Aalto University
School of Science
Degree programme in Systems and Operations Research

Marianne Honkasaari

Computation of Non-Dominated Alternatives in Spatial Decision Analysis

Instructor: Mikko Harju
Supervisor: Prof. Kai Virtanen

MS-E2108 Independent Research Projects in Systems Analysis
August 25, 2016

Contents

1 Introduction

People make decisions continuously, whether small, like which route to take to work, or big, like where to live. In the making of every decision, some analysis is involved. The decision maker must consider possible outcomes of different alternatives and strive to make the best decision according to his or her preferences.

Good decisions are not always easy to make. Preferences can be uncertain, outcomes unknown or there could be too many attributes or alternatives to even consider everything thoroughly. In some decision problems, it is important to find the best solution. For instance, when considering the number and locations of fire stations in a city, there are people's lives at stake.

To help find the preferred solution, there exist decision-making methods (see, e.g., Hwang and Yoon (2012)) that identify and assess important aspects of a decision and use mathematical models to find the most advantageous alternative. This study focuses on spatial decision analysis (see, e.g., Malczewski and Rinner (2015)) that takes incomplete information (see, e.g., Weber (1987)) into account. Many decisions have a spatial context, like the fire station example above. Consequences can vary across a geographic region depending on the chosen alternative.

In spatial decision making, data over a geographic region is used. In other words decision making occurs in the geographic information systems (GIS) context. GIS provides tools for information acquisition, storage, modeling, analysis, and management and integrates them in applications that solve problems related to spatial information (Chen, 2010). GIS is widely used in support of research in geography because it can associate locations in space-time with properties such as population density, temperature and water quality (Goodchild, 2009). The use and development of GIS is discussed, for instance, by Obermeyer and Pinto (2007) and Goodchild (1991) and the use of it in multicriteria decision making (MCDM) for instance by Jankowski (1995, 2006) and Malczewski and Rinner (2015).

Performances of alternatives are measured with attributes (see French (1988); Hwang and Yoon (2012)). For example, an attribute could be 'the time it takes for fire fighters to come from the nearest station' and the level of that attribute can be '5 minutes'. In decisions, different alternatives that can be selected have different outcomes that depend on the levels of the attributes and the GIS data. Those levels may depend on the alternative that is selected, the geographic location or the time instant that is observed, or some or all of them. Total value of a decision can be computed by aggregating valued outcomes that are weighted spatially, temporally, and attribute specifically (e.g., Simon et al. (2013)).

Most multiattribute decision analysis methods (e.g., Hwang and Yoon (2012)) are based on the assumption that there exist complete information about the model parameters. In practice this is challenging to achieve because decision maker's (DM) preferences rarely are precisely known. For instance, when the fire station example is considered, the DM should know exactly the importance of rapid action in every location whereas when using the model that accepts incomplete information (e.g., Weber (1987); Simon et al. (2013)) the DM has to provide only constraints for the weights in some regions, which is much more realistic.

With incomplete information it is not possible to compute exact values of different alternatives but their minimum and maximum values can be found, which can be used to limit the number of relevant alternatives through dominance (see Eisenführ et al. (2010); Kirkwood and Sarin (1985)). When dominance relations are found with the extreme values, dominated alternatives can be discarded, for non-dominated alternatives are preferred to the dominated ones.

This study presents spatial preference decision models and their extension to function with incomplete preference information. An example that illustrates the computation of non-dominated alternatives with incomplete information is constructed with a tool developed for the models introduced in this study. The implemented tool computes the values of different alternatives and finds the non-dominated alternatives.

The study is structured as follows. Section 2 presents the decision analysis through preference models and dominance. Different cases of spatial preference models with spatial, attribute, and temporal weights allowing information to be incomplete are reviewed in Section 3, and the experimental setup in Section 4 applies some of these models by illustrating the problem of choosing best fire station locations in Espoo. Section 5 provides conclusions.

2 Spatial decision analysis

The basis for spatial decision analysis is discussed in this section. The concept of preference and its connection to value functions are first defined. The additive multiattribute value function is then presented, after which preference models with a spatial aspect and the concept of dominance are introduced.

2.1 Preference models

Preference models are used to determine the relative desirability of alternatives. Alternatives z can be for example vectors or functions and $z \in Z$, where Z is the set of all possible alternatives. In practice there exists a subset $Z' \subset Z$ of feasible alternatives, among which the best alternative is sought. For more information about preference modeling, for instance Öztürké et al. (2005) and Keeney and Raiffa (1993) provide overview of the subject.

The DM's preferences are described by the relation $\succeq$ so that

$$z \succeq z',$$

which means that the DM weakly prefers object z to object z'. In other words, she holds z to be at least as good as z' (French, 1988).

Preferences can be presented in mathematical form by a value function but there are some conditions that must be met for a value function to exist. Two axioms are always needed:

A1: $\succeq$ is complete

$$\text{For any } z, z' \in Z, \text{ either } z \succeq z' \text{ or } z' \succeq z \text{ or both,}$$

A2: $\succeq$ is transitive

$$\text{If } z \succeq z' \text{ and } z' \succeq z'' \text{ then } z \succeq z'' \, .$$

Depending on the value function, there exists also other axioms (see, e.g., Simon et al. (2013)) but they are not discussed in this paper. If the axioms hold, there exists a value function $V(z)$ such that

$$V(z) \geq V(z') \Leftrightarrow z \succeq z', \tag{1}$$

which means that the alternative with the bigger value is preferred to the alternative with the smaller value.

2.2 Additive multiattribute value function

Often decision making requires multiple attributes (for more about multi-attribute decision making, see, e.g., Hwang and Yoon (2012); Keeney and Raiffa (1993)). To determine an additive multiattribute value function, it is assumed that z is a vector, where each element z_i represents the specific level of the attribute Z_i, $i = 1, 2, ..., n$. For every attribute Z_i, there exists a value function $v_i(z_i) \in [0, 1]$. Alternatives have different outcomes, and the preferences over the set of outcomes can be presented by aggregating the weighted values of outcomes. If the attributes are mutually preferentially independent (Dyer, 2005) and each attribute is difference independent of the others (Dyer, 2005), the aggregation function is an additive value function $V(z_1, z_2, ..., z_n)$ of the form

$$V(z) = \sum_{i=1}^{n} w_i v_i(z_i), \tag{2}$$

where $w_1, w_2, ..., w_n$ are non-negative weights and v_i is a value function over the ith attribute. Weights are scaled so that they add up to one $\sum_{i=1}^{n} w_i = 1$ (Keeney and Raiffa, 1993; Salo and Hämäläinen, 2010).

2.3 Spatial preference models

Outcomes of decision alternatives are modeled across the region in spatial decision making, and both the alternative that is selected and the spatial location can have an impact on the levels of the attributes. The discrete preference model assumes that the region of interest S is finite, that is, it consists of a specific number of subregions that are labeled $1, 2, ..., n$. Attribute levels do no vary within any subregion. The non-discrete model assumes that S is infinite, i.e., it consists of an infinite number of locations. Locations are dense and the model is continuous so the value can be computed everywhere within the region of interest. The consequence in the location $s \in S$ is determined by a function $z(s)$. For both discrete and non-discrete models, preferences can be described by a preference relation (1).

For the spatial preference model, there exists a value function (Simon et al., 2013) which can be presented as a discrete (3a) or a non-discrete (3b) function of the form

$$V(z) = \sum_{i=1}^{n} a(s_i) v(z(s_i)), \tag{3a}$$

$$V(z) = \int_S a(s) v(z(s)) \, \mathrm{d}s, \tag{3b}$$

where $v(z(s))$ is a value function over $z(s)$, $z(s)$ is a specific level of Z in the location s and $a(s)$ is a non-negative spatial weight in the location $s \in S$. In the discrete model, s_i is a location on subregion i. In the discrete model, $\sum_i a_i = 1$ and in the non-discrete model $\int_S a(s)\mathrm{d}s = 1$.

Some decisions require multiple attributes so the single-attribute preference model has to be expanded to a multiattribute model. One or more of these attributes can vary geographically. In this case, v in Equations (3a, 3b) is altered to a multiattribute value function, and there exists an additive value function $V(z)$ of the form (Simon et al., 2013)

$$V(\mathbf{z}) = \sum_{i=1}^{n} a(s_i) \sum_{j=1}^{m} b_j v_j(z^j(s_i)), \tag{4a}$$

$$V(\mathbf{z}) = \int_S a(s) \sum_{j=1}^{m} b_j v_j(z^j(s)) \, \mathrm{d}s. \tag{4b}$$

In this model, $\mathbf{z}$ represents a specified outcome where $\mathbf{z}(s) = [z^1(s), ..., z^m(s)]$ and m is the number of attributes. $\mathbf{z} \in \mathbf{Z}$ where $\mathbf{Z}$ represents the set of outcomes. Consequence in the location $s \in S$ for the jth attribute is determined by a function $z^j(s)$ and a is a function so that $a(s)$ is a spatial weight in location $s \in S$. In the discrete model, s_i is a location on subregion i. v_j is a single-attribute value function for the jth attribute and b_j is a positive weight for the jth attribute so that $\sum_j b_j = 1$.

2.4 Dominance

Complete preference information is discussed when the DM can provide exact spatial, attribute and temporal weights. However, there exist many situations when preference information is incomplete. For example, the DM can be a group of people who do not agree on preferences, the knowledge or experience of the subject can be imperfect, the DM has not made up her mind or she cannot for some other reason state her preferences, i.e., give exact weights (e.g., Weber (1987); Öztürké et al. (2005)). In that case a set of weights A consists of all the weights a that are feasible within the constraints the DM is able to provide. Methods to work with incomplete preference information are discussed in Section 3.

Even with incomplete preference information, one alternative can sometimes be identified as superior to another through dominance. Two different types of dominance relations between alternatives are considered in this study, absolute dominance and pairwise dominance (Eisenführ et al., 2010). Pairwise dominance is always followed when absolute dominance exists. Absolute dominance has two conditions:

$$\begin{cases} \min_{a \in A} V(z) & \geq & \max_{a \in A} V(z'), \\ \max_{a \in A} V(z) & > & \min_{a \in A} V(z'), \end{cases} \tag{5}$$

where $a \in A$ is a spatial weight and A is the set of all feasible spatial weights within the region of interest.

For this relation to hold, the minimum value of z must be greater than or equal to the maximum value of z' and the maximum value of z must be greater than the minimum value of z'. If both of these conditions are met, z dominates z' absolutely.

Pairwise dominance has two conditions as well:

$$\begin{cases} V(z) \geq V(z') & \text{for all } a \in A, \\ V(z) > V(z') & \text{for some } a \in A. \end{cases} \tag{6}$$

These conditions state that the overall value of z is greater than or equal to that of z' for all feasible spatial weights $a \in A$ and strictly greater at least for one. If both of these conditions are met, there exists a pairwise dominance so that z dominates z' which can be denoted $z \; D_A \; z'$ (Eisenführ et al., 2010).

If an alternative is not better than another in any attribute but it is worse at least in one attribute, that alternative is dominated by the other. The specific form of the additive value function or the exact weights of objectives are not required; the alternatives that are dominated can be discarded.

If no other alternative dominates some alternative, that alternative is called non-dominated. The set of non-dominated alternatives can be formulated as

$$Z_{ND} = \{z^j \in Z | \nexists k \text{ such that } z^k \succ z^j\}.$$

Dominance depends on the set A as seen in Equations (5) and (6), and so Z_{ND} depends also on the set A. The set of non-dominated alternatives Z_{ND} contains all good decision recommendations because alternatives $z \notin Z_{ND}$ have at most as high value for all feasible weights as a non-dominated alternative $z \in Z_{ND}$ and strictly smaller value for some. If there exists an alternative that dominates all other alternatives, it is the most preferred alternative.

Absolute dominance relations can be found by comparing two alternatives' minimum and maximum values according to Equation (5). Conditions of pairwise dominance are presented in (6). The first condition does not hold if there is at least one $a \in A$ for which $V(z) < V(z')$ applies. The simplest way to verify the first condition is to find the minimum of $V(z) - V(z')$ and see whether it is non-negative. If it is, then also all the other are non-negative and the first condition holds.

The other condition holds only if there is at least one $a \in A$ for which $V(z) > V(z')$ applies. This can be verified by finding the maximum of $V(z) - V(z')$ and by examining whether it is positive. Now these two conditions can be written in the following form:

$$\begin{cases} \min\limits_{a \in A} \big(V(z) - V(z')\big) \geq 0, \\ \max\limits_{a \in A} \big(V(z) - V(z')\big) > 0. \end{cases} \tag{7}$$

If these two conditions hold, there exists a pairwise dominance so that $z \; D_A \; z'$ (Eisenführ et al., 2010). This approach can be used to narrow down the number of relevant alternatives since the non-dominated alternatives are preferred ones and dominated alternatives can be discarded (Salo and Hämäläinen, 2010).

Pairwise dominance is always used unless the computational efficiency is not sufficient. To compute absolute dominances, only the maximum and minimum values of every alternative must be known and compared, whereas for pairwise dominances, the values that give the maximum and minimum remainders must be computed for every pair. Hence, pairwise dominance is much more demanding to compute than absolute dominance.

3 Incomplete preference information in spatial models

The spatial preference model was introduced in the previous section, but more complex models are sometimes needed. In this section, the spatial preference model is expanded so that spatial weights can be attribute specific or there can be independent or attribute specific temporal weights. There are altogether eight different cases that are presented.

Models with incomplete preference information are also discussed in this section. The eight cases are considered without knowing spatial weights in every location, but instead knowing the total weights of given subregions. Finally, some cases are examined where no weights are known exactly; there are instead linear constraints for them.

3.1 Spatial value functions

First the additive value functions for eight different cases with single or multiple attributes, independent or attribute specific spatial weights and independent or attribute specific temporal weights are introduced. Cases (8a) and (8b) are the same that were introduced in Section 3.1 and the rest are expansions of these (Harju et al., 2016):

$$V(z) = \sum_{i=1}^{n} a_i(s_i) v(z(s_i)), \tag{8a}$$

$$V(z) = \sum_{i=1}^{n} \sum_{j=i}^{m} a_i(s_i) b_j v_j(z^j(s_i)), \tag{8b}$$

$$V(z) = \sum_{i=1}^{n} \sum_{j=i}^{m} a_i^j(s_i) b_j v_j(z^j(s_i)), \tag{8c}$$

$$V(z) = \sum_{i=1}^{n} \sum_{k=1}^{l} a_i(s_i) c_k v(z_k(s_i)), \tag{8d}$$

$$V(z) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{l} a_i(s_i) b_j c_k v_j(z_k^j(s_i)), \tag{8e}$$

$$V(z) = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} a_i(s_i) b_j c_k^j v_j(z_k^j(s_i)), \tag{8f}$$

$$V(z) = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} a_i^j(s_i) b_j c_k v_j(z_k^j(s_i)), \tag{8g}$$

$$V(z) = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} a_i^j(s_i) b_j c_k^j v_j(z_k^j(s_i)). \tag{8h}$$

Spatial value functions are denoted by $V(z)$ where $s_i \in S$ is a location within the region S, $a_i^j(s_i)$ is a spatial weight for a location s_i for the jth attribute, b_j is a weight for the jth attribute, c_k^j is a temporal weight for the jth attribute on the kth time instant and $v_j(z_k^j(s_i))$ is a single-attribute value function over the level of the jth attribute in s_i on the kth time instant.

Equations (8a–8h) are introduced as discretized based on the non-discrete models and all the different cases are separated. For example, case (8b) is a multiattribute value function as is case (8c), but in case (8c) there are attribute spesific spatial weights, whereas in case (8b) there are independent spatial weights. When weights are attribute specific, there are different attributes and the weights are defined separately for each of them.

In case (8d), there is a new variable, temporal weight. That means there are more than one time instant to observe and they are assigned different weights. Case (8h) is a complete model where there are multiple attributes and time instants and attribute specific spatial and temporal weights.

3.2 Regional spatial weights

So far it is assumed that the DM's preferences are known exactly so that spatial, attribute, and temporal weights are all given. In real-world situations that is difficult to achieve. Thus, decision making should be possible also with incomplete information. Methods of processing incomplete information in the field of preference programming have been studied (e.g., Weber (1987)) and Salo and Hämäläinen (2010) provides a review of these methods.

It is much more challenging to elicit the spatial weights a than attribute weights b or temporal weights c because a must be determined in every location. Hence, the first step of spatial decision analysis with incomplete preference information is to assume the exact spatial weights are not known. In this situation, the whole region is divided into smaller subregions and the spatial weight is given for each subregion without knowing how it is distributed inside the subregion. The division into subregions can be made as the DM sees suitable. The spatial weight of the subregion i is denoted by $\alpha_i = \sum_{s \in S_i} a(s)$, when the whole region of interest S is divided into n different subregions $S_i \subset S$, $i = 1, 2, ..., n$.

Without knowing the exact spatial weights, the exact values cannot be known either but minimum and maximum values can be found. Hence, an interval in which the exact value belongs to is known and with extreme values it is possible to compute dominances, as was introduced in Section 2.4, and find out the preferred alternatives. Next the minimum values of the same eight cases that were presented in Equations (8a–8h) are introduced but knowing only the regional spatial weights:

$$\min V(z) = \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} v(z(s)), \tag{9a}$$

$$\min V(z) = \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} \sum_{j=1}^{m} b_j v_j(z^j(s)), \tag{9b}$$

$$\min V(z) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i^j b_j \min_{s \in S_i} v_j(z^j(s)), \tag{9c}$$

$$\min V(z) = \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} \sum_{k=1}^{l} c_k v(z_k(s)), \tag{9d}$$

$$\min V(z) = \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} \sum_{j=1}^{m} \sum_{k=1}^{l} b_j c_k v_j(z_k^j(s)), \tag{9e}$$

$$\min V(z) = \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} \sum_{j=1}^{m} \sum_{k=1}^{l} b_j c_k^j v_j(z_k^j(s)), \tag{9f}$$

$$\min V(z) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i^j b_j \min_{s \in S_i} \sum_{k=1}^{l} c_k v_j(z_k^j(s)), \tag{9g}$$

$$\min V(z) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i^j b_j \min_{s \in S_i} \sum_{k=1}^{l} c_k^j v_j(z_k^j(s)). \tag{9h}$$

In Equations (9a–9h), the minimum values are presented as functions, where s defines the locations within the subregions $S_i \subset S$, α_i^j is a regional spatial weight of the jth attribute in the subregion S_i, b_j is a weight of jth attribute, and c_k^j is a temporal weight for the jth attribute on the kth time instant. Furthermore, $v_j(z_k^j(s))$ is a single-attribute value function over the level of the jth attribute in s on the kth time instant.

Maximum values $\max V(z)$ can be computed with similar equations as presented above, but by maximizing the function with respect to s instead of minimizing it.

3.3 Incomplete preference information

Regional spatial weight models that allow spatial weight information to be incomplete were introduced in the previous section. Situations where the DM cannot determine exact weights are considered next; regional spatial weights are not known precisely and neither are exact weights of different attributes or temporal weights. An example of the application of incomplete information on weights can be found in Kirkwood and Sarin (1985).

A method used in this study to overcome the problem of incomplete preference information is to use linear constraints to define sets of feasible weights. That means exact weights for the decision model don't have to be known; instead some linear constraints for spatial weights, some for temporal weights and some for the weights of different attributes can be given to find the preferred decisions.

When the exact weights are unknown, no exact values can be computed either, but with constraints minimum and maximum values can be found so an interval in which the value belongs to can be defined and dominance relations can be computed.

Without any constraints, the set of feasible weights, for example spatial weights, consists of all $\alpha \in A$ where $\alpha_i \geq 0$ and $\sum_i \alpha_i = 1$. Linear constraints restrict feasible weights α to a subset $A' \subset A$ so that $\alpha \in A'$. For example, the DM can value one subregion (e.g., S_1) more valuable than another (e.g., S_2), in which case the linear constraint is $\alpha_1 \geq \alpha_2$. Or the DM can give a maximum weight for the sum of two subregions, e.g., the sum of weights of subregions S_1 and S_2 is at the most 0.5, in which case the linear constraint is of the form $\alpha_1 + \alpha_2 \leq 0.5$.

3.3.1 Linear constraints for spatial weights

First, the case where there is a single attribute and a single time instant is considered. The region of interest S is divided into subregions $S_i \subset S, i = 1, ..., n$ and α_i represents the spatial weight of the subregion S_i. The values of $\alpha = [\alpha_1, ..., \alpha_n]$ are not known, only linear constraints for them are given. The minimum value can be computed by

$$\min V(z) = \min_{\alpha \in A} \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} v(z(s)), \tag{10}$$

where s defines the locations within the subregions $S_i \subset S$, α_i is a spatial weight within the subregion S_i such that $a \in A$ where A is a set of all feasible spatial weights and v is a single-attribute value function over the consequence $z(s)$.

The maximum value $\max V(z)$ can be computed with a similar equation as presented above, but by maximizing the function with respect to α and s instead of minimizing it.

There exist different methods to find the extreme values (see, e.g., Dantzig and Thapa (2006)). One method is to solve the linear programming (LP) problem for α. Another is to find the vertices of a bounded polyhedron defined by linear constraints, compute the values using these vertices and select the minimum and maximum values. Vertices of the weights do not depend on $v(z(s))$. Hence, the vertices must be computed only once even when multiple alternatives are considered.

3.3.2 Linear constraints for spatial and attribute weights

In addition to spatial weights, exact attribute and temporal weights can also be unknown. There is a single time instant but multiple attributes in Model (11), for which there exist linear constraints, and a single-attribute value function v_j is attribute specific:

$$\min V(z) = \min_{\alpha \in A, b \in B} \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} \sum_{j=1}^{m} b_j v_j(z^j(s)). \tag{11}$$

The minimum value with linear constraints for both $\alpha \in A$, where A is a set of all feasible spatial weights, and $b \in B$, where B is a set of all feasible attribute weights, can be computed with Equation (11) when spatial weights are independent. s defines the locations within the subregions $S_i \subset S$, α_i is a spatial weight within the subregion S_i, b_j is the weight of the jth attribute and v_j is a single-attribute value function over the consequence $z^j(s)$ for the jth attribute.

Model (12) presents the situation where there are multiple attributes and one time instant, like in Model (11), but the spatial weights are attribute specific:

$$\min V(z) = \min_{b \in B} \sum_{j=1}^{m} \min_{\alpha^j \in A^j} \sum_{i=1}^{n} \alpha_i^j b_j \min_{s \in S_i} v_j(z^j(s)). \tag{12}$$

In Model (12), α_i^j is a spatial weight within the subregion S_i for the jth attribute.

The maximum value $\max V(z)$ can be computed with a similar equations as presented above, but by maximizing the function with respect to α, b, and s instead of minimizing it.

There exist different methods to find extreme values, for example by combining the two methods introduced in Section 3.3.1. One possible solution to find the extreme values of case (11) is to find vertices for both α and b and find the combinations that produce the minimum and maximum values. However, if there exists a big number of subregions computing can be inefficient. Likely the more efficient way is to compute vertices for b, solve the LP problem for α with different vertices and find the minimum and maximum values.

Attribute specific spatial weights must be taken into account in case (12). One method to find the extreme values for it is to solve the LP problem for α_j for every $j = 1, 2, ..., m$ separately and then solve the LP problem for b.

3.3.3 Linear constraints for spatial and temporal weights

Model (13) presents the situation where there exist multiple time instants whose weights are not known but there is only one attribute. This case is computationally similar with case (11) and is of the form

$$\min V(z) = \min_{\alpha \in A, c \in C} \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} \sum_{k=1}^{l} c_k v(z_k(s)). \tag{13}$$

In Equation (13) minimum value can be computed with linear constraints for spatial weights $\alpha \in A$ and temporal weights $c \in C$, where C is a set of all feasible temporal weights. The locations within the subregions $S_i \subset S$ are defined by s, α_i is a spatial weight within the subregion S_i, c_k is the temporal weight of the kth time instant and v is a single-attribute value function over the consequence $z_k(s)$ of the kth time instant.

The maximum value $\max V(z)$ can be computed also in this case with a similar equation as presented above, but by maximizing the function with respect to α, c, and s instead of minimizing it.

The extreme values in this case (13) can be computed the same way than in the case (11). That means, by computing the vertices for α and c and find the combinations that produce the minimum and the maximum or to compute vertices only for c and then solve the LP problem for α.

3.3.4 Linear constraints for spatial, temporal, and attribute weights

In the case where there exist multiple attributes and time instants but there are no exact values for spatial, temporal or attribute weights, there exist a minimum value function of the form

$$\min V(z) = \min_{\alpha \in A, b \in B, c \in C} \sum_{i=1}^{n} \alpha_i \min_{s \in S_i} \sum_{j=1}^{m} \sum_{k=1}^{l} b_j c_k v_j(z_k^j(s)). \tag{14}$$

where the minimum value is computed with linear constraints for spatial weights $\alpha \in A$, attribute weights $b \in B$ and temporal weights $c \in C$. Spatial and temporal weights are independent. The locations within the subregions $S_i \subset S$ are defined by s, α_i is a spatial weight within the subregion S_i, b_j is the weight of the jth attribute and c_k is the temporal weight of the kth time instant. v_j is a single-attribute value function over the consequence $z_k^j(s)$ of the jth attribute and the kth time instant.

The maximum value $\max V(z)$ can be computed also in this case with a similar equation as presented above, but by maximizing the function with respect to α, b, c, and s instead of minimizing it.

Again, there are several ways to find the minimum and maximum values. One method is to find vertices for all α, b, and c, compute all the combinations and find the minimum and the maximum. With many subregions, this can be inefficient, so another way is to find the vertices for b and c and then solve the LP problem for α.

3.4 Tool for spatial decision analysis

As part of this assignment, a tool for computing the values and finding the set of non-dominated alternatives was implemented with MATLAB. All the models presented in Sections 3.1, 3.2 and 3.3 were implemented. This tool is used in the example analysis in Section 4.

The region S is assumed to be a rectangle which is comprised of squares. Models presented in Section 3.1 give an exact spatial value as an output, and as inputs they take spatial, temporal, and attribute weights and alternatives. The alternatives are given as $v(z(s))$ since the single-attribute value function v can be any function the DM finds suitable.

Models presented in Section 3.2 take as inputs alternatives as $v(z(s))$, a division into subregions, and weights, whereas models presented in Section 3.3 take as an input, in addition to alternatives as $v(z(s))$ and a division into subregions, linear constraints for the weights. With this information they can compute minimum and maximum values, but the actual use of these models is to compute dominance relations of alternatives with the extreme values as presented in Section 2.4. Hence, the output is a list of non-dominated alternatives or a graph that shows the dominance relations of the alternatives.

4 Example analysis – fire station locations in Espoo

The example in this section is developed to demonstrate the use of the models presented in Section 3 in a real-world decision problem. It illustrates the use of the GIS data and incomplete preference information in spatial decision analysis, utilizes dominance relations to find the preferred alternatives, and shows how refining the information has an effect on the results.

4.1 Decision problem and preference model

This example illustrates the problem of how to choose the location of three fire stations in Espoo so that the fire fighting capability is maximized. However, there is no unambiguous optimum so some trade-offs must be made. The best decision is found by following the DM's preferences and it is based on the data over the geographic region. Spatial decision analysis can take both of them into account and find preferred alternatives with incomplete information.

The station that responds to a fire is the closest of the three stations. However, if the closest station is busy, the second closest station is the one that responds. Sometimes neither of the two closest stations cannot respond and in this case it is the furthest station that responds. Because the three stations have different response times, the multiattribute model (8b) is used. In the example, it is assumed that the response time of a fire station is directly proportional to the distance between the station and the fire.

There are multiple procedures for assessing a single-attribute value function v (see, for example, Keeney and Raiffa (1993); Dyer (2005); Keeney (1982)), but they are not discussed in this study. In this example, when defining v, it is assumed the damage caused by a fire depends on the response time non-linearly. The more time it takes to respond to the fire, the more damage it causes. There is defined a maximum response time T, after which the fire fighters will arrive too late to make a difference. As long as the response time $z^j(s) < T$, the single-attribute value function v is defined as in Simon et al. (2013). With these assumptions, the value function is of the form

$$v_j(z^j(s)) = \frac{1 - e^{-3.86(1-\min\{1, \frac{z^j(s)}{T}\})}}{1 - e^{-3.86}}, \tag{15}$$

where $z^j(s)$ is a response time from the jth closest station to the location s and T is a maximum response time.

Weight of the jth attribute, b_j, represents the likelihood that fire fighters arrive from the jth closest station. The importances of different locations can be determined by $a(s)$. It is a spatial weight in location s and presents the importance of that location.

In order to use the discrete model, the data has to be discretized. This means that the region of interest has to be divided into a finite number of locations and the single-attribute value is computed in every location using Equation (15). In this case spatial weights a should also be given for every location in order to compute the value using Equation (8b).

However, it is very challenging for the DM to determine the weights of all locations, so the model with regional spatial weights (9b) is useful. The whole region is divided into smaller subregions and the spatial weights are not given to every location, but to every subregion. Values $v(z(s))$ are still computed for every location.

If the DM is not able to state a spatial weight for every subregion or exact weights of different attributes, model (11) can be used. With this model, the DM has to give only linear constraints for weights. She can, for example, evaluate some subregion more valuable than another based on population or important places nearby.

4.2 Initial preference information

In this example, nine possible locations, that can be seen in Figure 1, are considered for the fire stations around Espoo. All the three-station combinations of these nine candidates are analyzed in the example to eliminate poor alternatives and to find the best alternatives.

In this example, the map of Espoo is divided into $67 \times 96 = 6432$ locations and the single-attribute value is computed in every location using Equation (15). T is defined to be a response time it takes from help to come a distance that is half of the maximum distance possible on the map in Figure 1, i.e., near to 20 km. An example of $v_j(z^j(z))$ can be seen in Figure 2. Observed station locations in the example are Otaniemi, Perusmäki, and Kauklahti, and the values are computed to the closest station when $j = 1$, to the second closest station when $j = 2$, and to the third closest station when $j = 3$.

For the regional spatial weights, the region is divided into nine subregions. Espoo is divided into eight subregions and the ninth subregion consists of the sea, the islands, and the areas that are not part of Espoo; their spatial value is defined to be zero. The subregions can be seen in Figure 1.

The DM gives linear constraints for the spatial weights α of the subregions $S_i, i = 1, 2, ..., 9$ as follows:

$$\begin{aligned} \alpha_1 &= 0, \\ \alpha_2 + \alpha_4 &\leq 0.08, \\ \alpha_3 &\leq 0.01, \\ \alpha_3 &\leq \alpha_4, \\ \alpha_6 &\geq 0.25, \\ \alpha_6 &\leq 0.4, \end{aligned}$$

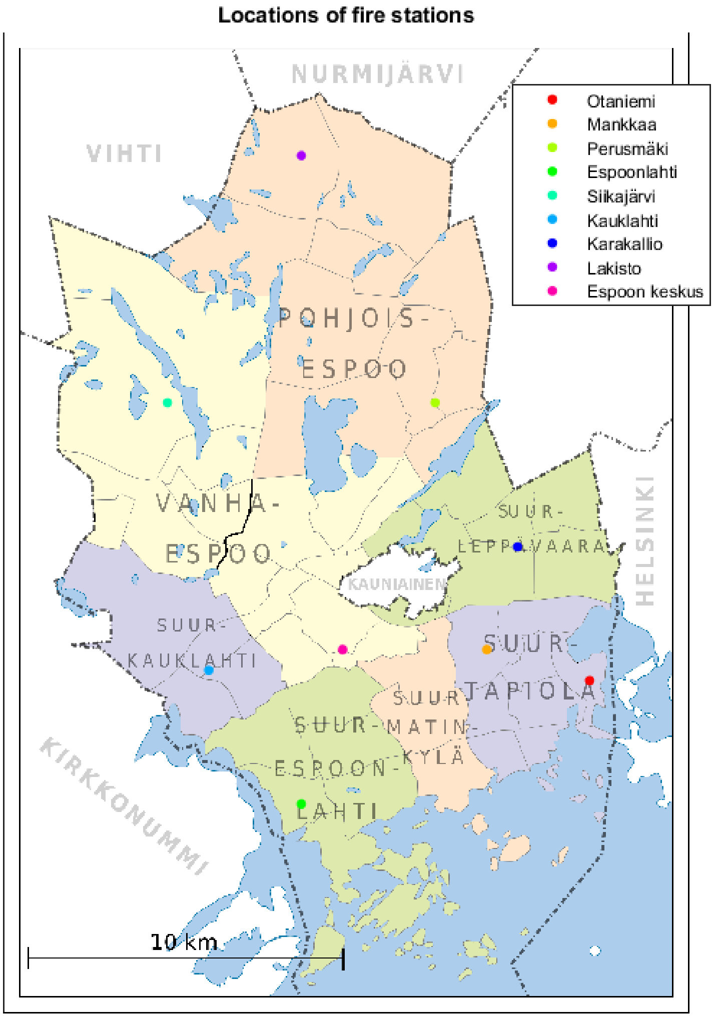

Figure 1: Potential fire station locations in Espoo and the division into subregions. Vanha-Espoo (the yellow region) is divided into two subregions with the black borderline.

$$
\begin{aligned}
\alpha_6 + \alpha_7 &\geq 0.5, \\
\alpha_6 &\geq 2\ \alpha_9, \\
\alpha_5 + \alpha_8 &\geq 0.15, \\
\alpha_6 &\geq \alpha_7, \\
\alpha_7 &\geq \alpha_5, \\
\alpha_4 + \alpha_5 &\geq 0.11, \\
\alpha_2 &\geq \alpha_4, \\
\alpha_2 + \alpha_4 &\leq \alpha_8, \\
\alpha_9 &\leq \alpha_7,
\end{aligned}
$$

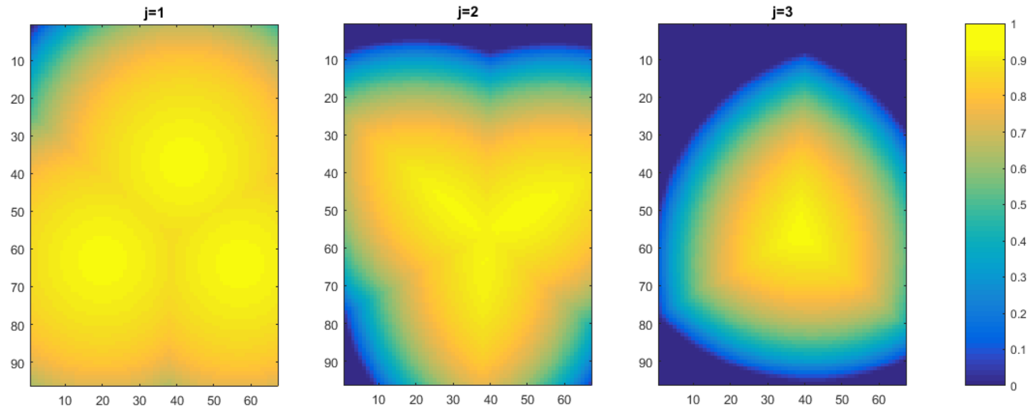

Figure 2: A single-attribute value function $v_j(z^j(s))$ to the jth closest station, $j = 1, 2, 3$, when the observed station locations are Otaniemi, Perusmäki, and Kauklahti.

where α_i is the spatial weight of subregion S_i so that the subregions shown in Figure 1 are numbered as shown in Table 1.

Table 1: Subregions $S_i, i = 1, 2, ..., 9$

Subregion	Area
S_1	The sea, the islands and the areas outside Espoo
S_2	Pohjois-Espoo
S_3	Western side of Vanha-Espoo
S_4	Suur-Kauklahti
S_5	Eastern side of Vanha-Espoo
S_6	Suur-Leppävaara
S_7	Suur-Tapiola
S_8	Suur-Matinkylä
S_9	Suur-Espoonlahti

The first of these constraints defines the spatial weight of S_1 to be zero and others determine the weights of different subregions of Espoo. For example, the spatial weight of Suur-Leppävaara (α_6) is defined to be between 0.25 and 0.4, bigger than the weight of Suur-Tapiola (α_7) and at least twice as big as the weight of Suur-Espoonlahti (α_9).

The DM weights subregions on the grounds of population, density (a fire spreads easily and causes more damage when population density is high), and important premises. Therefore, for example, Suur-Leppävaara (S_6) is weighted quite valuable: it's the subregion with the biggest population, people live there quite densely and a lot of important premises are located there. By contrast, subregions Pohjois-Espoo and western Vanha-Espoo (S_2 and S_4) are not so valuable for the DM; their population and density are small and there are not much premises.

The DM gives the linear constraints for weights of different attributes b as follows:

$$\begin{aligned} b_1 &\geq 0.55, \\ b_1 + b_2 &\geq 0.8, \\ b_2 + b_3 &\geq 0.3, \\ b_3 &\geq 0.1, \\ b_1 &\geq b_2, \\ b_2 &\geq b_3, \end{aligned}$$

where b_j is the weight of the jth attribute, in practice the likelihood that fire fighters arrive from jth closest station. The fifth and sixth conditions state that the likelihood that fire fighters come from the nearest station is the biggest and from the second nearest station the second biggest. The first four conditions give linear constraints for the magnitude of the weights, for example the first one determines the likelihood that help comes from the closest station must be at least 0.55. Extreme values can be found using vertices as described in Section 3.3. With these constraints, b has five vertices as shown in Table 2.

Table 2: Vertices of b with given linear constraints

b_1	b_2	b_3
0.7	0.15	0.15
0.7	0.2	0.1
0.6	0.2	0.2
0.55	0.25	0.2
0.55	0.35	0.1

With these constraints for regional spatial weights α and attribute weights b, extreme values are computed with the implemented tool for all 84 different three-station combinations using Equation (11). Preferred alternatives are found by pairwise dominance using Equation (7). There are 43 dominated alternatives and 41 non-dominated alternatives, so a little more than half of the alternatives can be eliminated.

Figure 3 visualizes dominance relations of a set of alternatives. A graph of all 84 alternatives' dominance relations would be messy, so only 13 alternatives are shown in the figure: alternatives whose one station location is Otaniemi, one is Mankkaa or Perusjärvi and the last one chosen freely. Every number matches with one station, for example, '1' means that one of three stations is located in Otaniemi in that alternative. All station numbers are seen in Table 3.

An arrow is drawn from the alternative that dominates to the dominated alternative. For example, when observing this set of alternatives, alternatives '129', '139', '137', '123', '124', and '127' are non-dominated, whereas an alternative '128' is dominated by almost every other alternatives. Overlapping dominance relations are not shown, for example, '129' dominates '128' even though there is not an arrow from '129' straight to '128' but through '125'.

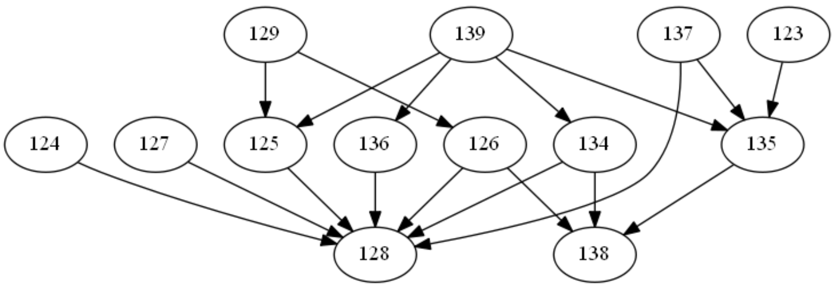

Figure 3: Graph of dominance relations of a set of alternatives.

There are still 41 non-dominated alternatives whose preferences over each others are not known. By analyzing the non-dominated alternatives, some conclusions can be made of which are good locations for the stations and which are not. The appearances of stations among non-dominated alternatives are found in Table 3.

For example, it can be seen from Table 3 that Lakisto is not a good station location because it only appears in dominated alternatives. Then again, Espoon keskus seems quite a good location for a station because it appears in 20 non-dominated alternatives.

Table 3: Number of times an examined station appears in non-dominated alternatives

Station number	Station	Number of appearances
1	Otaniemi	12
2	Mankkaa	19
3	Perusmäki	16
4	Espoonlahti	14
5	Siikajärvi	10
6	Kauklahti	14
7	Karakallio	18
8	Lakisto	0
9	Espoon keskus	20

4.3 Refined preference information

Many alternatives were eliminated with incomplete information. If fewer non-dominated alternatives are wanted, stricter constraints or exact weights are needed. Next it is assumed that the DM has received more information and is able to give exact weights for subregions and attributes that are the following:

$$\begin{aligned} \alpha &= \begin{bmatrix} 0 & 0.0352 & 0.0088 & 0.0308 & 0.0881 & 0.3524 & 0.2643 & 0.0881 & 0.1322 \end{bmatrix}, \\ b &= \begin{bmatrix} 0.6 & 0.25 & 0.15 \end{bmatrix}. \end{aligned}$$

The weights are scaled so that $\sum_i \alpha_i = 1$ and $\sum_j b_j = 1$. With these weights, Equation (9b) can be used to compute extreme values and then pairwise dominances can be found. By using this method, there exist 66 dominated alternatives and 18 non-dominated. So better results can be gained by specifying the preferences.

There are still 18 non-dominated alternatives whose relative order is not known. To gain more precise results some specifying should be made. One possibility is to divide the region into smaller subregions so the spatial weights could be more precisely defined.

Espoo is divided into 13 subregions and the one that consists of the sea, the islands, and the areas outside of Espoo; these 14 subregions can be seen in Figure 4. In addition to the smaller subregions, these 13 subregions of Espoo differ from the previous 8 subregions so that now the area inside a subregion is more homogeneous. For example, the density may previously have varied

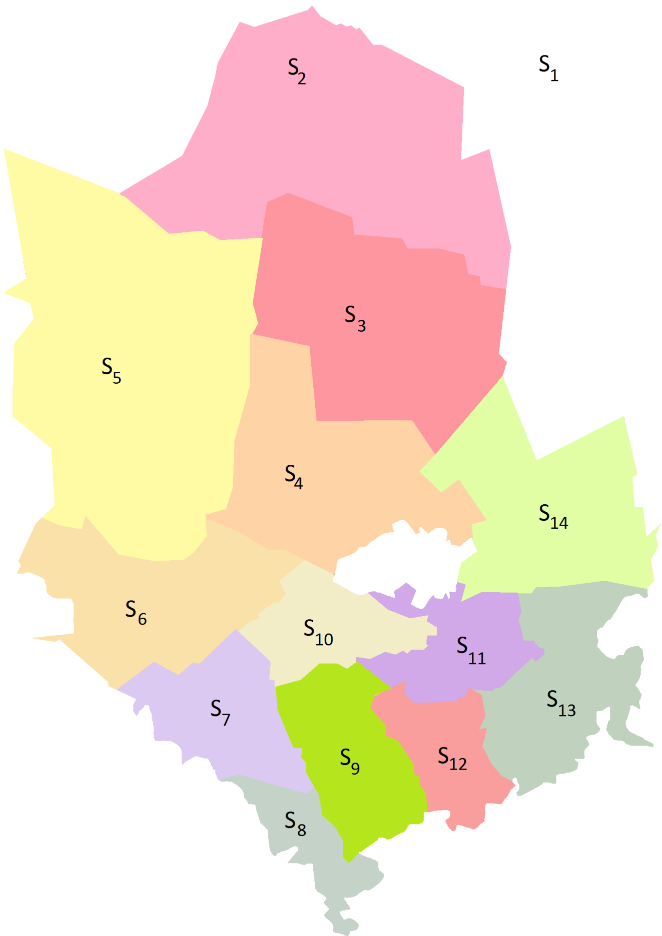

Figure 4: Espoo divided into 13 subregions and the one with the sea, the islands and the areas outside of Espoo.

a lot within a subregion while in the current division these types of varieties are striven to minimize.

It is assumed that the DM is able to give spatial weights for the subregions weighting by the same reasons than previously and the weights of different attributes have not changed. Equation (9b) can be used to compute the extreme values, but now α_i is the weight of subregion S_i so that $i = 1, 2, ..., 14$. The spatial weights are the following:

$$\alpha = [0 \;\; 0.0088 \;\; 0.022 \;\; 0.0264 \;\; 0.0088 \;\; 0.0308 \;\; 0.0308 \;\; 0.0661 \;\; 0.0441 \;\; 0.0661 \;\; 0.0352 \;\; 0.1101 \;\; 0.2203 \;\; 0.3304].$$

This method eliminates 75 dominated alternatives and leaves 9 non-dominated alternatives. The numbers of stations' appearances in non-dominated alternatives are shown in Table 4.

Table 4: Number of times an examined station appears in non-dominated alternatives

Station	Number of appearances
Otaniemi	3
Mankkaa	7
Perusmäki	3
Espoonlahti	1
Siikajärvi	0
Kauklahti	1
Karakallio	5
Lakisto	0
Espoon keskus	7

As can be seen in Table 4, Lakisto and Siikajärvi are not in any of the non-dominated alternatives so they can be discarded. Espoon keskus and Mankkaa are in seven non-dominated alternatives, so for both of them, there exist only two alternatives that they are not included in. Hence, they are good decision recommendations. All non-dominated three-station combinations are itemized in Table 5.

Table 5: Non-dominated alternatives

Station locations		
Otaniemi	Mankkaa	Karakallio
Otaniemi	Mankkaa	Espoon keskus
Otaniemi	Karakallio	Espoon keskus
Mankkaa	Perusmäki	Karakallio
Mankkaa	Perusmäki	Espoon keskus
Mankkaa	Espoonlahti	Espoon keskus
Mankkaa	Kauklahti	Espoon keskus
Mankkaa	Karakallio	Espoon keskus
Perusmäki	Karakallio	Espoon keskus

Non-dominated station locations concentrate near the subregions that are weighted the most valuable. At least two of Mankkaa, Espoon keskus, and

Karakallio should always be locations of fire stations in the three-station location combination. The third location can be a little further from the areas with biggest weights, for example in Otaniemi, Perusmäki or even in Espoonlahti or Kauklahti. However, Siikajärvi and Lakisto are too far for them to be good fire station locations. Most likely optimal fire station locations, when maximizing a geographically weighted value function, are Mankkaa, Espoon keskus, and a third one, which could be Otaniemi, Perusmäki, Espoonlahti, Kauklahti or Karakallio.

In this example, only one case of models presented in Sections 3.2 and 3.3 was demonstrated, that is the multiattribute model with one time instant and independent spatial weights. The other models work with the same principles.

5 Conclusions

After the basics of the spatial decision analysis, eight different cases for spatial preference models were presented in this study, first with exact weights and then with regional spatial weights. Next, five cases with incomplete preference information were discussed. To use these models in real-world decision problems, a tool for computing spatial values and for finding non-dominated alternatives was implemented. Lastly, an example to demonstrate presented models and to use the developed tool was constructed. The example illustrated the decision problem of finding the best locations for three fire stations in Espoo.

Spatial preference models are useful in the decision problems based on GIS data to eliminate alternatives and find the preferred decisions, as shown by the illustrative example in this study. An extensive decision analysis can be made because the models include also multiple attributes and time instants, division into subregions as the DM prefers and incomplete preference information.

As was shown, it is possible to eliminate many alternatives with incomplete preference information, but the more specified results are wanted, the more information must be given. Decision analysis can be conducted with big subregions and with linear constraints for weights, but by increasing the number of subregions, refining constraints of weights or giving exact weights, better results can be gained. This was demonstrated in the example; first big subregions and linear constraints eliminated a bit more than half of the alternatives. Then, exact weights narrowed down the number of relevant alternatives to almost one fifth of the original and finally the division into smaller subregions halved the count of remaining alternatives.

With incomplete preference information, the exact values for different alternatives cannot be obtained, but extreme values are found. This enables to eliminate relevant alternatives through dominance, which is fairly straightforward and efficient. The most preferred alternative is in the set of non-dominated alternatives. However, that set often consists of multiple alternatives whose relative order is unknown. The tool implemented for computing the extreme values and finding the non-dominated alternatives was used in the example and proved to be workable.

Preference models that allow information to be incomplete are useful in many decision problems and as the example shows, the models presented in this study can be applied to real-world decisions with spatially varying consequences. The capability of these models to work with multiple attributes is often necessary since many decisions demand consideration of many aspects.

The development of techniques and procedures in preference decision making have enabled the contributions of mathematical models on many decision problems. To use the full potential of decision analysis, for instance in companies, the development of decision analysis methods is required so that they are understandable and practical. Future research is needed especially in the GIS context, for the impact of the data over geographic region is often relevant but not taken into consideration. Another aspect, time, is not either much discussed in the field of decision making, even though many decisions could gain better results by considering multiple time instants.

References

Jim X Chen. Geographic information systems. *Computing in science & engineering*, 12(1):8–9, 2010.

George B Dantzig and Mukund N Thapa. *Linear programming 1: introduction*. Springer Science & Business Media, 2006.

James S Dyer. MAUT — multiattribute utility theory. In *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 265–292. Springer New York, 2005.

Franz Eisenführ, Thomas Langer, and Martin Weber. *Rational decision making*. Springer, 2010.

Simon French. *Decision theory: an introduction to the mathematics of rationality*. Ellis Horwood, 1988.

Michael F Goodchild. Geographic information systems. *Journal of Retailing*, 67(1):3–15, 1991.

Michael F Goodchild. Geographic information system. In *Encyclopedia of Database Systems*, pages 1231–1236. Springer, 2009.

Mikko Harju, Juuso Liesiö, and Kai Virtanen. Spatial preference models. Memorandum, 2016.

Ching-Lai Hwang and Kwangsun Yoon. *Multiple attribute decision making: methods and applications a state-of-the-art survey*, volume 186. Springer Science & Business Media, 2012.

Piotr Jankowski. Integrating geographical information systems and multiple criteria decision-making methods. *International journal of geographical information systems*, 9(3):251–273, 1995.

Piotr Jankowski. Integrating geographical information systems and multiple criteria decision-making methods: Ten years after. *Classics from IJGIS: Twenty years of the International Journal of Geographical Information Science and Systems*, pages 291–296, 2006.

Ralph L Keeney. Decision analysis: an overview. *Operations research*, 30(5): 803–838, 1982.

Ralph L Keeney and Howard Raiffa. *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.

Craig W Kirkwood and Rakesh K Sarin. Ranking with partial information: A method and an application. *Operations Research*, 33(1):38–48, 1985.

Jacek Malczewski and Claus Rinner. *Multicriteria decision analysis in geographic information science*. Springer, 2015.

Nancy J Obermeyer and Jeffrey K Pinto. *Managing geographic information systems*. Guilford Press, 2007.

Meltem Öztürké, Alexis Tsoukiàs, and Philippe Vincke. Preference modelling. In *Multiple criteria decision analysis: State of the art surveys*, pages 27–59. Springer, 2005.

Ahti Salo and Raimo P Hämäläinen. Preference programming–multicriteria weighting models under incomplete information. In *Handbook of Multicriteria Analysis*, pages 167–187. Springer, 2010.

Jay Simon, Craig W Kirkwood, and L Robin Keller. Decision analysis with geographically varying outcomes: Preference models and illustrative applications. *Operations Research*, 62(1):182–194, 2013.

Martin Weber. Decision making with incomplete information. *European Journal of Operational Research*, 28(1):44–57, 1987.