Integrated Assessment **2:** 219–224, 2001.

Assessing water quality improvement schemes: The multi-attribute technique of the UK's environment agency

Chris Hope [a] and Ronan Palmer [b]

[a] *Judge Institute of Management, University of Cambridge, UK*
[b] *Environment Agency, Bristol, UK*

The Environment Agency for England and Wales is required to take account of likely costs and benefits in carrying out its duties. Given the complex nature of environmental problems, this task requires sensitivity to issues such as uncertainty, multiple objectives and conflicting value systems. This paper describes a multi-attribute methodology used to carry out this duty in one area of the Agency's work, the regulation of the water industry. The method includes nine attributes measuring the benefits from water quality improvement schemes, and one attribute measuring scheme costs. It is a workable method that clearly satisfies the requirements for the Environment Agency to take account of the costs and benefits of its actions. Refinements are suggested to improve the individual attribute scores, the weights used in prioritisation and the incorporation of costs.

Keywords: multi-attribute techniques, water quality

1. Background

The water industry in England & Wales is regulated by a system of price controls. The economic regulator of the industry (OFWAT) reviews the controls every five years. Part of that review process is determining the level of operating and capital costs that the industry will face. In view of the industry's direct dependence on the environment, both as a source of raw water, and as a destination for waste water and sludge, environmental protection is a significant driver of water industry costs. The Environment Agency is required to judge a programme of expenditure on environmental improvements, as part of the review process.

Of itself, the environmental investment programme brings the companies no financial benefit, as it does not increase revenues or reduce costs. Therefore the costs of the investment have to be passed through to customers, via the price control. The role of the Agency is to verify that the programme is justified. Some of the expenditure is determined by statutory considerations. For example, companies are required to comply with legislation, such as the European Bathing Waters Directive. The balance of the programme is made up of projects driven by non-statutory social benefits, such as enhanced recreation benefits to the general public, or reduced stress on specific habitats. In those cases the Agency is required to identify the relative benefits of schemes. The role of OFWAT is to protect consumers and as such they will evaluate the cost estimates of the companies. While both OFWAT and the Agency are executive regulators, both have networks of regionally based consultative committees. These advise on issues of importance to consumers, in the case of OFWAT, and the environment in the case of the Agency. In a periodic review, the regulators would consult both networks. The ultimate decision on the programme is taken by the Secretary of State for Environment, Transport and the Regions (in England) and the Secretary of State for Wales. Once the decision has been made to proceed with a given project, the company's environmental licence for that activity will be amended, making it a licence condition, with which the company must comply.

Choosing a method of analysis for the task of ranking schemes by their relative merits raised a number of issues for the Agency. Some of these were practical questions. The scale was immense. Over 900 schemes had to be reviewed, ranging in cost from a few thousand pounds to £70 million, over a period of 9 months. This meant that the analysis could not be exhaustive. Moreover the schemes tended to be local in nature, requiring some substantial knowledge of local conditions. Both these features meant that much of the work had to be done by locally based staff, not technically trained in sophisticated appraisal methods. There were also theoretical questions raised. The benefits, and costs, of the schemes affected different parties, and were measured in different units. Hence the analysis had to cope with questions of equity of impact (both in terms of who pays for improvements, and who benefits) and non-commensurability. The Agency also has a duty to contribute to sustainable development and so wished to look at the wider social issues and at the longer-term impact of its actions.

Standard guidance in the UK suggests the use of cost-benefit analysis for issues such as this, although the guidance also makes reference to alternatives, such as cost-effectiveness analysis (where targets are pre-determined) and scoring and weighting (a simplified form of multi-criteria analysis) [1]. For those schemes involving water *quantity* (for example, restoring flows to rivers that ran low in summer) there was an existing cost–benefit methodology [3]. In any case such schemes were fewer in number. For water *quality* schemes, there was also a cost–benefit methodology available in a document called the FWR man-

ual [5]. This was applied to a number of the larger schemes, but it was thought impracticable to apply it to all schemes. Hence the Agency opted to develop a multi-attribute approach.

2. Cost–benefit analysis

The fundamental elements of economic appraisal as applied to issues such as the environment, are opportunity cost and externalities [11]. An opportunity cost is a measure of the value of a resource foregone when it is used for a specific purpose. In a market situation, this cost can be measured by the market price. That is because markets allocate resources in an economically "efficient" manner. However certain resources are outside the market. Such resources, including many environmental resources such as clean water, are said to be external to the market. There are no market prices to measure their opportunity cost.

Hence cost–benefit analysis looks to monetary valuation to provide estimates of the opportunity costs of externalities. The process of valuation makes a number of assumptions, derived from expected utility theory. Values are given, in that people know, or can establish, the value they have for a resource. They have some stability, in that while they may change over time, they will not do so in a random manner. The value of a resource to society is the linear sum of the value of that resource to individuals in that society [9].

Uncertainty over valuation is usually addressed by sensitivity analysis, identifying those parameters to which the decision is sensitive, and determining how it would change if the parameters changed.

A frequent criticism with CBA approaches is that they can be expensive to produce and to tailor to specific projects. In the case of the water periodic review, with over 900 projects, this was a crucial consideration. The FWR manual used a technique called "Benefit Transfer" to overcome this problem. Benefit transfer takes values for environmental goods or services derived from one study, to apply to relevantly similar goods or services in another case. This allows analysis to proceed without a tailor-made study. However this approach is itself open to criticism, as it can be very difficult to justify transfers in specific cases. This can arise for a number of reasons, including doubts about the validity of the original values (which may be based on a single study conducted many years previously) or because the attributes in question are not relevantly similar to the attributes in the original study.

3. Multi-attribute value theory

Multi-attribute value theory (MAVT) is the special branch of utility theory [12] that can be used when a decision needs to be made with multiple objectives but no uncertainty. Its axioms are set out in Keeney and Raiffa [8], and a popular but rigorous description is given in Watson and Buede [13].

MAVT can be used by a single decision-maker in any circumstances where the options for action are well defined, as are the attributes, or factors, which discriminate amongst them. In its most general form, it asserts that each option can be given a score on each attribute and that a value function can be constructed so that the option with the most preferred set of scores has the highest value.

Applications to US child development research programs [2], the Mexican electricity system [10], and UK energy policy [7], amongst many others, show that the technique can be useful for public policy decision making in elucidating issues and in helping identify the implications and trade-offs involved in pursuing different options.

Applying MAVT in its most general form can be a very taxing exercise. The comparisons needed to obtain the value function can be no easier than completely enumerating the set of options. Fortunately there are simplifications that can lead to special forms of the value function that are easier to work with. It is these simplified forms that are most often used in practice.

By far the most common form is the linear additive value function. In this form, a value function is constructed for each attribute separately. The overall value of an option is then the sum of its value on each attribute, multiplied by a weight representing the importance of the attribute [13].

It is not certain that a linear additive value function will be appropriate in every case. In fact, the theory behind MAVT tells us that it will only be appropriate if the decision-maker's preference structure exhibits a particular characteristic called mutual preference independence. A pair of attributes is said to be preference independent of all the other attributes if preferences between different combinations of levels of these two attributes, with all other attributes being held at constant values, do not depend on what those constant values are. The whole set of attributes is said to show mutual preference independence if all pairs of attributes are preference independent of all the others [13, p. 26]. This is similar to the requirement in cost–benefit analysis for the goods or services in question to be mutually exclusive.

Mutual preference independence is something that can be tested with a decision-maker in any particular case, although in practice it is often just assumed to exist. If it is not satisfied, then more complicated value functions than the linear additive one can be used. It is worth noticing in passing that cost–benefit analysis, by its use of money as a common numeraire, is constrained to use a linear additive value function in all cases.

4. A multi-attribute technique for water quality improvements

The multi attribute based methodology used by the Environment Agency in assessing the benefits of surface water quality improvements as part of the third Periodic Review was a linear additive form of MAVT (with mutual preference independence assumed), with the ten attributes shown

Table 1
The ten attributes of water quality improvement schemes.

Informal recreation
Coarse angling
Salmonid angling
In-stream recreation
Agriculture
Industrial abstractions
Drinking water supply abstractions
Amenity value
Nature conservation
Scheme costs

in table 1 [4]. Nine of the attributes measure benefits that are expected as a result of the scheme, while the tenth is the cost of the scheme. These attributes were chosen to cover the whole range of water quality concerns while being as consistent as possible with an earlier FWR assessment method [5].

The methodology was used to assist in screening, ranking, prioritising and justifying proposed schemes for submission as part of the discretionary expenditure. Initially it was intended merely to rank schemes on environmental criteria alone, for use in a cost–effectiveness framework. For any given budget constraint (set in the periodic review) this would give a ranked list of projects for each company, with projects proceeding until the budget constraint was used up. However it was decided to add the tenth attribute, scheme costs, to enable a comparison between costs and benefits.

The value functions for some of the attributes were quite complex. For instance, for informal recreation the value function included scores for the length of the river, the quality of access, the potential for visitors, the improvement in water quality, and the improvement in aesthetic quality.

Detailed guidance was given describing how each of these scores could vary between 1 and 5 (for instance, for quality of access, a score of 1 was given if less than 10% of the river length was accessible, a score of 3 if between 10 and 30% was accessible, and a score of 5 if more than 30% was accessible). The 1 to 5 scoring system was adopted to avoid overly complex and time consuming scoring requirements that might have resulted from using the more common 0 to 100 scale. The scores were combined together using the formula:

$$(\text{Length} \times \text{access} \times \text{visitor potential}) \times (\text{improvement in water quality} + \text{improvement in aesthetics}) = \text{benefit}.$$

The maximum potential benefit score for a scheme on this attribute is $(5 \times 5 \times 5) \times (5 + 5) = 1250$. So the benefit score for each scheme on this attribute was normalized by dividing by 1250 and multiplying by 100. The basic idea is that on each attribute the best possible scheme could get a score of 100, and the worst a score of zero.

To be consistent with the underlying theory, an intermediate scheme with a score of 50 on this attribute must be exactly half as preferred as the best possible scheme – in the sense that if there were a choice between carrying out two schemes with a score of 50 or one scheme with a score of 100, the decision-maker would be indifferent as to which was chosen.

For scheme costs, the same basic idea applied, but reversed, so that the scheme with the highest cost was given a score of −100 and a scheme with no cost would have scored zero.

Experts in the Environment Agency regional offices provided scores for each of the 900 schemes on each of the components that made up the nine benefit attributes. Costs were derived from the Water Companies' costs estimates, as submitted to OFWAT.

To provide an aggregated assessment of each scheme, weights were required to reflect the relative importance of the attributes. It was appreciated from the start that different stakeholders would be likely to have different views about this. So several sets of weights were collected during regional focus groups, using a combined ranking and swing weighting procedure [13, p. 201]. No attempt was made to force a consensus amongst the participants.

A more informal procedure was also used to collect approximate importance weights from the Agency's Regional Environmental Protection Advisory Committee (REPAC) members.

The weights were used in two ways:

The first approach was to aggregate across all impact categories, including costs, to develop an overall measure of net worth. The most expensive scheme (which had a cost of £70 million) was given a costs score of -100, and each focus group participant was asked to provide a swing weight for the attribute "scheme costs" which took this information into account. Any scheme with a cost of £c million was given a score of $-100 \times c/70$. The results provide an indication of whether or not a scheme is worthwhile if there is no budget constraint.

The second approach was to aggregate only the nine environmental benefit scores, excluding scheme costs. This measure of net environmental benefit could then be divided by a measure of scheme costs to prioritise schemes on the basis of benefits per unit of expenditure.

When carrying out the first approach, the weights from the focus groups were used to generate three lists of schemes:

List A: Schemes with a positive net worth no matter whose weights were adopted;

List B: Schemes with a positive net worth using some participants' weights, but a negative net worth using others;

List C: Schemes with a negative net worth no matter whose weights were adopted.

Within each of these lists, the second approach was used to prioritise the schemes. A single set of weights had to be used for this, and those developed through consultation with the REPACs were chosen.

Within the second approach, the scheme costs were also not expressed in monetary terms, but as a roughly geometric

Table 2
Scores for scheme costs used in the prioritisation.

Scheme cost (£)	Score
<20k	10
<50k	20
<125k	30
<300k	40
<750k	50
<1.8m	60
<4.5m	70
<11m	80
<27m	90
<70m	100

score as shown in table 2. The adoption of the 1–5 scoring system for benefits placed a limit on the maximum benefit that any scheme could achieve. Scheme costs in monetary terms are not limited in the same manner, so the geometric cost score was adopted to avoid penalising large schemes relative to small ones.

The three ranked lists were then combined to provide a long list of schemes that was given a final review, with some limited re-ordering to include any local factors not included in the attributes, and then submitted to the Secretary of State for consideration.

5. Lessons from the case study

The MAVT technique used by the Environment Agency clearly satisfied the requirements for it to take account of the costs and benefits of its actions. As constructed, it enable the Agency to identify both costs and benefits, and to make a basic estimate of the relative trade-offs.

The use of carefully elicited swing weights from focus groups marked a clear step forward from the reliance on experts, or on more vaguely defined importance weights, as criticised by Helm [6]. The application of a range of weights, rather than a single aggregated set, did ensure that a range of stakeholder concerns were taken seriously into account.

Having said this, the technique as finally applied by the Environment Agency differed in several ways from a "textbook" MAVT application. The Agency's need to deal with the evaluation of a wide range of projects, within a limited timescale, meant in effect that it was constrained to develop and pilot a technique simultaneously. This means that the details of the technique did not reflect either the state of the art, nor all the core tenets of MAVT. Even at the time of development, it was evident that assumptions were being made in order to deliver the basics, which meant that the final product is still in need of considerable refinement.

5.1. Individual attribute scores

It is not possible to be confident that the value functions conform to one of the basic axioms of the theory: that equal increments in the scores represent equal increases in perceived worth.

At the start of the project, the Agency was particularly keen to maintain a link with previous work in this field, the FWR manual. In order to maintain this link it had to make the value functions in the MAVT approach reflect the underlying functions in the FWR manual. This has led to very complex forms for the value functions, making it particularly difficult to check them against this axiom.

For instance, with the informal recreation attribute discussed above, a maximum score of 100 would be obtained by a scheme:

- affecting more than 30 km of river,
- with more than 30% of its length accessible,
- with a population of more than 50,000 within 3 km of the river and few alternative recreational sites,
- giving an improvement in water quality of more than 1 class, and
- an improvement in aesthetic quality from grade 4 to grade 3 or better.

A score of 50 could be obtained by a scheme affecting the same stretch of river, but only improving the water quality from the bottom of a class to the top of the same class, and improving the aesthetic quality from grade 3 to grade 2 or better. Or it could be obtained by a scheme giving the same improvement in water quality and aesthetics, but only affecting between 15 and 30 km of river with between 10 and 30% of its length accessible and the same visitor potential.

It is actually very hard to say whether the two schemes with a score of 50 do indeed have the same informal recreation benefit as each other. Or whether the scheme with a score of 100 gives double that informal recreation benefit. Yet this is what is required for the axioms of the technique to be obeyed.

Further research may be needed to evaluate this, and to see if the earlier, cost–benefit, work suffers from the same problem. Without this research, one cannot judge whether this led to over- or under-estimates of value.

5.2. Weights used in prioritisation

The use of REPAC weights in the final prioritisation stage was justified by the Agency as respecting the consultation mechanisms defined for it by statute, as opposed to the more informal, if more rigorous, forms of consultation with the focus groups. However, the method of elicitation used with the REPACs did not follow the same rigorous procedures used with the focus groups, and was not as clearly documented. This makes it difficult to test whether the weights obtained are consistent with those that would have been obtained with the rather rigorous procedures needed to obtain the swing weights as required by the theory.

Again it is difficult to judge the implications of this. On the one hand, the outcome may be no worse than using cost–benefit values derived from a benefits transfer model. However equally, it would be preferable to have the assurance of the rigour of the focus group process.

5.3. Incorporation of costs

In creating the three lists of schemes, the scheme costs were included in a straightforward and theoretically defensible way: the disbenefit of scheme cost was assumed to be linear with the cost incurred. This would seem to be a reasonable assumption for public works schemes of such a relatively small size compared to the overall investment budgets of either the water companies or the government.

The information contained in the lists of schemes was potentially very powerful. The schemes on list A provided benefits in excess of their costs, for every set of weights provided by a focus group member. If list A were too long to be approved all at once, it would seem that this should have given a clear indication for increasing the budget for such discretionary schemes as soon as possible.

However, with the budget for the schemes already agreed, it was necessary to find some way of prioritising so that a shorter list of schemes could be sent for immediate approval. Again, the overall procedure of ordering the schemes by net environmental benefit per unit cost was a sensible way of allocating a scarce budget.

But the detail of dividing not by the actual cost, but by a score based in a very non-linear way upon the cost, is questionable. For it to be correct, the scores shown in table 2 above must satisfy the usual MAVT requirement that equal increments imply equal changes in worth. In other words, it must be just as bad to go from a scheme costing nothing to one costing between £300k and £750k, as it is to go from a scheme costing between £300k and £750k to one costing between £27 million and £70 million. It is only necessary to observe that the latter scheme costs at least 36 times as much as the former to see how unlikely it is for this to be true.

The justification given for adopting this procedure was that the scoring method limits the maximum value of benefits that any scheme can achieve, so large schemes are penalised relative to smaller schemes. In order to overcome this, it is necessary to penalise small schemes when it comes to the consideration of costs.

Ideally this would be corrected by changing the benefit scores. If the technique was not giving adequate recognition of the benefits of large schemes, then it is not complying with the underlying theory, and the benefit scores should be extended to allow the large schemes to get full credit for the improvements they bring. The solution adopted was clearly a second best option, driven in part by time pressure, and by the belief that, as a similar adjustment would be needed for all benefits, the final outcome would be mathematically the same. This ad hoc adjustment is rather like saying that because you know your car's accelerator is faulty, you are always going to drive with the brake on. A better remedy might be to fix the accelerator! Clearly future applications need to take account of scheme size, or of distribution of schemes, more explicitly.

In practice, the final decisions did not hinge on the relative benefits, as to a large extent, the proposals made by the Agency were accepted in full at aggregate level by the Secretary of State.

6. Next steps

The MAVT approach proved to be workable in practice, but requires substantial development for future use. These developments should address a number of areas:

- The theoretical issues highlighted above, in particular simplifying the value functions, and ensuring that they comply with the underlying theory.
- The difficulties with scheme size & distribution of benefits must be addressed explicitly, and the way in which schemes are scored must be changed to do this.
- Include other disbenefits (e.g., noise, energy use by plant, etc.) as well as the costs of the scheme. This is easy to do in principle with a multi-attribute approach.
- Establish a link with sustainable development as the overall objective.
- Emphasise net worth including costs rather than cost–effectiveness, reflecting a move towards first-best rather than second-best assessment.
- Collect more representative sets of weights, either by using specially constituted focus groups, or by a revised process of using the REPACs, possibly combining them with a wider cross-section of the population.
- Cross-check the results with the results of the most recent contingent valuation exercises.

References

[1] Department of the Environment, *Policy Appraisal and the Environment* (HMSO, London, 1991).
[2] W. Edwards, How to use multi-attribute utility measurement for social decision-making, IEEE Transactions on Systems, Man and Cybernetics SMC-7 (1977) 326–40.
[3] Environment Agency, *Low Flow Alleviation: Benefit Assessment Guidelines* (Environment Agency, Bristol, 1997).
[4] Environment Agency, *MAT Scoring and Weighting System for Water Quality Improvements* (Environment Agency, Bristol, 1999).
[5] FWR, *Assessing the Benefits of Surface Water Quality Improvements Manual* (Foundation for Water Research, Marlow, 1996).
[6] D. Helm, The Assessment: Environmental policy – objectives, instruments and institutions, Oxford Review of Economic Policy 14(4) (1999) 1–18.
[7] M.R. Jones, C.W. Hope and R. Hughes, A multi-attribute value model for the study of UK energy policy, J. Opl. Res. Soc. 41(10) (1990) 919–929.
[8] R.L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* (Wiley, New York, 1976).
[9] R. Laslett, The assumptions of cost–benefit analysis, in: *Environmental Valuation: New Perspectives*, eds. K.G. Willis and J.T. Corkindale (CAB International, Wallingford, 1995).
[10] J.E. Matheson, Decision analysis practice: examples and insights, in: *The Principles and Applications of Decision Analysis*, eds. R.A. Howard and J.E. Matheson (Strategic Decision Group, Menlo Park, CA, 1983).

[11] R.K. Turner, D. Pearce and I. Bateman, *Environmental Economics: An Elementary Introduction* (Harvester Wheatsheaf, Hemel Hempstead, 1994).
[12] J. von Neumann and O. Morgenstern, *The Theory of Games and Economic Behaviour* (Princeton University Press, 1944).
[13] S.R. Watson and D.M. Buede, *Decision Synthesis* (Cambridge University Press, 1987).