with equal weights are highlighted with white square symbols and white boundary lines. Comparing both figures the reader can see that seven out of eight highlighted counties received topmost scores in the second evaluation. The remaining county received a lower score and moved to the near-fundable class. One of the two counties that was near fundable according to the first evaluation moved to the group of fundable (top 8) counties in the result of the second evaluation. These changes in the evaluation results are not dramatic and show a rather low sensitivity of the evaluation results.

It is important to note that the outcomes of using the C4.5 algorithm, as applied in our experiments, can be driven by a large number of parameters. For example, the form of decision tree derived from the same data set may vary from one run of the algorithm to another. However, for our particular purposes the form of the tree given by the number of branches and nodes was irrelevant. We were interested in the attributes exposed at the upper tree levels. These attributes were expected to be the most informative for the discrimination between the classes. Based on our experimentation with the C4.5 algorithm we found that changes of its parameters did not change its effectiveness in discrimination between the classes. Hence, one may expect that the most informative attributes will almost always be found near the top of the tree.

In principle, there may be several equally informative attributes with respect to a subset of objects. The settings of the algorithm may influence which of these attributes will be selected at a given stage of the division of the subset and included in the corresponding tree node. Therefore it may be appropriate to determine the subset of important (discriminant) attributes by running the algorithm several times with different settings. Another option is to use for this purpose other data mining techniques. Andrienko *et al.* (2000) describe the application of the so called 'feature weighting' technique which explicitly assigns numeric weights to attributes according to their importance. In the experiment described here we verified the results obtained by the C4.5 algorithm using the feature weighting technique. Changes of the decision tree due to modification of the parameters were insignificant. The 'feature weighting' algorithm assigned the highest weights to the attributes exposed in the decision tree thus confirming the selection of three discriminant criteria with the C4.5 algorithm.

4. Conclusion and directions for further research

We have shown how exploratory data techniques such as interactive maps and data mining methods can support spatial decision making. The research reported here was conducted utilizing the capabilities of Descartes—a mapping system designed to support exploratory analysis of spatially referenced data with interactive, dynamically changeable maps. We also employed the link of Descartes to a data mining system Kepler. For the study we used a real decision-making problem of allocating funds for primary health care services to Idaho counties. We described four multicriteria decision cases, each employing different decision support tools. The first three cases demonstrate the application of interactive maps and the fourth one the use of data mining algorithm integrated with interactive map displays. The first three cases differ in the pattern of using maps as aids in decision making. The order in which they are presented corresponds to the growing role played by the map.

In the first case the map helps the DM analyse data characterizing the decision problem. The DM has an opportunity to estimate the performance of any option with regard to all criteria by viewing its 'value path' in the parallel coordinates plot as well as to compare performances of two or more options. The map is dynamically linked with the plot and serves as a 'visual index'; through which the user points at decision options, that can also be viewed and analysed in the plot. The map also allows an easy identification of options with particular 'value paths' selected from the plot.

In the second example the map is used to capture the decision maker's target values for the criteria (aspiration levels). It is important that the DM does not have to enter the aspiration levels explicitly. Instead s/he selects suitable compromise variants directly in the map or in the supplementary scatterplot. In so doing, the DM can express implicit geography-based preferences and heuristics.

In the third case the interactive map serves as a tool to capture and encode geography-induced knowledge the user believes to be important for decision making. The knowledge is expressed through an ordinal scale attribute reflecting the user's spatial preferences. This attribute can be used as a criterion in the evaluation of options with the use of MCDM procedures. Such a map use exploits the human's capability of seeing: the human's eye may capture from a map a lot of valuable information that can hardly be obtained using any kind of automatic processing. The map is a rich source and good organizer of information, and it also acts as a catalyzer of preferences and heuristics important in selecting spatial options. The emphasis on visual, map-based analysis does not contradict, however, the use of geocomputational functions available in many GIS, e.g. generation of buffer zones to support visual observation. On the contrary, it complements spatial decision analysis.

The fourth case demonstrates the possibility of supporting spatial decision making with data mining techniques. The experiment described in the paper demonstrates the potential of data mining techniques to help reduce cognitive complexity of a decision-making problem. Interactive maps were also involved in the experiment; they helped to interpret and analyse the results of data mining-based criterion selection.

The results of our research demonstrate that exploratory data techniques can be successfully applied to support multicriteria spatial decision making. The high level of interaction between maps and attribute data graphs opens up new possibilities for the integration of criterion outcome and geographical decision spaces, thus allowing the decision maker to better understand the structure of the decision problem at hand. Reduction of the decision problem cognitive complexity through the elimination of 'weak' evaluation criteria also serves the purpose of structuring the decision problem, which is the basic tenet of decision support.

In the future we plan to continue research on the use of various data mining methods in spatial decision making. We also intend to undertake an investigation of map-based support for group decision making, in particular, the distributed (in space and time) decision making by a group of stakeholders, exports and decision makers. The activities we consider as likely to benefit from map support are exchange of opinions, arguments, and proposed solutions, comparison of solutions proposed by different group members, and aggregation of individual solutions into a consensus-based decision.

Acknowledgments

We would like to thank Drs Hans Voss and Klaus Rinner (GMD—German National Research Centre on Information Technology) and Professor Alexander Lotov (Moscow State University, Russia) for useful comments and fruitful discussion about the content of the paper.

References

Andrienko, G., and Andrienko, N., 1999a, Interactive maps for visual data exploration. *International Journal of Geographical Information Science*, **13**, 355–374.

Andrienko, G., and Andrienko, N., 1999b, Knowledge-Based Visualization to Support Spatial Data Mining. In *Advances in Intelligent Data Analysis*, *Proceedings of the 3rd International Symposium, IDA-99, Amsterdam*, Lecture Notes in Computer Science, vol. 1642, edited by D. J. Hand, J. N. Kok and M. R. Berthold (Berlin: Springer-Verlag), pp. 149–160.

Andrienko, G., and Andrienko, N., 1999c, Data Mining with C4.5 and Cartographic Visualization. In *User Interfaces to Data Intensive Systems*, edited by N. W. Paton and T. Griffiths (Los Alamitos, CA: IEEE Computer Society), pp. 162–165.

Andrienko, G., Jankowski, P., and Andrienko, N., 2000, Data mining approach to reducing cognitive complexity of multiple criteria spatial decision problem. In *Proceedings of Spatial Data Handling SDH'2000* (Beijing, China: The Study Group on Geographical Information Science of the International Geographical Union), pp. 3015–3028.

Armstrong, M. P., Densham, P. J., Lolonis, P., and Rushton, G., 1992, Cartographic displays to support locational decision making. *Cartography and Geographic Information Systems*, **19**, 154–164.

Buja, A., McDonald, J. A., Michalak, J., and Stuetzle, W., 1991, Interactive data visualization using focusing and linking. In *Proceedings IEEE Visualization'91* (Washington: IEEE Computer Society Press), pp. 156–163.

Buja, A., Cook, D., and Swayne, D. F., 1996, Interactive high-dimensional data visualization. *Journal of Computational and Graphical Statistics*, **5**, 78–99.

Carver, S., 1991, Integrating multicriteria evaluation with GIS. *International Journal of Geographical Information Systems*, **5**, 321–339.

Casner, S. M., 1991, A task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics*, **10**, (April 1991), pp. 111–151.

Chun, K. J., and Park, H. K., 1998. Examining the conflicting results of GDSS research. *Information & Management*, **33**, 313–325.

Church, R. L., Loban, S. R., and Lombard, K., 1992, An interface for exploring spatial alternatives for a corridor location problem. *Computers & Geosciences*, **8**, 1095–1105.

Cohon, J. L., 1978, *Multiobjective programming and planning* (New York: Academic Press).

Dykes, J. A., 1997, Exploring spatial data representation with dynamic graphics. *Computers & Geosciences*, **23**, 345–370.

Eastman, J. R., 1997, IDRISI for Windows, version 2.0: tutorial exercises. Worcester, MA, Graduate School of Geography, Clark University.

Eastman, J. R., Jin, W., Kyem, P. A. K., and Toledano, J., 1995. Raster procedures for multi-criteria/multi objective decisions. *Photogrammetric Engineering & Remote Sensing*, **61**, 539–547.

Faber, B., Wallace, W., and Cuthbertson, J., 1995, Advances in collaborative GIS for land resource negotiation. *Proceedings, GIS'95, Ninth Annual Symposium on Geographic Information Systems, Vancouver, B. C., March 1995* (Fort Collins, CO: GIS World, Inc.), **1**, 183–189.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P., 1996, The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, **39**, 27–34.

Fisher, G., Markowski, M., and Antoine, J., 1996, Multiple criteria land use analysis. Working Paper WP-96-006. Laxenburg, Austria, International Institute for Applied Systems Analysis.

GeoChoice Inc. 1999, Accessed 30 September, 1999, http://www.geochoice.com/software.html.

Hwang, C. L., and Yoon, K., 1981, *Multiple Attribute Decision Making: Methods and Applications* (Berlin: Springer-Verlag).

Inselberg, A., 1998, Visual data mining with parallel coordinates. *Computational Statistics*, **13**, 47–63.

Jankowski, P., 1995, Integrating geographic information systems and multiple criteria decision making methods. *International Journal of Geographical Information Systems*, **9**, 251–273.

Jankowski, P., and Ewart, G., 1996, Spatial decision support system for health practitioners: selecting a location of rural health practice. *Geographical Systems*, **3**, 279–299.

Jankowski, P., Nyerges, T. L., Smith, A., Moore, T. J., and Horvath, E., 1997, Spatial group choice: a SDSS tool for collaborative spatial decision-making. *International Journal of Geographical Information Systems*, **11**, 577–602.

Jankowski, P., Lotov, A., and Gusev. D., 1999, Application of multiple criteria trade-off approach to spatial decision making. In *Spatial Multicriteria Decision Making and Analysis: A geographic information sciences approach,* edited by Jean-Claude Thill (Aldershot: Ashgate), pp. 127–148.

Jankowski, P., and Nyerges, T. L., in press, GIS-supported collaborative decision making: results of an experiment. *Annals of the Associations of American Geographers.*

Janssen, R., and Rietveld, P., 1990, Multicriteria analysis and GIS: an application to agricultural landuse in The Netherlands. In *Geographical Systems and Regional Planning*, edited by H. J. Scholten and J. C. Stilwell (Dordrecht: Kluwer).

Kirkwood, C. W., 1997, *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets* (Belmont, CA: Duxbury Press).

Lotfi, V., Stewart, T. J., and Zionts, S., 1992, An aspiration-level interactive model for multiple criteria decision making. *Computers Operations Research*, **19**, 671–681.

Lotov, A. V., Bourmistrova, L. V., and Bushenkov, V. A.,1999, Efficient strategies. An application in water quality planning. In *Decision Analysis and support for sustainable development*, edited by G. Kersten, Z. Mikolajuk and A. Yeh (Dordrecht: Kluwer), pp.145–166.

Lotov, A. V., Bushenkov, V. A., Chernov, A. V., Gusev, D. V., and Kamenev, G. K., 1997, Internet, GIS, and interactive decision maps. *Journal of Geographical Information and Decision Analysis*, **1**, 118–143

MacDougall, E. B., 1992, Exploratory analysis, dynamic statistical visualization, and geographic information systems. *Cartography and Geographic Information Systems*, **19,** 237–246.
MacEachren, A. M., 1994, Visualization in modern cartography: setting the agenda. In *Visualisation in Modern Cartography* (New York: Elsevier Science Inc.), pp. 1–12.
MacEachren, A. M., and Kraak, M.-J., 1997, Exploratory cartographic visualization: advancing the agenda. *Computers & Geosciences*, **23**, 335–344.
Malczewski, J., 1999a, *GIS and Multicriteria Decision Analysis* (New York: John Wiley & Sons).
Malczewski, J., 1999b, Visualization in multicriteria spatial decision support systems. Geomatica, **53**, 139–147
Monmonier, M., 1989, Geographic brushing: enhancing exploratory analysis of the scatterplot matrix. *Cartographycal Analysis*, **21**, 81–84.
Quinlan, J. R., 1993, *C4.5: Programs for Machine Learning* (San Mateo: Morgan Kaufmann Publishers).
Simon, H. A., 1960, *The New Science of Management Decision* (New York: Harper & Row).
Tukey, J. W., 1977, *Exploratory Data Analysis* (Reading: Addison-Wesley)
Unwin, A., and Hofmann, H., 1998, New interactive graphic tools for exploratory analysis of spatial data. In *Innovations in GIS–5* (London: Taylor and Francis), pp.46–55.
Wrobel, S., Wettschereck, D., Sommer, E., and Emde, W., 1996, Extensibility in Data Mining Systems. In *Proceedings of KDD'96 2nd International Conference on Knowledge Discovery and Data Mining* (Menlo Park, CA: AAAI Press), pp. 214–219.
Wu, F., 1998, SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *International Journal of Geographical Information Science*, **12**, 63–82.