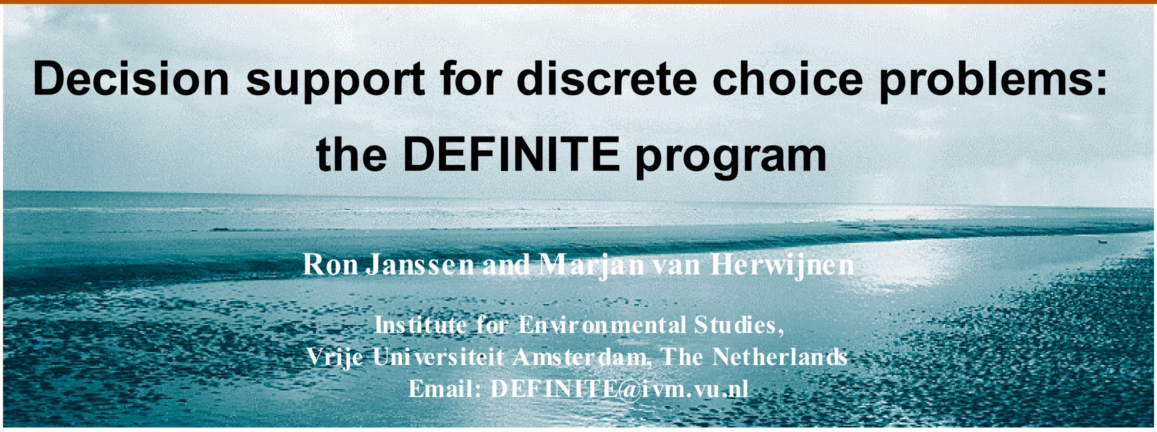


Decision support for discrete choice problems: the DEFINITE program

Ron Janssen and Marjan van Herwijnen

**Institute for Environmental Studies,
Vrije Universiteit Amsterdam, The Netherlands
Email: DEFINITE@ivm.vu.nl**

1 Introduction

Everybody makes decisions, many times a day. Most decisions come naturally, a well-trained reaction to familiar stimuli to which people apply habitual responses. Some decisions are a little harder, because they are not a routine business and have more important consequences. Buying a new car, changing job or leaving for an expensive holiday are decisions which are worth some attention. For these decisions, it seems obvious that we should gather information and ask people for advice before "making-up" our minds. This requires time, effort, and perhaps money. The resources allocated for the analysis of the decision depend on the magnitude of its consequences: choosing where to go on holiday is likely to be far less demanding than deciding in which country to settle for the next ten years.

Few decisions have a single objective. The very idea of making decisions suggests the need for considering multiple aspects and achieving a successful blend of performances. Decisions with multiple objectives are common in almost every private or public decision context:

- Choosing the means of transport to the office. Objectives may be the transport costs, the transport time, the safety of the trip, and so on.
- Constructing a new airport. Relevant concerns may be the construction and maintenance costs, the capacity of the airport, the access time, the safety of the system, the social disruption caused by locating the new facility and the effects of noise pollution.
- Selecting the best candidate for a job vacancy. The committee may be willing to consider the training, experience, communication and language skills etc.

These situations are strikingly different. Nevertheless, they share important similarities. First, individuals evaluate a set of alternatives, which represent the possible choices. The objectives to be achieved drive the design (or screening) of candidate alternatives and determine their overall evaluation. Attributes are the measurement rods for the objectives and specify the degree to which each alternative matches the objectives. Factual information and value judgements jointly establish the overall merits of each option and highlight the best solution.

DEFINITE (decisions on a finite set of alternatives) is a decision support software package that has been developed to improve the quality of decision-making. DEFINITE is, in fact, a whole toolkit of methods that can be used on a wide variety of problems. If you have a problem to solve, and you can identify alternative solutions, then DEFINITE can weigh up the alternatives for you and select the best alternative. The program contains a number of methods for supporting problem definition as well as graphical methods to support representation. To be able to deal with all types of information DEFINITE includes multicriteria methods, cost-benefit analysis and graphical evaluation methods. Related procedures such as weight assessment, standardization, discounting and a large variety of methods for sensitivity analysis are also available. A unique feature of DEFINITE is a procedure that systematically leads an expert through a number of rounds of an interactive assessment session and uses an optimisation approach to integrate all information provided by the experts to a full set of value functions. DEFINITE supports the whole decision process, from problem definition to report generation. The structured approach ensures that the decisions arrived at are systematic and consistent. DEFINITE can be used by the busy professional with no prior experience of such software, as well as the sophisticated user. A tutorial example and examples from the practice of environmental decision making are provided. Menus, information screens and help screens will lead you through the program and will very rapidly make you familiar with its features.

The first version, DEFINITE for MSDOS, appeared in 1994. A wide range of users has applied the program. Within the Dutch government users are almost all ministries, provinces, public bodies and a number of larger cities. Outside government the main users are consultancy and engineering firms. Finally, DEFINITE is used in universities and other schools of higher education for teaching purposes. DEFINITE for WINDOWS is built upon the concepts and ideas of DEFINITE for MSDOS and uses the experiences from all users of this version. However, the software is rewritten from scratch to produce a genuine WINDOWS program. The development of a WINDOWS program also provides the opportunity to include better graphics and new developments from the field of multicriteria decision-making. In summary, DEFINITE is a program that looks good, is easy to use and includes the latest theoretical developments.

Five case studies are used to demonstrate the potential of DEFINITE for windows. The first case study " A journey to Paris" provides an overview of a DEFINITE session. This session shows how a user proceeds trough the various steps in the evaluation. The other four case studies focus on one aspect of a session. The "Corridor study Amsterdam-Utrecht" shows the use of graphical evaluation. The case studies "Soil remediation Nieuwerkerk a/d IJssel" and "the Hondsbossche sea-wall" demonstrate the use of multicriteria analysis and cost-benefit analysis respectively. Finally, the case study Highway 73-South illustrates the potential of different types of presentation of the results.

2 From problem definition to report: A journey to Paris

The first case study concerns a journey to Paris. An Amsterdam couple has decided to spend the weekend in Paris and want to know the best way to travel from Amsterdam to Paris. This case is used to show how a user moves through the four evaluation steps: 1.Problem definition, 2.Multicriteria analysis, 3.Sensitivity analysis, and, 4.Report. Figure 1 shows part of the main menu including these four steps.

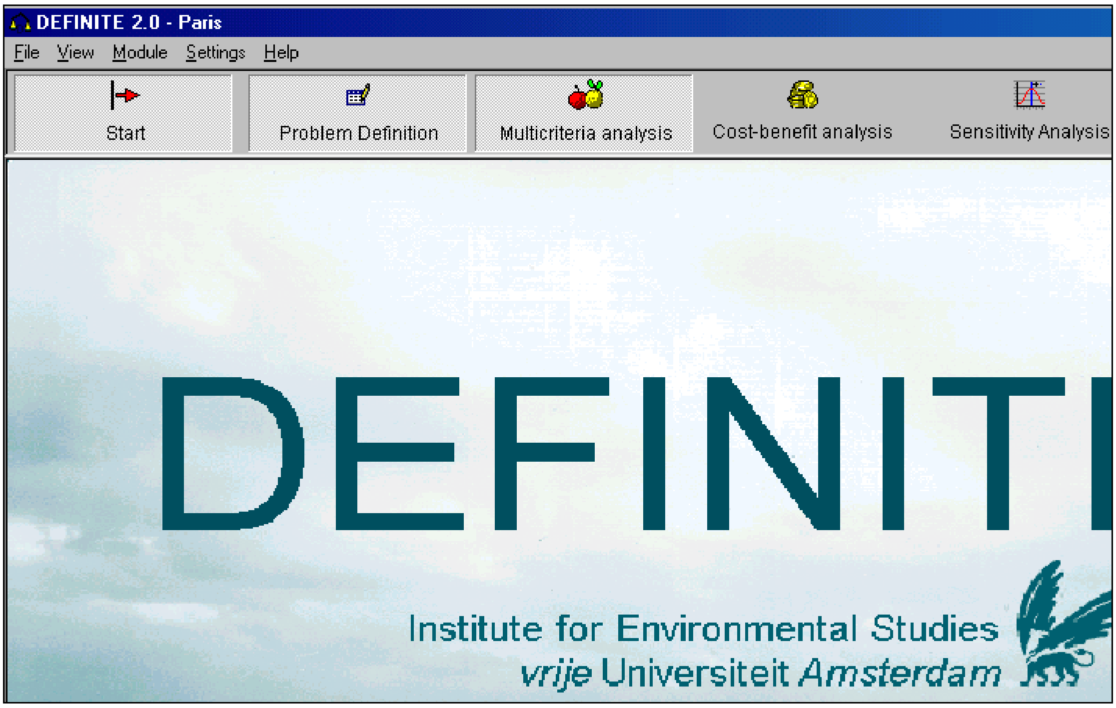


Figure 1. Main menu: Steps of an evaluation procedure.

Problem definition is the first step. The purpose of this step is to set up an effects table. The effects table of "A journey to Paris" is shown in Figure 2. The effects table includes five alternatives for the journey: 1.Car, 2.Bus, 3.Night train, 4.Thalys (high speed train) and 5. Aeroplane. These alternatives are evaluated according to five criteria: Comfort, Costs, Environment, Privacy and Travel time. Comfort and Privacy are measured on a ---/+++ scale; the other three on a quantitative scale. The second column indicates whether a criterion is a benefit criterion (the higher the better) or a cost criterion (the lower the better).

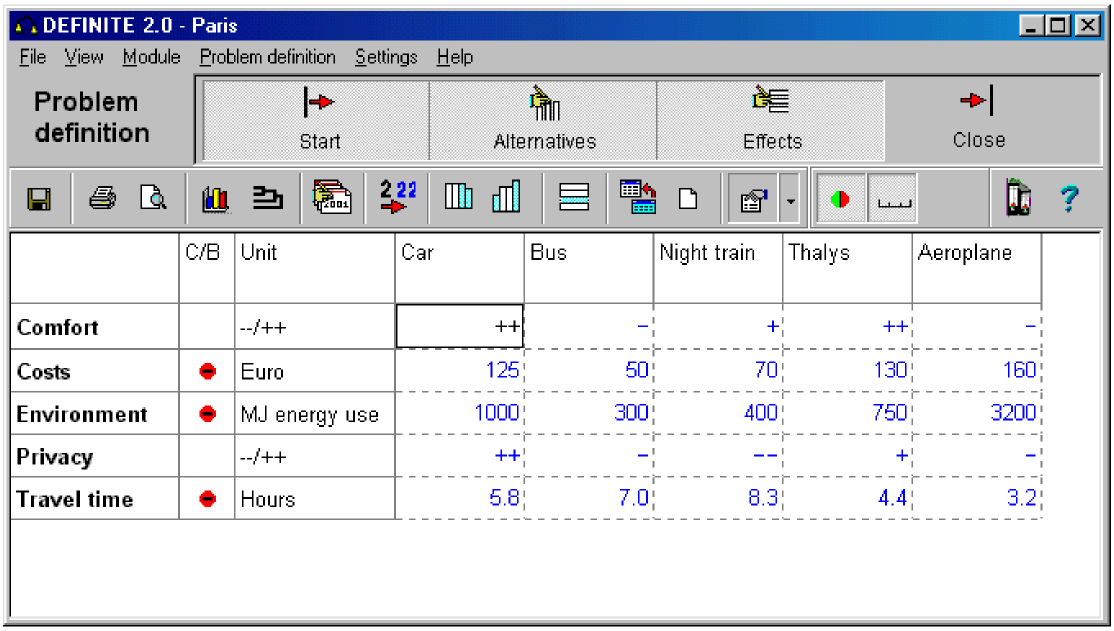


Figure 2. Problem definition: the effects table of "A journey to Paris".

Multicriteria analysis is the second step in this evaluation procedure. The purpose of this step is to derive a ranking of the alternatives. To do this the scores must first be standardized to make them comparable, and the relative importance, the weight of each score, must be established. Figure 3 shows the three steps of multicriteria analysis: 1.Standardize, 2.Weights and 3.Rank. In this case study all criteria are standardized using interval standardization and all criteria have equal weight. The columns minimum and maximum range show the lowest and highest score for each criterion. This information is important to assess the weights.

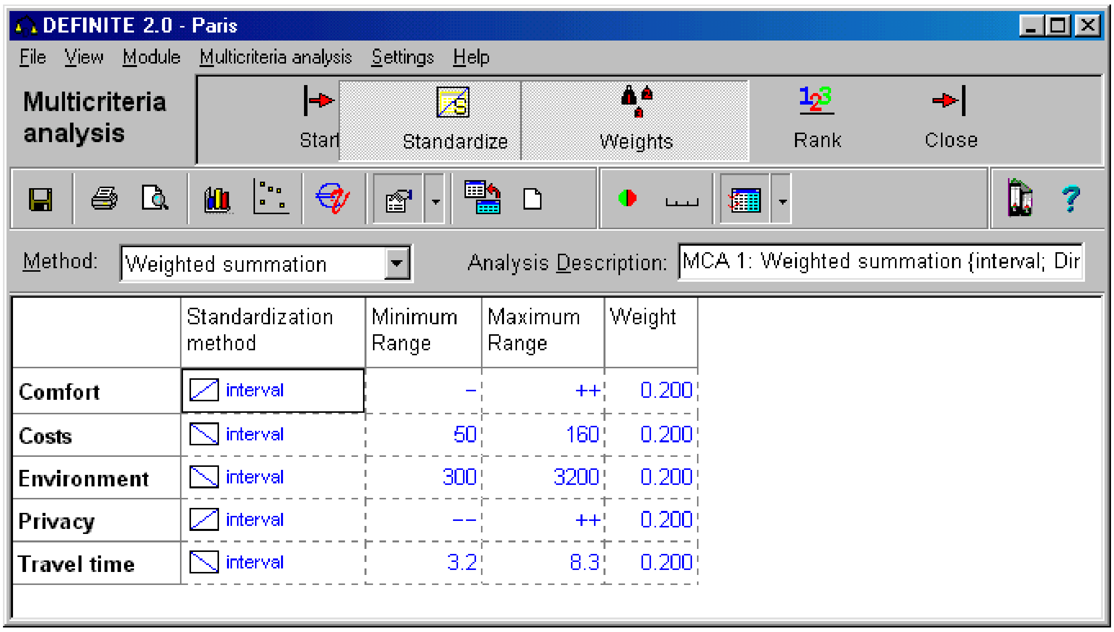


Figure 3. Multicriteria analysis: standardization and weights for "A journey to Paris".

In this example weights are set by direct assessment. It is also possible to use pair wise comparison or to provide an ordinal ranking of importance. Figure 3 shows that all criteria are standardized and all weights are set. This is all that is needed for the next step: Ranking.

Figure 4 shows that, if all criteria are equally important, the Thalys is the best alternative closely followed by Car. For both the Thalys and Car the performance on comfort contributes the most to the total score. The Aeroplane is by far the worst alternative. The total score of the Aeroplane is almost equal to its performance on Travel time. Weighted summation was used to calculate this ranking. Weighted summation is recommended by the Dutch committee on environmental impact assessment because it is well founded in theory and it is easy to explain to all participants in the procedure (Bonte et al. 1997). Other multicriteria methods available in DEFINITE are: the Electre 2 method, the Regime method and the Evamix method.

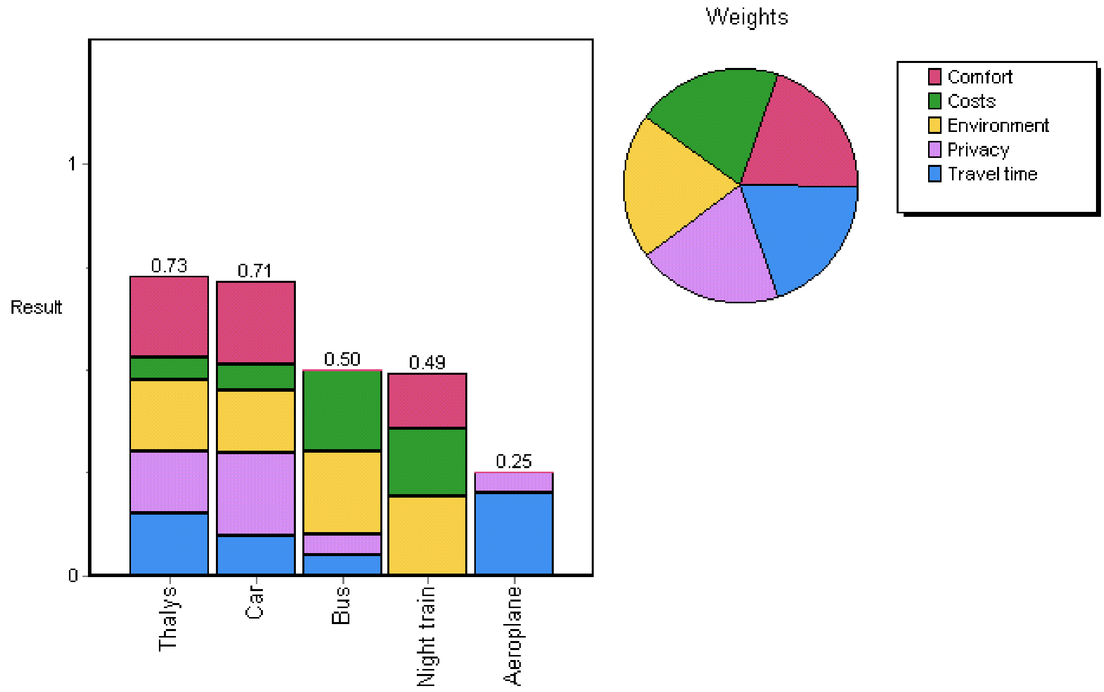


Figure 4. *Multicriteria analysis: ranking and total scores using weighted summation.*

The Thalys comes out best. The reliability of this result can be analysed in step 3: Sensitivity analysis. The ranking is dependent on the scores and weights. Changes in scores or weights may therefore result in changes in the ranking. The travel time of Car was estimated at 5.8 hours. If this estimate is correct the Thalys is the best alternative. The question whether the Thalys is still the best if travel time of the Car proves to be lower can be answered using Figure 5.

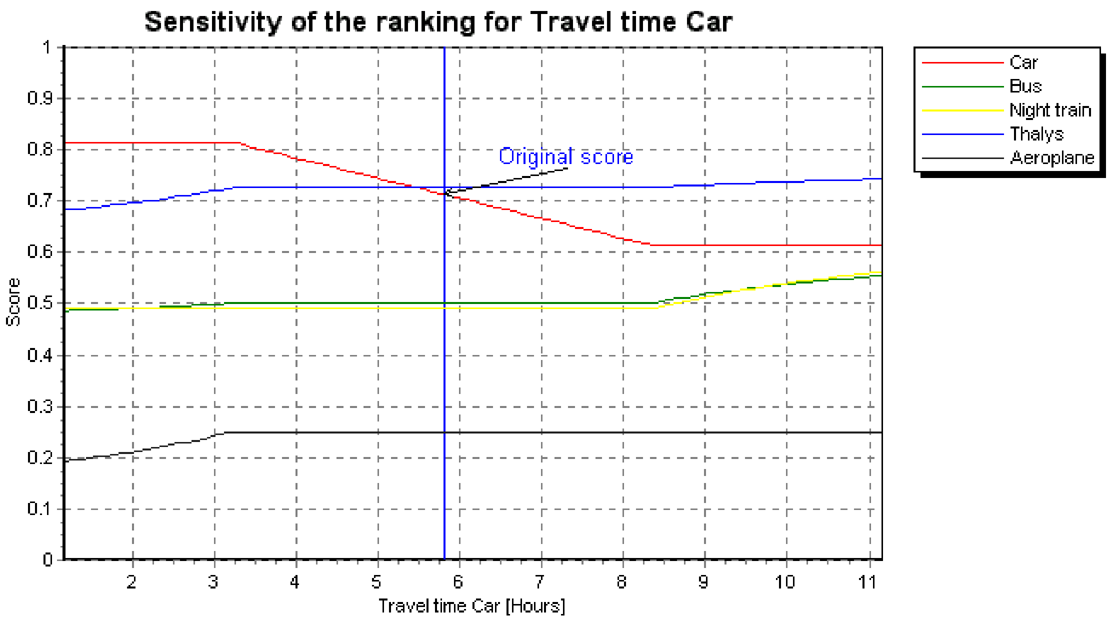


Figure 5. *Sensitivity analysis: sensitivity of the ranking for changes in travel time of the car.*

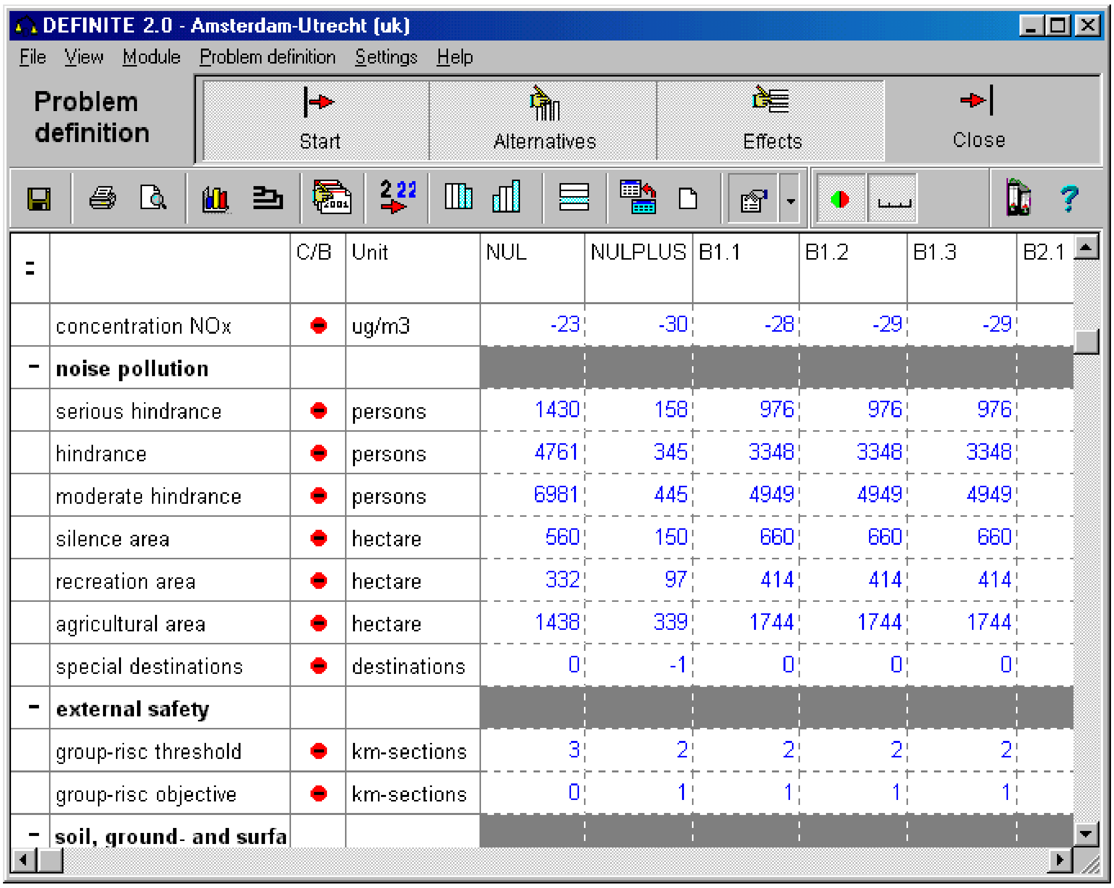
The original estimate, a travel time of 5.8 hours, is marked with a vertical line. At 5.8 hours the total score of the Thalys is just above the total score of Car. It is clear that the difference is only small and that a small reduction in the travel time of Car will make Car the better alternative. Reversal occurs at a travel time by Car of 5.5 hours. The ranking is shown to be very sensitive to changes in this score. The sensitivity of the ranking to changes in weights can be analysed in a similar manner. Monte Carlo analysis is available to analyse the sensitivity of the ranking to stochastic changes in scores and weights (see also section 5). The last step, Report, brings together all results in a standard evaluation report. This report includes all results specified by the user and can be edited using MS Word. In this example a quick tour through the four evaluation steps was provided. Many more tools are available to support each of these steps. These tools will be shown in the case studies presented below.

3 Graphical evaluation: Corridor study Amsterdam-Utrecht

In 1990 the Dutch Railway Company (NS) together with the Ministry of Transport initiated the Corridor study Amsterdam-Utrecht (CAU). The main reason for this project was serious capacity problems by road and by train between Amsterdam and Utrecht. Alternatives were designed to increase capacity in combination with measures to reduce or redirect mobility. For the highway A2, situated between Amsterdam and Utrecht, this resulted in the following nine alternatives.

- Nul autonomous development; car mobility increase by 70%
- Nulplus As Nul with minor adjustments; car mobility increases by about 35 %
- B1.1 improvement A2 (4+4); car mobility +35%
- B1.2 improvement A2 (2+2+2+2); car mobility +35%
- B1.3 improvement A2 (3+2+3); car mobility +35%
- B2.1 improvement A2 (5+5); car mobility +50%
- B2.2 improvement A2 (3+2+2+3); car mobility +50%
- O120.1 no improvement A2; car mobility +20%
- O120.2 improvement A2 (4+4); car mobility +20%

For these alternatives the effects on air, noise, safety, soil, ground and surface water, landscape, ecology and living environment are predicted. As a first step alternatives, criteria and scores are combined in the effects table. Part of this effects table is presented in Figure 6. This figure shows the expected noise pollution produced by the various alternatives. The column C/B shows that all effects on noise pollution are costs: the higher the score the worse the alternative. The effects table also shows that in all alternatives the number of people suffering serious hindrance increases.



		C/B	Unit	NUL	NULPLUS	B1.1	B1.2	B1.3	B2.1
	concentration NOx	●	ug/m3	-23	-30	-28	-29	-29	
-	**noise pollution**								
	serious hindrance	●	persons	1430	158	976	976	976	
	hindrance	●	persons	4761	345	3348	3348	3348	
	moderate hindrance	●	persons	6981	445	4949	4949	4949	
	silence area	●	hectare	560	150	660	660	660	
	recreation area	●	hectare	332	97	414	414	414	
	agricultural area	●	hectare	1438	339	1744	1744	1744	
	special destinations	●	destinations	0	-1	0	0	0	
-	**external safety**								
	group-risc threshold	●	km-sections	3	2	2	2	2	
	group-risc objective	●	km-sections	0	1	1	1	1	
-	**soil, ground- and surfa**								

Figure 6. Problem definition Noise pollution caused by the highway alternatives.

In the corridor study there were differences of opinion among stakeholders. The steering committee did not want these differences in the report. This implied that they did not allow weighting of the criteria nor the use of multicriteria analysis. A graphical presentation was used instead to support comparison of alternatives. As a first step scores were standardized between zero (worst alternative) and one (best alternative) and presented as a series of bar graphs (Figure 7).

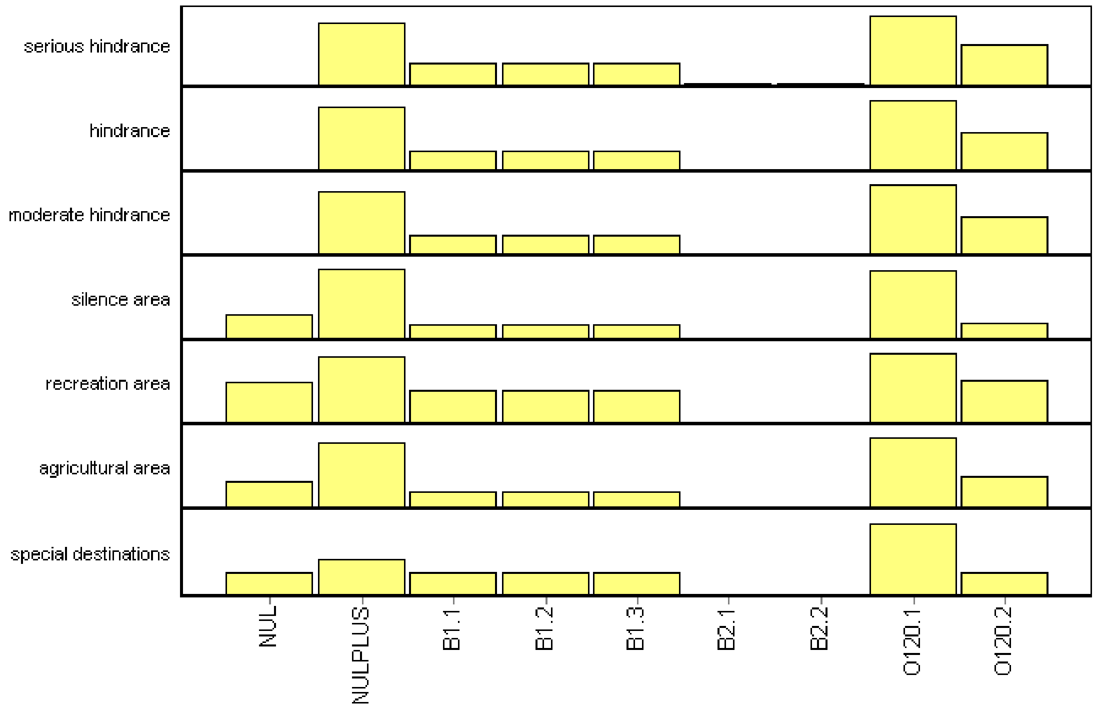


Figure 7. *Graphical presentation: Noise pollution caused by the highway alternatives (the highest bar represents the best alternative).*

By exchanging columns in such a way that the highest bars are moved to the left the alternatives can be ranked graphically (Figure 8).

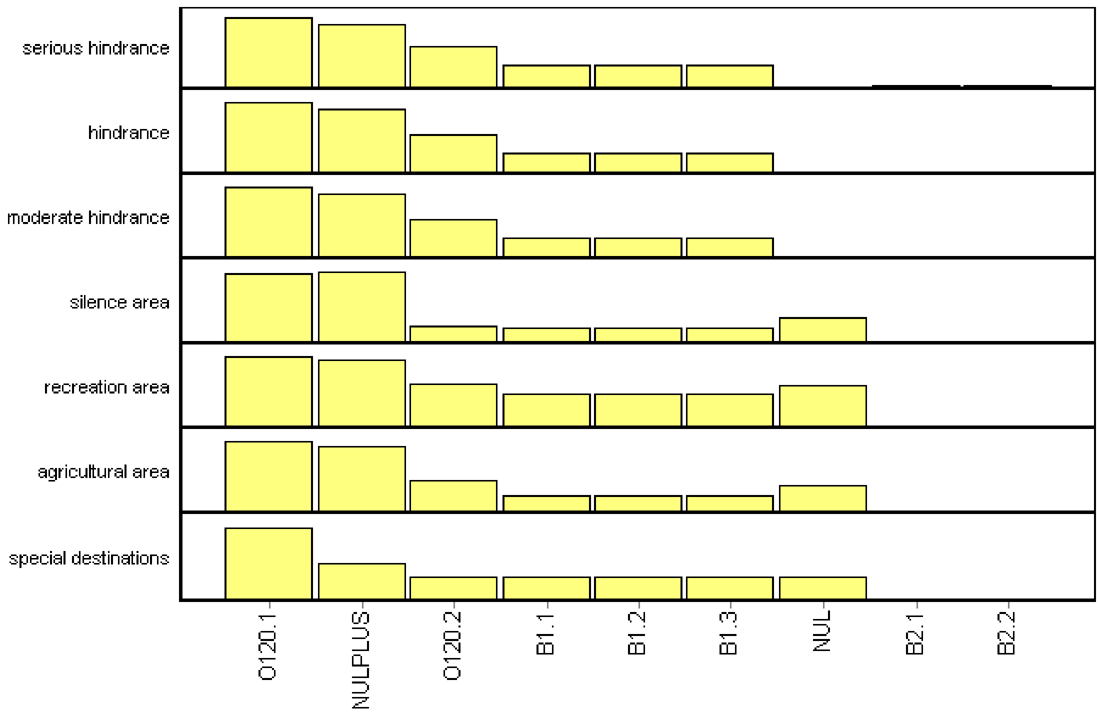


Figure 8. *Graphical ranking: Noise pollution caused by the highway alternatives (the highest bar represents the best alternative).*

On the left in the graphical presentation is alternative O120.1: the alternative with the smallest effects for noise hindrance. In second place is the Nulplus alternative, which is followed by O120.2. The alternatives B1.1, B1.2, B1.3 share fourth place followed by the Nul alternative. The alternatives B2.1 and B2.2 together occupy last place. The ranking of the B1 alternatives ahead of the Nul alternative implies that the first three effects are more important than the last three. Note that it is not necessary to weigh the various effects for noise hindrance to determine a ranking for all the other alternatives. The bar graph above clearly shows that the ranking of the remaining alternatives holds for all effects. Alternative O120.1, for example, is better than or equal to the Nulplus alternative for all effects.

Graphical presentation of all the categories was sufficient in this case to rank the alternatives. This had the advantage that no weights or prices were needed to rank the alternatives. In practice, graphical presentation is often sufficient to communicate the relative performance of alternatives. This is especially so, if a ranking per objective or per category is presented. In this example linear standardization was used to translate the scores to a value between zero and one. DEFINITE includes various types of linear and non-linear standardization methods. The following example illustrates the use of different types of standardization.

4 Multicriteria analysis: Soil remediation Nieuwerkerk a/d IJssel

A former industrial area near Rotterdam is heavily polluted with mineral oil and cadmium. It has to be cleaned up. There are two basic clean-up strategies and four alternatives. *On-site* strategies treat the soil on the site without excavation. Suitable alternatives are biological treatment and ventilation. *Off-site* strategies require firstly removal of the soil, and then treatment of the soil at a specialized plant. Suitable alternatives are thermal treatment and washing treatment. The objective of the clean-up is to eliminate, or reduce (to safe levels), the concentrations of cadmium and mineral oil.

The effects table includes scores measured in different measurement units. To make the scores comparable they need to be standardized to a common scale, for example between zero and one. In the previous sections interval standardization was used: the best score received a value of one, the worst zero and all others were scaled in between. Figure 9 shows the residual concentrations of cadmium after clean-up. In this case the scores are standardized between the concentration before clean-up (30 mg/kg) and the official policy target (0.8 mg/kg; Staatscourant 1998, 127). The ranges are not determined by the scores of the alternatives but by the extent to which the policy objective is reached. In DEFINITE this is called “goal standardization”.

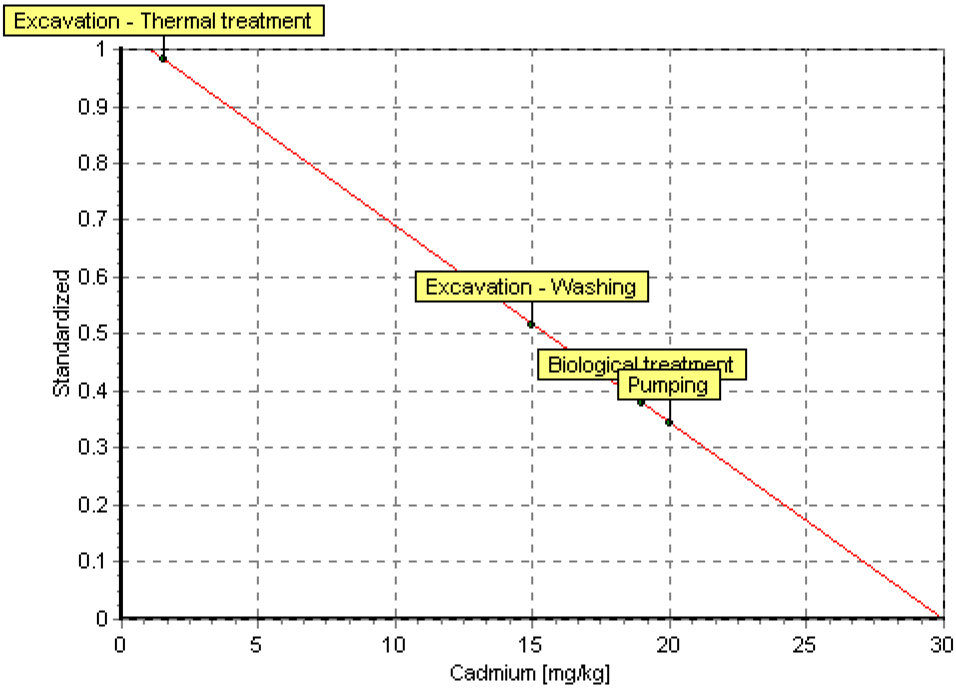


Figure 9. Linear standardization of the residual concentrations cadmium.

Although in practice the relation between a criterion score and its value (utility) is usually more complex, a linear standardization is often an acceptable approximation if the range of the scores is not too large. In those cases that a linear approximation is not acceptable, other, non-linear, standardization or value functions should be used. An example of a S-shaped value function is presented in Figure 10.

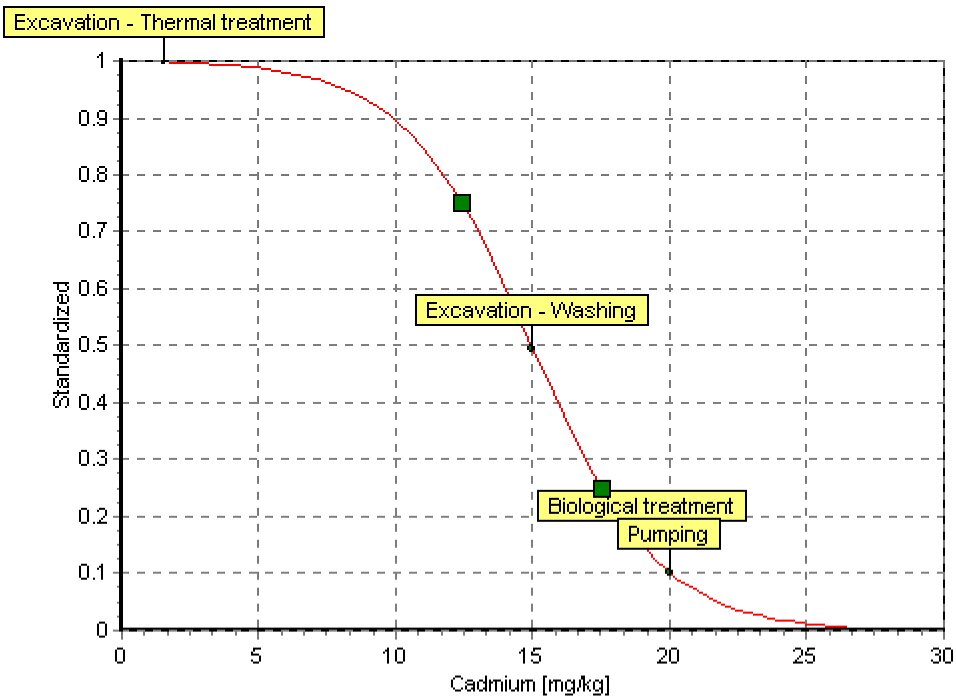


Figure 10. Standardization of the residual concentrations cadmium using an S-shape value function.

In Figure 10 the range of standardization is the same as in Figure 9: the scores are standardized between the concentration before clean-up (30 mg/kg) and the official policy target (0.8 mg/kg). However, the shape of the value function is in this case based on interviews with experts (Beinat 1997). By law, intervention of some sort is required, in this case, if the concentration is above 12 mg/kg (Staatscourant 1998, 127). Figure 10 shows that the function decreases sharply to the right of this score. The value of the alternatives extraction, biological treatment and ventilation is therefore low for this criterion. The shape of the curve is dependent on the type of criterion to be standardized. In addition to linear and S-shaped value functions DEFINITE includes convex, concave and free form value functions. An example of a concave value function is presented in Figure 11.

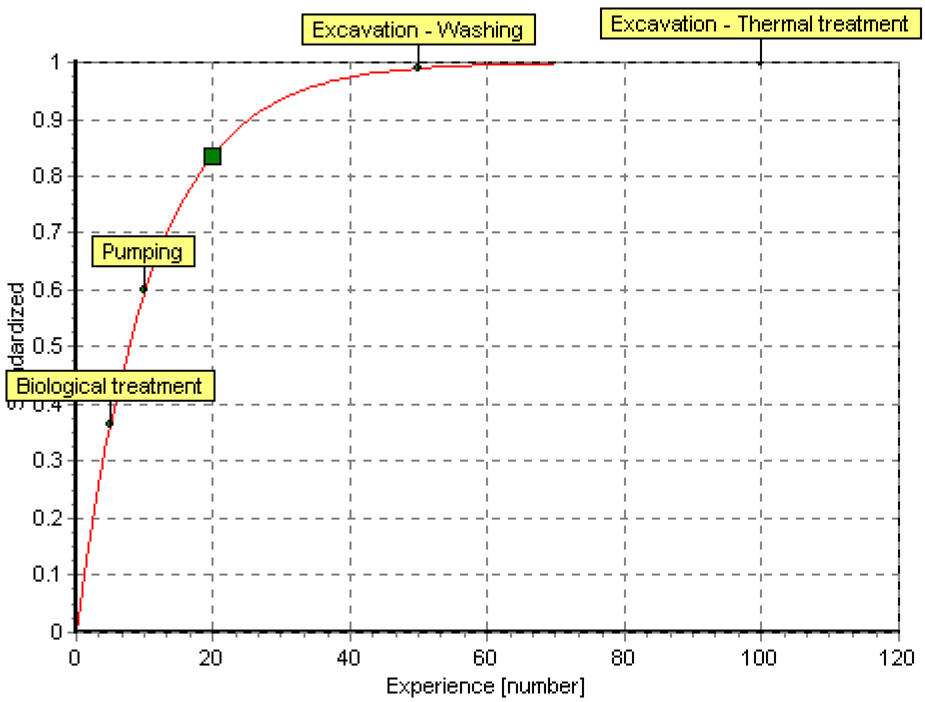


Figure 11. Concave standardization of experience measured as the number of times a technique has been applied.

Certainty that the required result is achieved is an important element in the selection of a clean-up technique. The choice of a new technique is therefore often considered as a risk and therefore experience with a technique is considered important. In Figure 11 experience is operationalized as the number of times a technique has been applied. A new technique receives a value of zero. The value function increases sharply and reaches a value of almost one for Extraction, a technique that has been applied 50 times. Above this score additional experience has little impact on its value.

Assessment of value functions and corresponding weights is a difficult task. DEFINITE includes the EValue procedure that can be used to assess these functions interactively (Beinat 1997, Stewart et al in prep.). This procedure is especially suitable to translate expert knowledge on the importance of effects into value functions and weights. A description of the EValue procedure can be found in the appendix. In the example presented in this section the value functions and weights linked to the various types of residual concentrations after clean-up were assessed in separate interview sessions with five experts. Assessment of non-linear value functions this way is rather time consuming and therefore expensive. This approach is most suitable for decision problems that occur several times in similar form, such as the clean-up of oil around petrol stations. If the ranges of scores are not too large linear standardization will be an acceptable approximation of reality in most cases. Standardization and assigning weights can be understood as valuation of scores in value or utility units. Another type of valuation is valuation in monetary units as presented in the next section on cost-benefit analysis.

5 Cost-benefit analysis: the Hondsbossche sea-wall

The “Hondsbossche” sea-wall is a potential bottle-neck in the coast defence of North Holland. In time, the risk of flooding will become an issue as a result of erosion and of the expected rise in lea levels. These problems have an expected urgency of 50 to 100 years. The main goal of the Ministry of Water and Transport is primarily: ’to maintain the safety against floods combined with preservation, if possible enlargement, of the natural area’.

Solutions can be found according to the principles of dynamic coast management (maintain the basic coastline with sand supplements) and through enlargement of the natural adaptability and recovery of natural processes. Three alternative solutions have been distinguished: Bastion, Dunes and Tidal gully (see Figure 12). In *Bastion* the present-day policy is continued, that means that the basic coastline is maintained according to the principles of dynamic coast management (sand supplements). In *Dunes* the sea-wall is demolished and replaced by a dune area. This alternative is primarily focussed on a better utilisation of the potentials for coast recreation. In *Tidal gully* the sea-wall is demolished without being replaced by an alternative sea dike. This alternative is focussed on maximal development of the natural values by enlarging the natural adaptability and offering more possibilities to recover the natural processes.

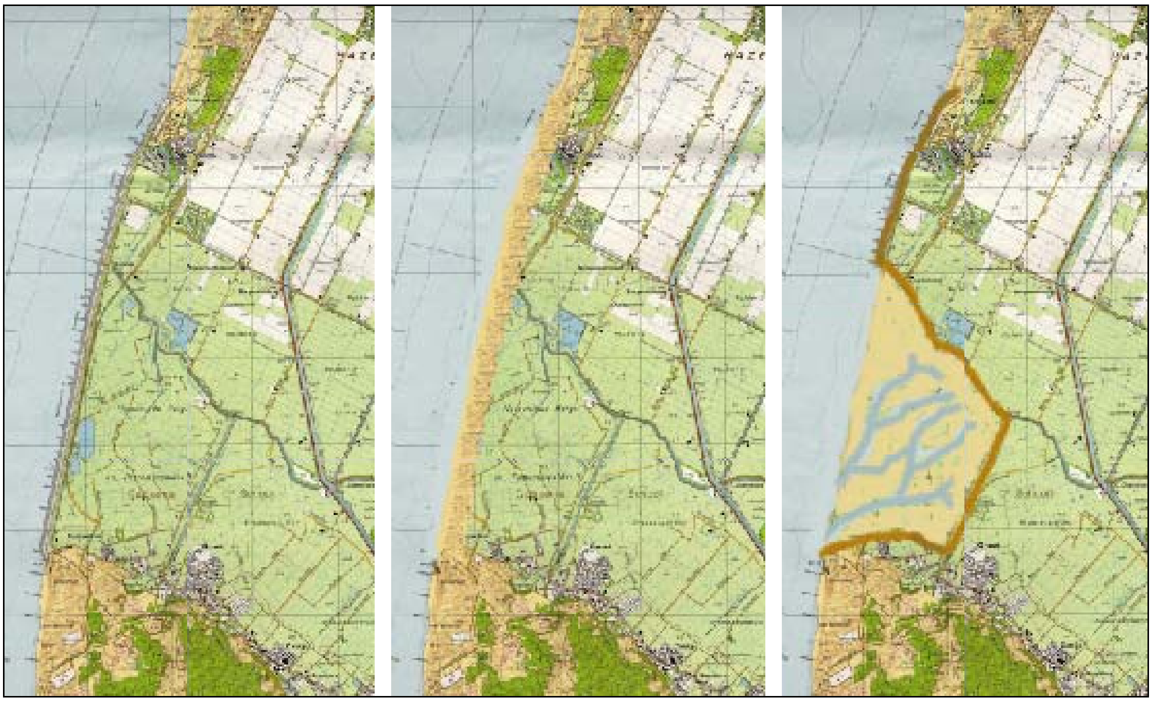


Figure 12. The three alternatives: Bastion, Dunes and Tidal gully.

The costs of these alternatives are investment and operational costs. Benefits are estimated for agriculture, recreation, fishery, biodiversity, fresh water storage and extraction, and, immovable properties. Benefits for the perception of safety and the prevention of damage are also calculated. All effects are studied for the period between 2000 and 2050. The cost-benefit analysis uses a discount rate of 4% and values all effects in monetary terms. An overview of the cost and benefits for the alternatives Dunes and Tidal gully are presented in the cost-benefit sheet in Figure 13. All effects are valued and therefore no pro memory effects are present. Dunes and Tidal gully differ with respect to: basic investment costs (Tidal gully better), benefits to agriculture, perception of safety and fresh water storage and winning (Dunes better) and benefits to fishery and biodiversity (Tidal gully better).

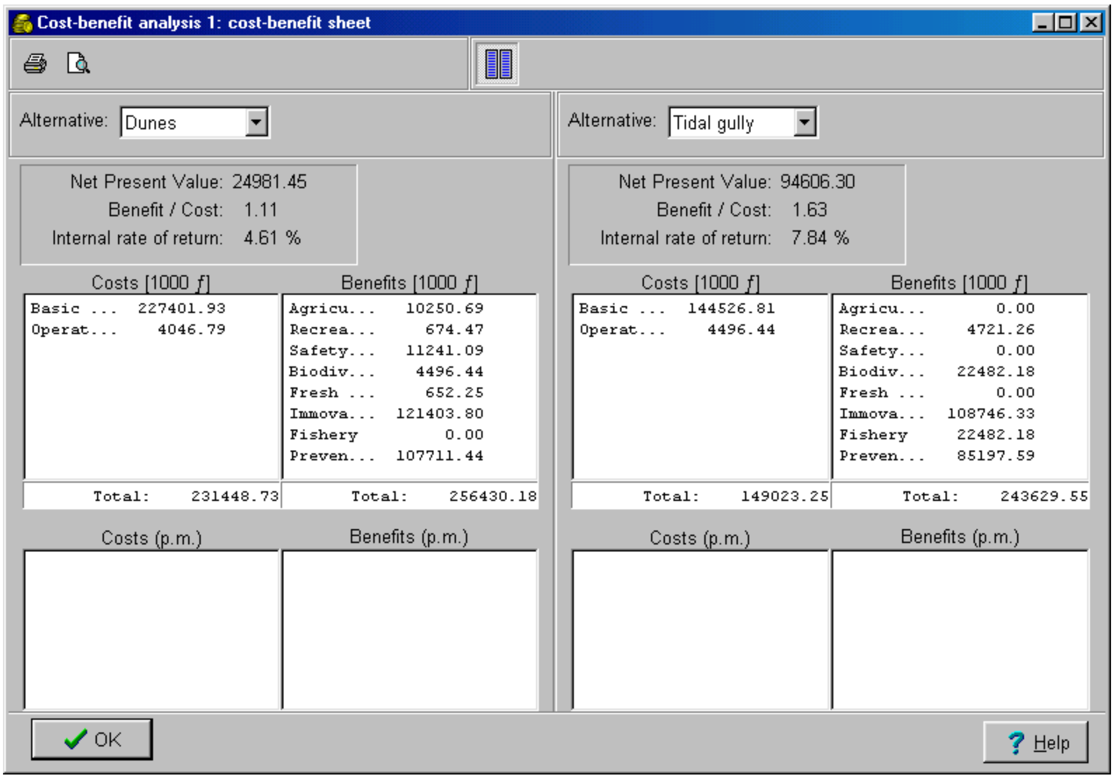


Figure 13. Cost-benefit sheet of the alternatives Dunes and Sea inlet.

The cost-benefit sheet also presents the three indices net present value, benefit-cost ratio and internal rate of return. All three indices show a better result for alternative Tidal gully. This result is based on total certainty of the predictions made for the expected cost and benefit. Because predictions are never totally certain, it is advised to test how robust the ranking of the three alternatives is for uncertainties in the used scores. Figure 14 shows an example of a result of sensitivity analysis on the ranking of alternatives based on the net present value.

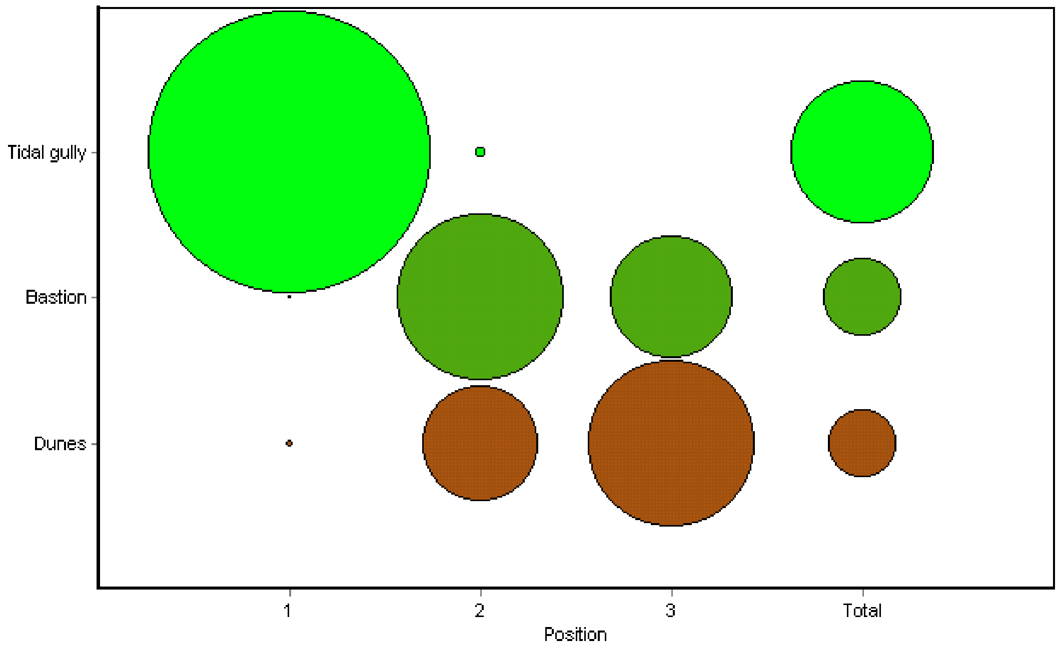


Figure 14. Uncertainty analysis: Ranking of the alternatives based on the net present value for an uncertainty in the assigned scores of 25 %.

To create Figure 14 the net present value of the three alternatives was calculated 2000 times, assuming that the scores can be 25% higher or lower than the assigned scores and that this deviation is normally distributed. The dependencies between the scores are taken into account (Herwijnen et al 1995). The figure shows that alternative Tidal gully is almost always ranked on the first position. Numbers associated with this figure (e.g. Tidal gully is found at first position 97% of the time) can also be derived. Alternative Bastion ranks with 57% on the second place and alternative Dunes with 58% on the third place. The difference between these two alternatives is so small that it is not quite certain that alternative Bastion ranks better than alternative Dunes. Further research can be done on the sensitivity of the ranking for the discount rate. Figure 15 shows the sensitivity of the ranking based on the benefit-cost ratio for a discount rate ranging from 0% to 25%.

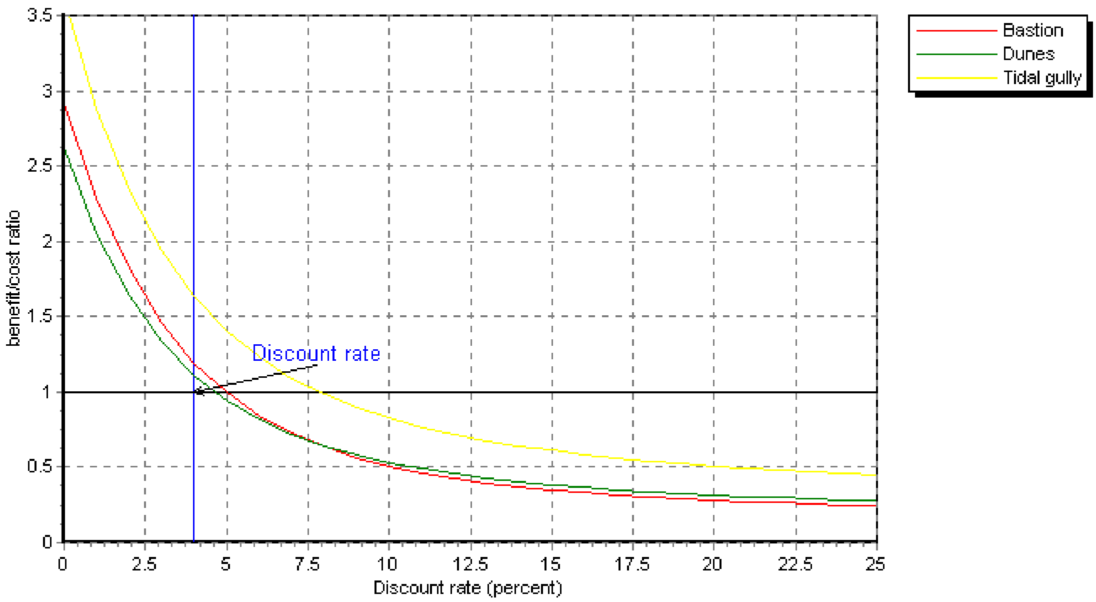


Figure 15. Uncertainty analysis: Sensitivity of the ranking of the alternatives based on the benefit-cost ratio for the discount rate.

The original discount rate, 4%, is marked with a vertical line. If the value of the discount rate increases, then the benefit-cost ratio of all alternatives decreases. The benefit-cost ratio of alternative Tidal gully is highest for all discount rates presented in the figure. The ranking of the two alternatives Dunes and Bastion, however, changes at a discount rate of about 7.5%. If the discount rate is higher than 7.5%, the alternative Dunes becomes better than Bastion, although the difference between these two alternatives remains very small. It can be concluded that alternative Tidal gully is the best alternative from an economic point of view and that there is little to distinguish between Dunes and Bastion. Further research is needed to clarify the preference between these two alternatives.

6 Presentation of results: Highway 73-south[1]

The area between Venlo and St Joost (see Figure 16) has problems with its living environment, accessibility and economic growth. Most of the problems are linked to insufficient capacity of the available infrastructure, an increase of car mobility, a decrease of the share of public transport and an increase of freight transport by road. After a long period of political debate a policy document was produced in preparation for a decision on new infrastructure. This policy document relied heavily on the results of multicriteria analysis (Heidemij advies 1995).

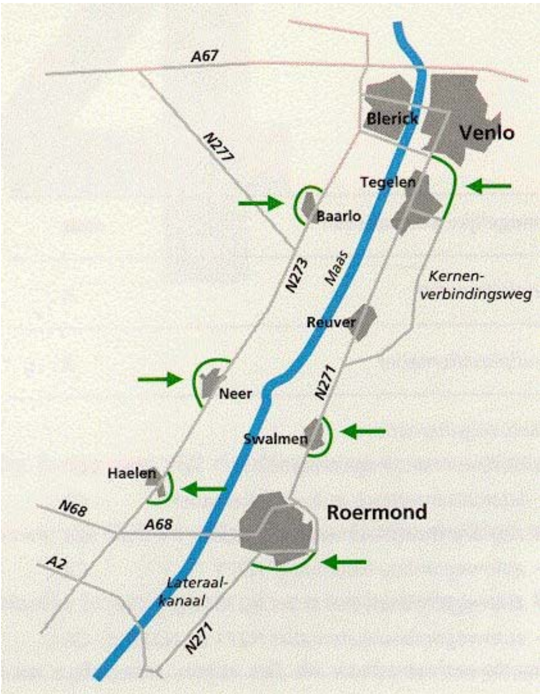


Figure 16. Principal roads in the plan area.

To solve these problems seven alternatives were developed. The Zero alternative represents the situation in 2010 without the construction of new infrastructure. The highway alternatives are new two lane roads with fly-overs and a maximum speed of 100 km/hr. The motorway alternatives are new 2x2 lane roads with fly-overs and a maximum speed of 120 km/hr. The environment alternative is a two lane highway designed in such a way that adverse environmental effects are minimised.

---

[1] The authors would like to thank Heidemij Advies BV. for using their study and figures. The full study can be found in Heidemij Advies, 1993.

Zero alternative
Highway alternatives -Highway east bank -Highway west bank -Highway east and west (upgrade N271 and N273)
Motorway alternatives -Motorway east bank -Motorway west bank
Environment alternative

A total of 73 effects are used to describe the performance of the alternatives. These effects are grouped into the following seven criterion groups: 1.Transport; 2.Soil and water; 3.Noise and air; 4.Flora, fauna and ecosystems; 5.Landscape, 6.Economy; and 7. Land use. The number of criteria in each group ranges from 2 for economy to 8 for flora, fauna and ecosystems. Figure 17 shows the objective tree of this example. The groups Economy and Soil and water are opened which means that the underlying criteria are visible. Economy includes two criteria: Access to the region and Employment. The icon in front of these criteria shows the measurement scale: a ratio scale for Access to the region and a -/+ scale for Employment. A tree structure like this is recommended if the effects table includes more than 10 criteria. The tree in this example has two levels. Any number of levels can be specified in DEFINITE. However, for ease of presentation two levels are recommended.

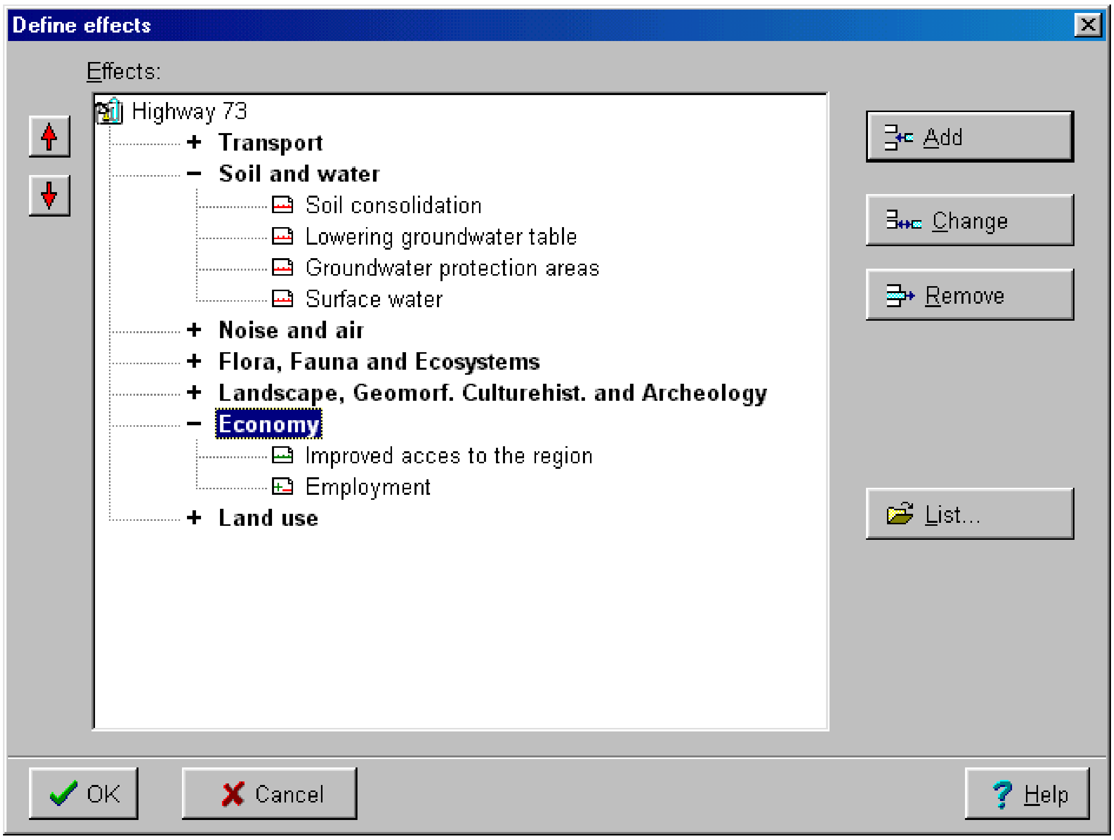


Figure 17. The criteria tree of Highway 73 South

The tree structure is used to assign weights to criteria within a group and to assess the weights between groups. Figure 18 shows the ranking of the alternatives. This ranking is generated using weighted summation. Experts have specified criterion weights.

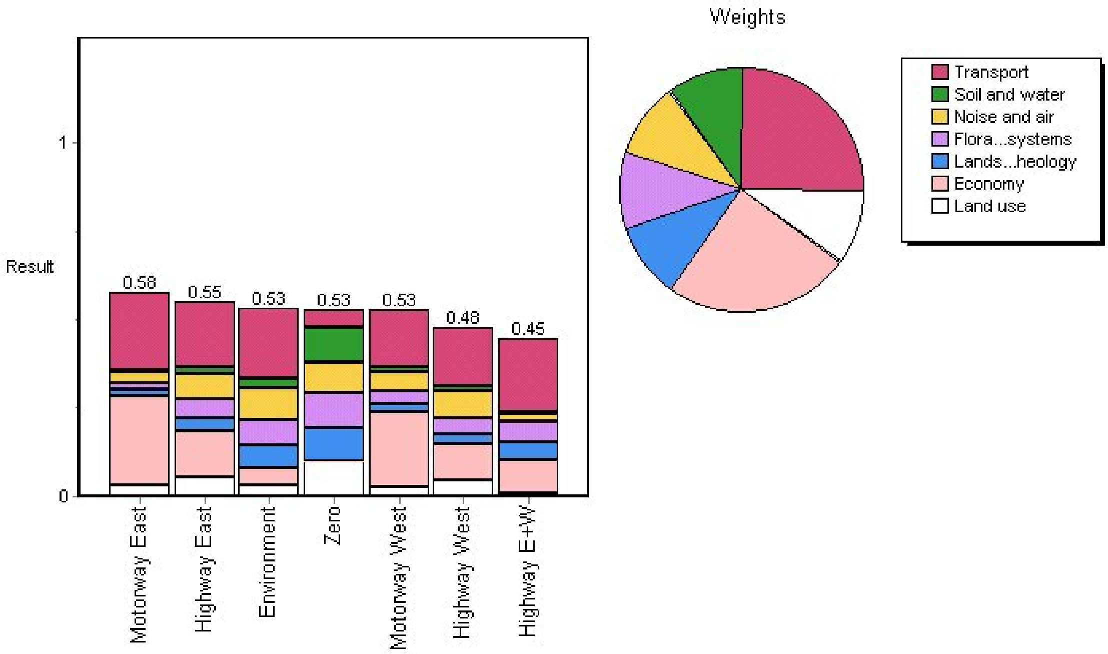


Figure 18. Total scores of the alternatives, the relative contributions of the criterion groups to these totals and group weights.

The pie chart shows that 25% of the weight is assigned to transport and also 25% to economy. The remaining 50% is distributed evenly among the five environmental criterion groups. Using these weights Motorway East is the preferred alternative followed by Highway East. The stacked bars show how the overall score is calculated as the sum of the weighted group scores. The graph shows that Economy and Transport are the groups that contribute the most to the overall score of Motorway East. The environmental groups contribute very little. This is in contrast to the Zero alternative where all groups contribute almost equally to the overall score. In addition to these graphs, DEFINITE has other graphical techniques available. For example a scatter diagram could be used to analyse further the conflict between economy and environment.

Assessment of weights is always to some extent subjective and political. This holds in particular for the group weights. The ranking presented in Figure 18 is calculated with 50% of the weights for economy and transport and 50% for the other groups. From the perspective of nature conservation groups, effects on the local natural areas should receive a higher weight. On the other hand from the perspective of the local population Noise should probably receive higher weight. DEFINITE has a procedure to explore systematically the relations between political perspectives and the ranking of the alternatives. To do this each perspective is translated into a set of weights. Figure 19 shows the result of this procedure. It is clear that thorough consultation with all stakeholders is required to specify these perspectives. The first row in Figure 19 reproduces the original ranking, in this example 50% for economy and transport and 50% for the other groups. The remaining rankings are linked to perspectives, each emphasizing a particular interest in the decision. In this example perspectives are linked to groups of effects. The second ranking is linked to a high weight (0.50) for Transport, the third to a high weight for Soil and Water etc. The procedure shows the influence of perspectives on the rankings. It is clear that the Zero alternative ranks best for the environmental perspectives. Not surprisingly, since most things stay the same. Exceptions are the economy perspective resulting in a first position for Motorway-east and Motorway-west and the transport perspective resulting in a first position for the environment alternative.

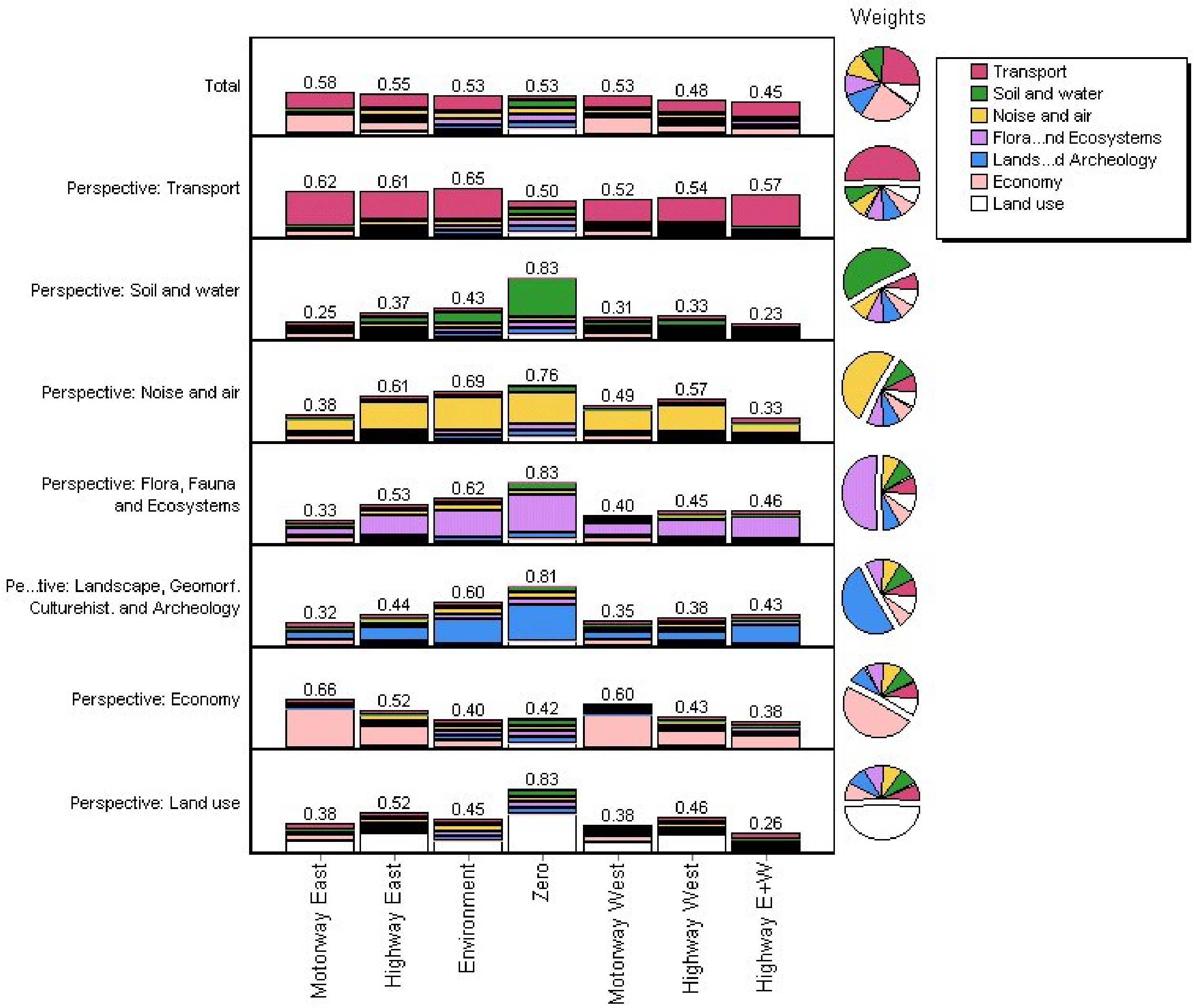


Figure 19. Total scores of the alternatives using various perspectives on group weights.

Figure 19 shows that the ranking is rather sensitive to changes in the group weights. All perspectives that emphasize one of the environmental criterion groups shift the Zero alternative to a clear first position. The Flora, Fauna and Ecosystems perspective, for example, has the Zero alternative on the first position with a total score of 0.83 and Motorway East even on the last position with a score of 0.33. After extensive debate the government decided for Motorway East in 1995. The motivation for this choice was linked to the economic effects. This is in line with the information presented in Figure 19.

7 In conclusion

Five examples were used to demonstrate the potential of the recently published DEFINITE program. Special features of the program are the graphical procedures, the procedure to assess value functions and the extensive procedures for sensitivity analysis. DEFINITE is not the only program available to support discrete choice problems. Other examples are Qualiflex (Stijnen and Smit 1996), TOPSYS (Wijnmalen 1997), Expert Choice (1999), VISA (Belton 1999), Rembrandt (Rog en Lootsma 1996), Decision explorer (Banxia) and DecisionPro (Vanguards). These programs are designed around one multicriteria technique. This has the advantage that it results in a relatively simple structure and a limited need for the user of the program to make choices. DEFINITE, on the contrary, is designed as a toolbox. The leading principle is that the choice of a method should be determined by the characteristics of the decision problem and by the preferences of the user. This results in a relatively large program with many options. To prevent users from getting lost in the program simple step menus have been designed. Using these menus the user is guided step by step through all the necessary procedures of the program.

DEFINITE is used by a large number and a large variety of users. These users appreciate the wide range of procedures offered. The improved user friendliness of the program and the additional functionality of the Windows version is intended to enlarge the group of users. This article offers a first introduction to DEFINITE. The best way to get to know the program is to experiment with the program itself. Hopefully this article will provide you with the incentive to make the next step.

References

Beinat, E. (1997) *Value functions for environmental management*, Dordrecht: Kluwer Academic Publishers.

Belton V. (1999) *Visual interactive sensitivity analysis for multicriteria decision making(VISA).* Visual thinking international, Glasgow

Bonte, R.J., J.v.d. Burg, R. Janssen, R.H.J. Mooren and J.T. de Smidt (1997) *Notitie over multi-criteria analyse in milieu-effectrapportage*, Utrecht: Commissie voor de milieu-effectrapportage.

CAU (1993). *Corridorstudie Amsterdam-Utrecht*. Rijkswaterstaat/Nederlandse Spoorwegen, Utrecht.

Expert Choice (1999). Expert choice inc., Pittsburgh.

Heidemij advies (1995) *Projectnota/MER Rijksweg 73-Zuid.* Rijkswaterstaat, Directie Limburg. Maastricht.

Herwijnen, M.van (1999) *Spatial decision support for environmental management,* Amsterdam: PhD dissertation, Vrije Universiteit.

Herwijnen, M., P. Rietveld, K. Thevenet and R. Tol (1995) *Sensitivity analysis with interdependent criteria for multi criteria decision making*. Multi Criteria Decision Making Volume 4 Nr 1, 57-70.

Janssen, R. (1992) *Multiobjective decision support for environmental management*, Dordrecht: Kluwer Academic Publishers.

Janssen, R., M. van Herwijnen and E. Beinat (2001). *DEFINITE for Windows. A system to support decisions on a finite set of alternatives (Software package and user manual)*. Institute for Environmental Studies (IVM), Vrije Universiteit, Amsterdam.

Janssen, R. and G. Munda (1999) Multicriteria methods for quantitative, qualitative and fuzzy evaluation problems. In: Bergh, J.C.J.M.v.d., (Ed.) *Handbook of Environmental and Resource Economics*, pp. 1000-1012. Edgar Elgar.

Stewart, T. J., E. Beinat and R. Janssen (in prep.), *Integrated value function assessments for impact assessments*, *EJOR*.

Stijnen, H. and H.J.Smit (1996). Qualiflex versie 2.3 een software pakket voor multicriteria analyse, Nederlands Economisch Instituut, Rotterdam.

Rog,L. and F.A.Lootsma (1996) Rembrandt program for multi-criteria decision analysis. Faculteit ITS, TU Delft.

Nijkamp, P., P. Rietveld and H. Voogd (1990) *Multicriteria evaluation in physical planning*, Amsterdam: North Holland.

Tufte, E. R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative,* Graphics press, Cheshire, Connecticut.

Werkgroep beleidsanalyse 0011 (1984). *Evaluatie Uniform Meldnummer 0011*, Ministerie van Financiën, Den Haag.

Wijnmalen, D. (1997). TOPSYStem. 3.0 voor Windows. TNO Fysisch en Electronisch Laboratorium, Den Haag.

Annex: The EVALUE procedure

A value function can be used to translate effect scores into a value scores. An example of a value function was shown in the soil remediation example. The function used to value the concentrations of Cadmium is shown again in Figure A1. In this figure the scores are standardized between the concentration before clean-up (30 mg/kg) and the official policy target (0.8 mg/kg). The shape of the value function is in this case based on interviews with experts (Beinat 1997). By law, intervention of some sort is required, in this case, if the concentration is above 12 mg/kg (Staatscourant 1998, 127). Figure A1 shows that the function decreases sharply to the right of this score. The value of the alternatives extraction, biological treatment and ventilation is therefore low for this criterion. The shape of the curve is dependent on the type of criterion to be standardized. In addition to linear and S-shaped value functions DEFINITE includes convex, concave and free form value functions. A DEFINITE user can assess a specific value function by specifying the range, selecting a shape and finally shape the function by click and drag.

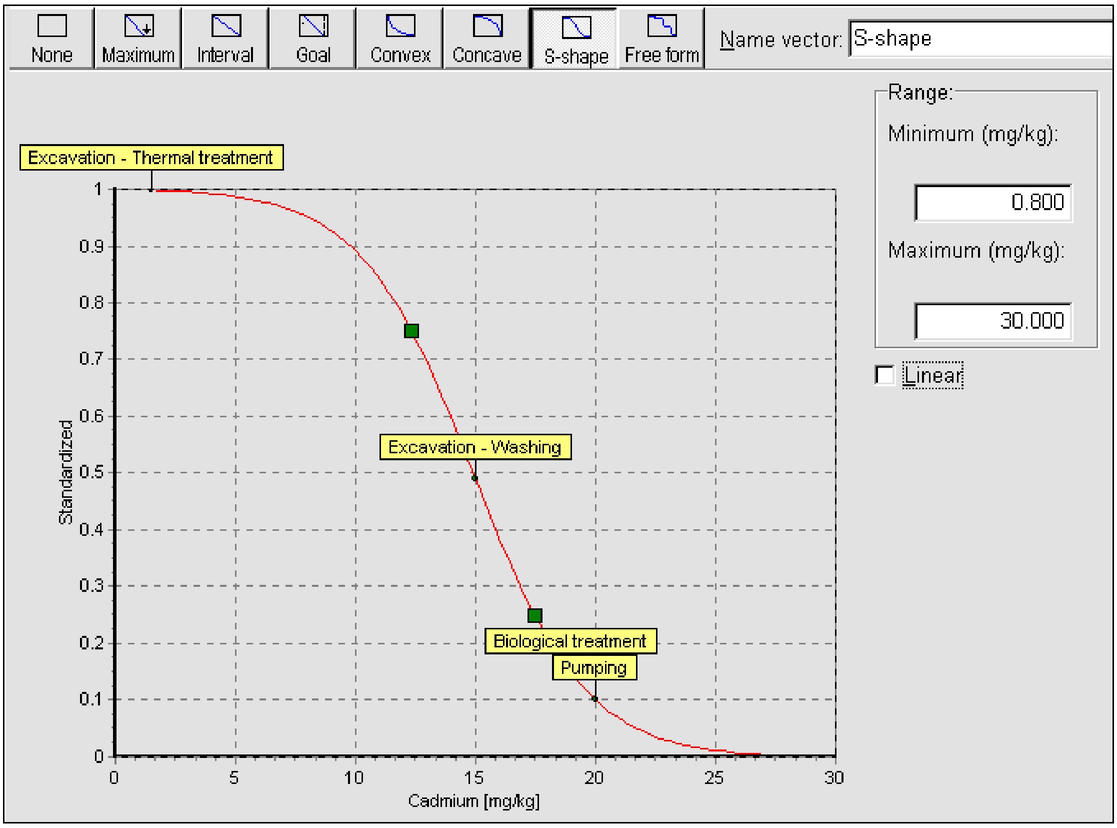


Figure A1 Direct assessment of a value function: 1.set the range, 2. select a shape and 3. click and drag.

In this example the evaluation is based on three criteria: the concentrations of Cadmium, Mineral oil and Zinc. Since weights and values are linked, value functions and weights for these three criteria are interdependent and have to be assessed simultaneously.

However, assessment of value functions and corresponding weights is a difficult task. DEFINITE includes the EValue procedure that can be used to assess these functions interactively (Beinat 1997, Stewart et al in prep.). This procedure is especially suitable to translate expert knowledge on the importance of effects into value functions and weights.

EValue requires three types of input:
1. A first estimate of the value functions;
2. Qualitative weights;
3. Rankings of test alternatives with two criteria (conjoint scaling).

An optimisation procedure is used to translate this information into a set of precise value functions and weights. The user is asked to evaluate the result and to suggest changes. It takes usually three or four rounds of inputs, calculations and feedback to reach a result that satisfies the user.

1. A first estimate of the value function
As a first step the user is asked to specify the range of the scores to be valued, the shape of the function (optional) and a value region (see Figure A2). In this example the range of scores is between 0.8 and 30 mg/kg, an S shaped function is selected and a relatively narrow value region has been defined. The clean-up alternatives are included to provide a reference. The value functions for Zinc and Mineral oil are assessed in a similar manner.

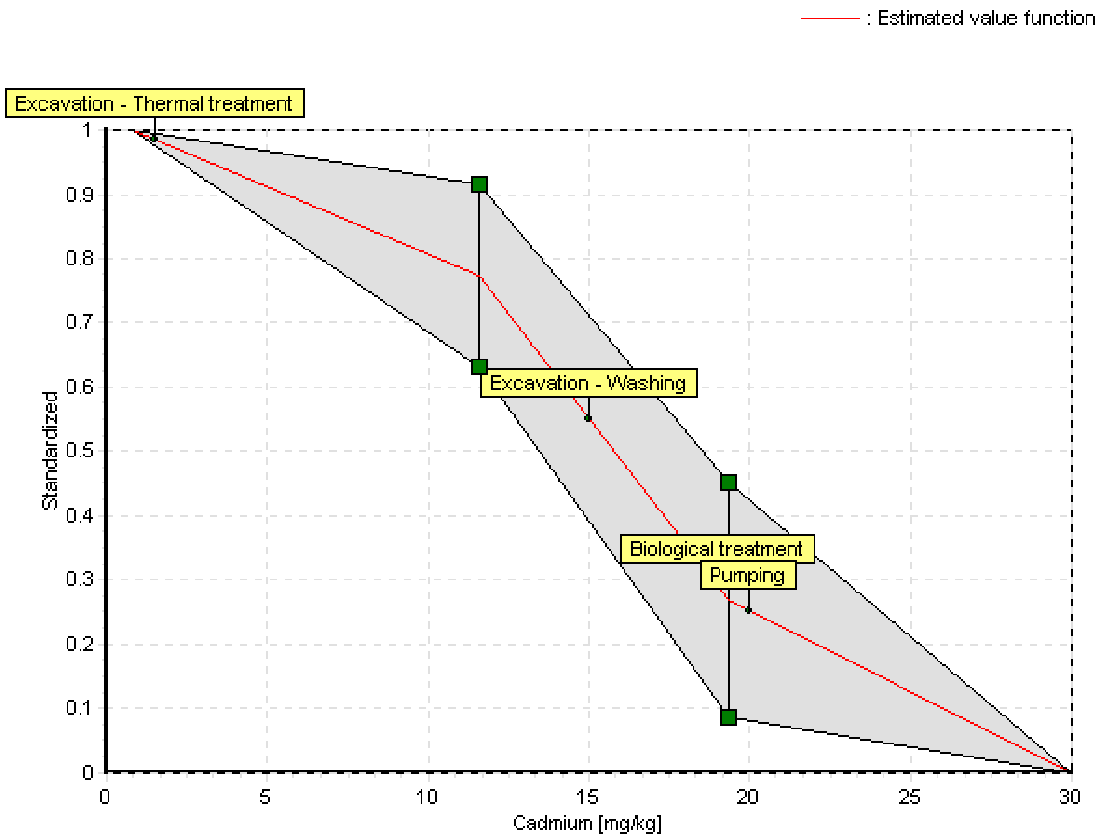


Figure A2. A first estimate of a value function for Cadmium.

2. Qualitative weights
The user is asked to rank the effects from most to least important taking into account the minimum and maximum values of the effects. The weights are interpreted as swing weights. In this example the priority order of the criteria is: Cadmium is more important then Zinc and Zinc is more important then Mineral oil.

3. Conjoint scaling

Conjoint scaling is used to collect information on value functions and weights indirectly through rankings of simple test alternatives. The input screen is shown in Figure A3. Starting point is a group of test alternatives (on the left). These are clean-up alternatives resulting in certain concentrations of Cadmium and Mineral oil. The expert is asked which is the best alternative. It is clear that alternative 7 which has the lowest concentrations on both pollutants is the most preferred. Next the expert is asked to compare alternatives 8 and 4. In this case the expert selects 9. This is followed by a comparison of alternatives 9 and 4 and again the low Cadmium is preferred and the expert selects 9. Since alternative 4 dominates alternatives 5 and 1, alternative 4 is now automatically added to the list. The next step is shown in Figure A3: the expert is asked to compare alternatives 5 and 1. After completing ordering the combinations of Cadmium and Mineral oil, EValue continues with combinations of Cadmium and Zinc and Mineral oil and Zinc. The expert is free to provide only part of the requested information.

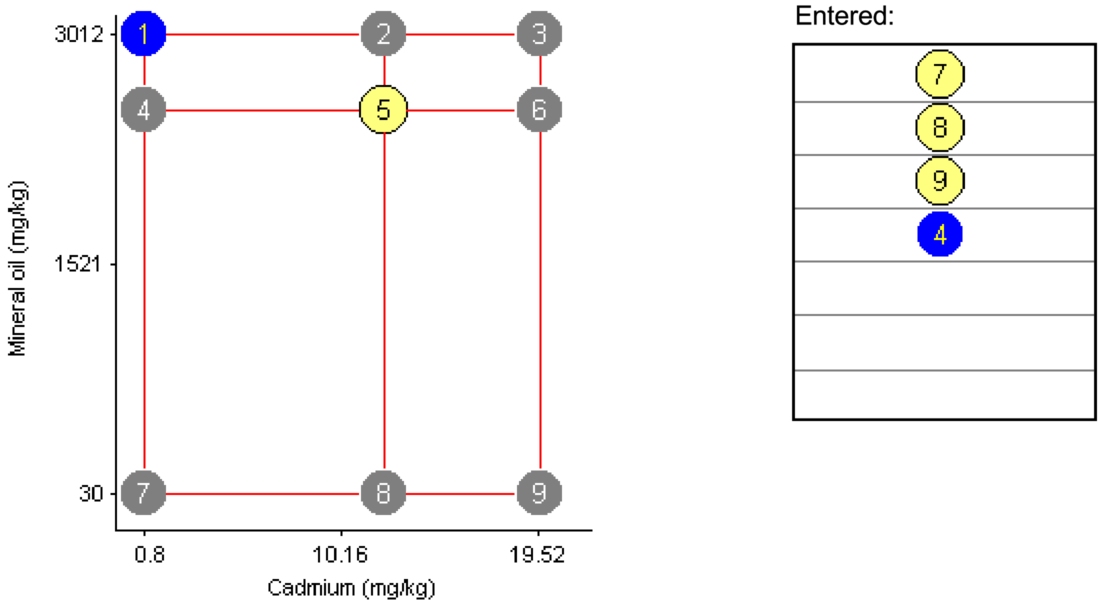


Figure A3. Ranking test alternatives based on concentrations of Cadmium and Mineral oil.

Generate the best estimate of value functions and weights

The expert is asked to provide information on value functions and weights directly through specification of value ranges and qualitative weights and indirectly through conjoint scaling. Both approaches contain in principle the same information and therefore, if the expert provides all information, the problem is in principle over-defined.

To generate the best estimate of value functions and weights based on all available information EValue uses an optimisation procedure that attempts to find value functions and weights for Cadmium, Mineral oil and Zinc such that:

1. The value functions stay within the specified value regions AND
2. The weights respect the qualitative weights AND;
3. The rankings are consistent with the rankings specified with conjoint scaling.

Since the model is over-defined it may not be possible to meet all specifications. In this case the optimisation procedure produces an estimate that best fits the input. The result is presented to the expert. The expert may decide to accept the result, alter part of the input or specify which part of the input is more important then the rest. EValue then produces a new set of results. This is repeated until the expert is satisfied with the result. An acceptable set of value functions is usually reached in about four rounds.

Figure A4 and A5 show part of the results of the first assessment round. Figure A4 shows that the value region of Cadmium as specified by the expert and the value function as estimated by EValue. The figure shows that part of the function is outside the specified region. The value functions for Mineral oil and Zinc (not shown) are within the specified regions. The weights estimated by EValue (0.4, 0.3 and 0.3) match the qualitative weights as specified by the expert. Figure A5 compares the ranking of the test alternatives entered by the expert (left) with the ranking computed using the value functions and weights estimated by EValue (right). It is clear that there are differences between the two. It is now up to the expert to accept this result, to edit the input or to give priority to part of the input.

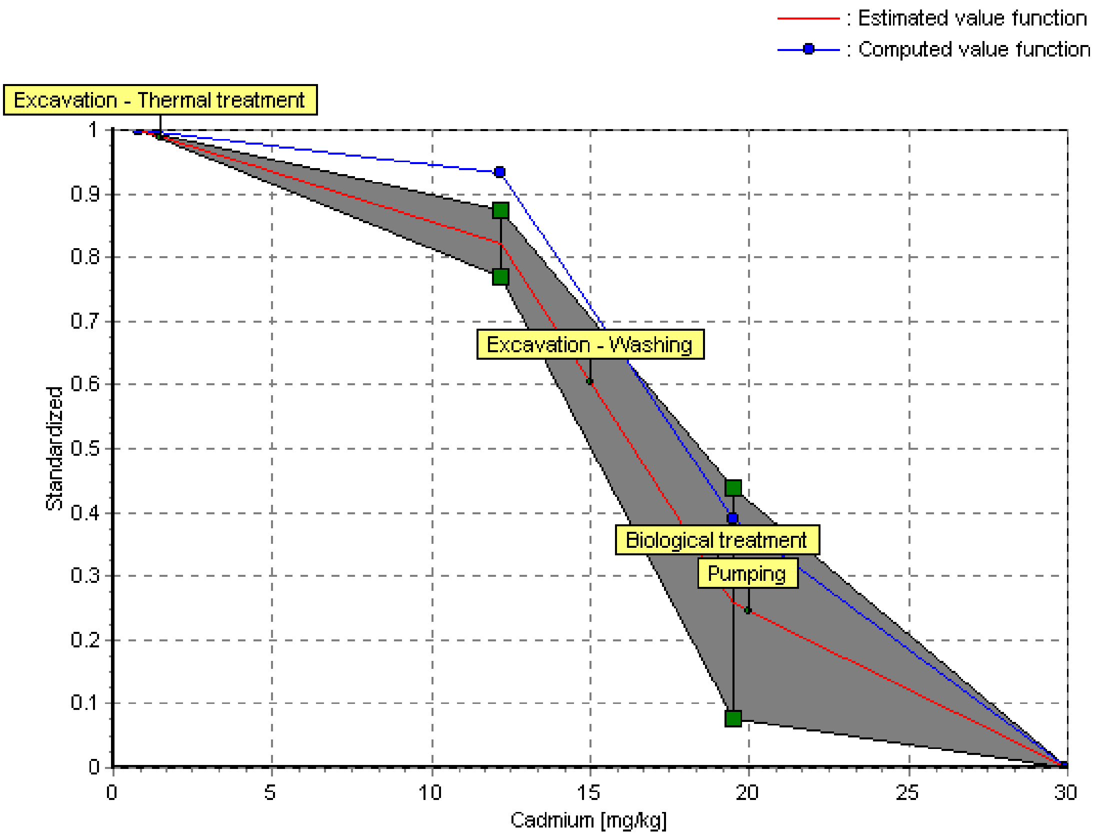


Figure A4. A first estimate of the value function of Cadmium.

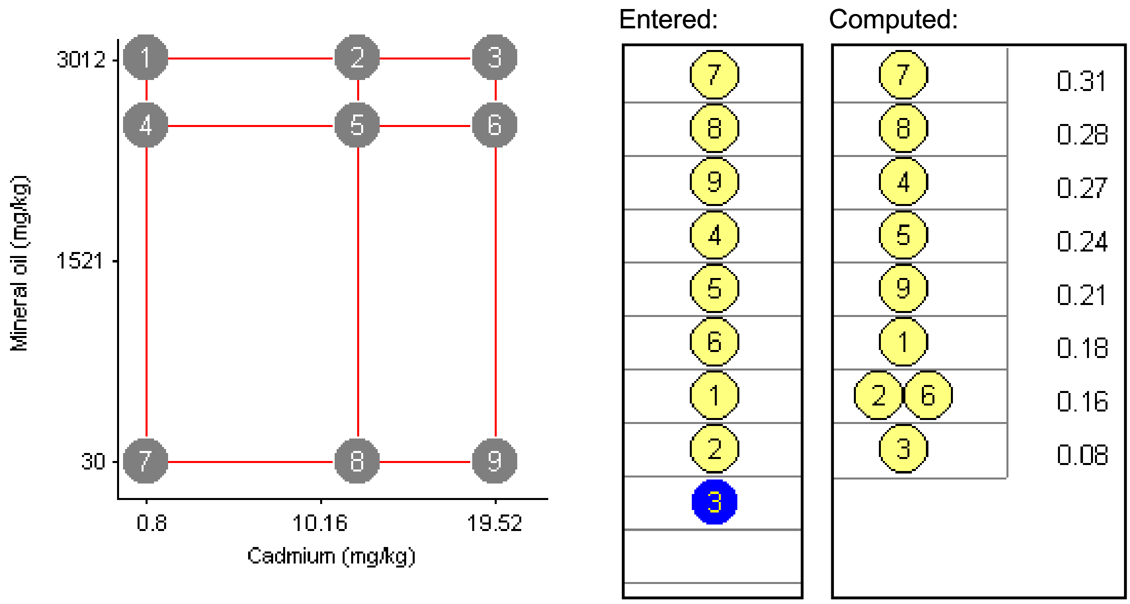


Figure A5. Ranking entered by the expert (left) and computed using the EValue value functions (right).

In the example presented above the value functions and weights linked to the various types of residual concentrations after clean-up were assessed in separate interview sessions with five experts. Assessment of non-linear value functions this way proved to be time consuming and therefore expensive. This approach is therefore most suitable for decision problems with large effects or decision problems that occur several times in similar form, such as the clean-up of oil around petrol stations. If the ranges of scores are not too large linear standardization will be an acceptable approximation of reality in most cases.