International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Regret Analysis for River Water Quality Management

Jolma, A.

IIASA Working Paper

January 1995

Jolma, A. (1995) Regret Analysis for River Water Quality Management. IIASA Working Paper. Copyright © 1995 by the author(s). http://pure.iiasa.ac.at/4588/

Working Paper

Regret Analysis for River Water Quality Management

Ari Jolma

WP-95-4
January 1995

IIASA
International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

Regret Analysis for River Water Quality Management

Ari Jolma

WP-95-4
January 1995

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

ABSTRACT

The treatment of uncertainty inherent in water quality policy development and related modeling is a key issue. This is particularly true in the Central and Eastern European region which is characterized by high pollution levels, strong economic transition, lack of financial resources and scarce and imprecise data availability. Possibilities to account for uncertainties in policy development by a regret analysis approach were studied. The risks associated with regret in decision making were qualitatively and quantitatively defined. The design of a decision making procedure where these risks are accounted for was analyzed from a theoretical and practical viewpoint. A regret analysis procedure was implemented within an experimental decision support system (DSS) for water quality management of a river. This system was tested with a hypothetical situation where the best option of upgrading of a number of waste water treatment plants was to be selected from a small set of possible options. The main problem of analyses of this kind is the curse of dimensionality. The study shows that a modular approach, which uses already available software components and models, speeds up the building of a DSS.

REGRET ANALYSIS FOR RIVER WATER QUALITY MANAGEMENT

Jolma, A.[1]

1 INTRODUCTION

It will not be an easy task to improve the water quality of rivers of countries in the Central and Eastern European (CEE) region. Reasons are manifold: the high level of existing pollution, the co-existence of contamination problems of differing origins (which were handled sequentially in the West in the course of past decades), strong political, economic, social and institutional transition, the lack of financing of environmental remedial measures and so forth (see Somlyódy, 1993). Stemming from all these features such policies are looked for that are affordable in the short term, yet flexible, allowing for a step-by-step tightening of water quality standards as economies in the region improve. As an element of such a strategy, the upgrading schemes of existing waste water treatment plants or the construction of new ones require careful river basin planning by using water quality models. In spite of all the difficulties, the situation has also promising aspects, namely all the experiences and methodological advances gained in past decades can be utilized to develop innovative, efficient policies. It seems possible and inevitable to introduce more ingenious emission reduction schemes as the traditional uniform emission control, considering jointly existing economic constraints and opportunities given by new legislation which are under preparation in many countries in the region. A general description of this subject is given by Somlyódy (1993).

The treatment of uncertainty inherent in water quality modeling is a key issue in particular with water quality of rivers of CEE. The main reason for this is the overall scarcity of information and data. This fact result in difficulties in identifiability, calibrationability, and verifiability of water quality models unless special monitoring programs are launched. The uncertainty also introduces the element of risk into the decision making. The results of the water quality models are not certain, the model might be inadequate, the system itself might change, and the conditions, which are used for the design, may be realized differently than assumed.

One case study of the water project at IIASA addressed the issue of developing water quality management strategies for the Nitra River basin in Slovakia (Somlyódy et al. 1994). The Nitra is a tributary of the Váh, which joins the Danube downstream of Bratislava. Uncertainty in many levels of modeling has been studied; these include parameter uncertainty and its implications in optimization models and scenario analysis approach in a regret analysis framework. The "regret" defined as an over-investment or as a failure to meet the target water quality. In the report of the Nitra study, Somlyódy et al. (1994) refer to Burn and Lence (1992) for broader regret analysis. In both of these studies the set of decisions, which is taken to regret analysis, originates from optimization models and they are the "best" decisions assuming certain design scenario and model parameters.

In this paper a policy problem is addressed where one should choose from a number of decisions, assuming a number of possible conditions, or states including the present and the future alike. This

---

[1] Helsinki University of Technology, Laboratory of Water Resources Engineering, FIN-02150 Espoo, Finland

situation and its analysis are first described and considered out of the water quality management context. This part of the study is based largely on the regret theory by Loomes and Sugden (1982). The main contributions here are the qualitative and quantitative definitions for the two aspects of risk present in decision making. The first aspect is the possibility and amount of disappointment one feels when the result of a decision is worse than s/he expected or targeted when making it. The second aspect of risk defined is the possibility and amount of regret which arises from one's observation that, by having made some other decision, the impacts would be better than they are (as the result of the decision s/he actually made). Loomes' and Sugden's regret theory is also applied to water quality management by Cardwell and Ellis (1993). They define regret as "a quantified measure of avoidable water quality deterioration". Crousillat et al. (1993) report of a decision making procedure which tries to minimize the risk associated with a decision.

The rest of the paper describes an implementation of a regret analysis within an experimental decision support system (DSS) for water quality management of a river. The DSS was tested with a hypothetical situation where the best option of upgrading of a number of existing waste water treatment plants was to be selected from a small set of possible options. The test setup also included a hypothetical decision maker because the whole procedure of choosing the upgrading scheme at various locations in the river basin is considered to require a significant amount of interaction between the analyst and the decision maker. It is also a rather case specific procedure.

2 DEFINITIONS AND MATHEMATICAL BACKGROUND

Suppose we have a simulation model M_S of some system to determine the consequences $\mathbf{q}$ of a decision $\mathbf{d}$ under conditions $\mathbf{f}$:

$$\mathbf{q} = M_S(\mathbf{d},\mathbf{f}) \tag{2.1}$$

In the following a particular realization of the vector $\mathbf{f}$ will be called a future. A future can be defined as a subset of all non-controllable input variables and parameters of the model. Let us also suppose that the purpose of using the simulation model is to find the best decision $\mathbf{d}^*$, the "best" defined in some way. This problem can be formulated as an optimization model M_O:

$$\mathbf{d}^* = M_O(\mathscr{D},\mathscr{F}), \tag{2.2}$$

where

$\mathscr{D}$ = the set of considered decisions
$\mathscr{F}$ = the set of considered futures

The problem of finding $\mathbf{d}^*$ from a given $\mathscr{D}$ and $\mathscr{F}$ will be studied in the following.

The stochasticity of the nature and our imperfect knowledge and measurements introduce different kinds of randomness into this problem. The simulation model is based on knowledge and measurements that we have of the system under study and similar systems. Because the simulation model is based on imperfect information it contains uncertainties. A future contains, besides uncertainties, also stochasticity because in it there is future values of variables.

2.1 The regret matrix

By applying the simulation model M_S for each considered decision $\mathbf{d}_i$, i=1...m, and for each separate future $\mathbf{f}_j$, j=1...n, we get a matrix $\mathbf{Q}$:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_{0,0} & \cdots & \mathbf{q}_{0,j} & \cdots & \mathbf{q}_{0,n} \\ \ldots & \ldots & & & \ldots \\ \mathbf{q}_{i,0} & & \mathbf{q}_{i,j} & & \mathbf{q}_{i,n} \\ \ldots & & & \ldots & \ldots \\ \mathbf{q}_{m,0} & \cdots & \mathbf{q}_{m,j} & \cdots & \mathbf{q}_{m,n} \end{bmatrix},$$

where each row is associated with a decision and each column is associated with a future. $\mathbf{q}_{i,j}$ is the consequences of decision $\mathbf{d}_i$ and future $\mathbf{f}_j$. **Q** is traditionally called a regret matrix. If we want to handle **Q** numerically in an analysis, m and n must be small enough.

One way to get a reasonable number of initial decisions is to break the optimization model M_O into two models: M^1_O and M^2_O, where M^1_O is used to find the best decision under one specified future and M^2_O, is used to analyze the consequences of different futures. If we now have a set of futures $\{\mathbf{f}_i\}$ we get, by applying the model M^1_O for each separate future $\mathbf{f}_i$ at a time, a set of conditionally optimal decisions $\{\mathbf{d}_{ic}^*\}$. Applying the model M^1_O also leads to a square matrix because in that case each decision is associated with a future. If we then concentrate our interest only on the differences between consequences of a decision at the design future and at some other future, i.e. let $\mathbf{r}_{i,j} = \mathbf{q}_{i,j} - \mathbf{q}_{i,\mathrm{design}}$, we get a new matrix **R**, which has zeros in the diagonal vector. The matrix **R** obtained this way, is the regret matrix used e.g. by Burn and Lence (1992). Each row of **R** is associated with a conditionally optimal decision $\mathbf{d}_{ic}^*$, and the analysis of **R** should lead to a decision to pick one of them as the overall optimum decision.

In this paper, however, each decision is generally *not* associated with a future. This is because of the notion that the best overall decision might not be the best decision assuming some future. This fact is described more in the next chapter. In this case generally $m \neq n$ and **Q** is not a square matrix. In the simplest case there is only one future, which reduces the **Q** into a column vector. This is of course discouraging from the point of view of finding a reasonable number of decisions to the regret analysis.

The size of the **Q** and subsequently the size of the problem must be made small enough by using some strategy to shrink the number of futures. This can be done for instance by using some averaging scheme. If the decision maker can assign a weight or probability to each scenario, a weighted average could be used. This reducing of columns is equivalent to searching for a center point for the set of points in objective space.

If we define a measure for the nearness of two decisions then we can group decisions and maybe pick the group with most decisions because that can be thought as a robust decision. Actually the grouping of decisions is not generally possible using only one dimension.

It is also possible to use a threshold for non-acceptable behavior under any future and that way remove unsuitable decisions.

2.2 An extended multicriteria optimization problem

Owing to the fact the elements of regret matrix **Q** are not scalars, in general; the analysis of it involves a multicriteria optimization problem. The additional element in the analysis is because of the many futures - columns in **Q**. In the following the resulting extended multicriteria optimization problem is first defined. Secondly, some prospects of reducing the dimensionality of the problem are discussed.

The principle of solving a decision problem by mathematical optimization is based on finding the best decision, which is represented by a numerical value of for instance a vector, or by a function. The "best" is defined in the objective space $\mathcal{Q}$. The objective space is the space of the consequences **q**.

The difference of using many futures instead of only one is that, one decision is not represented by a single point but as a series of points in $\mathcal{Q}$.

The optimization model compares different decisions by comparing their consequences in the objective space. Comparison of single points in one-dimensional space is always unambiguous. Comparison of single points in multi-dimensional space can be done unambiguously at least in one dimension at a time. Comparing a set of points in one- or multi-dimensional space is never guaranteed to be unambiguous, not even in one dimension at a time. In figure 2.1 there is an illustration of this fact.

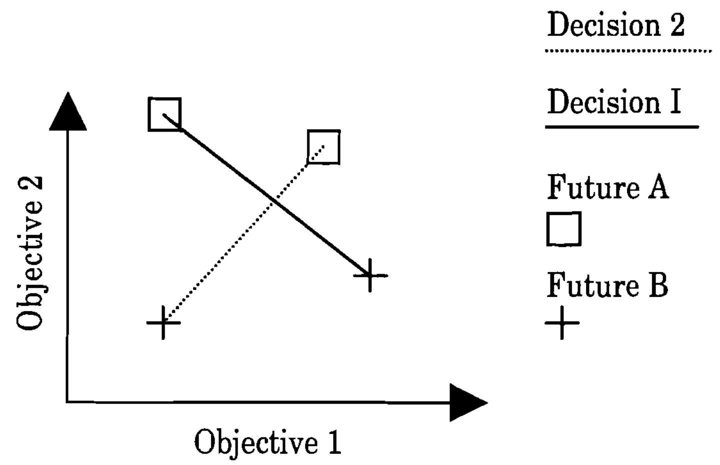

Figure 2.1 The effects of two decisions and two different futures in two-dimensional objective space.

The first step in a multicriteria decision making is to find the pareto-optimal subset of all decisions. In order to find the pareto-optimal set of decisions we need a way to compare effects of decisions in the objective space. In the case of multiple futures this can be done by comparing effects of decisions one future at a time. This leads to a situation where the pareto-optimal set does not form a hyper-plane in the objective space but rather something like a "fuzzy" hyper-plane.

2.4 An extension to risk analysis

The decision maker faces two kinds of risk when making a decision. The first one is the risk of disappointment. The second one is the risk of regret. These two aspects of risk are discussed in this chapter.

When making a decision, a decision maker usually has a goal or satisfaction level that he wants to achieve - as a high jumper when he decides to participate in a sports meeting has a target height he wants to clear. If it then happens that the future happens which leads to a situation that is worse than the goal, the risk of disappointment associated with that decision also realizes. We define the risk of disappointment as the set of (negative) differences of utilities and their probabilities.

If a decision maker makes a certain decision and a certain future happens there is some amount of regret if the result would be better if he/she had made some other decision. If we assume that a person regrets not making some other decision only in the case when that decision would have resulted all criteria having better values, then the risk of regret making that decision is in the case that certain future the differences in utilities between those decisions. Because we cannot assign different probabilities to different decisions we must assume that they are the same for each. The calculation of these numbers over the whole set of considered futures gives the total risk of regret associated with that decision.

Again, to reduce the dimensionality of the problem, one may want to calculate a weighted sum of these two types of risk and then calculate a single number to represent the total subjective risk associated with a decision. The single number might be the maximum, the integral over the probability range [0...1], the center of gravity of the area, or some other statistic. After the risk assessment a two-dimensional plot with risk and expected utility of different decisions is a good way to summarize the risk analysis.

3 A SIMPLE, ILLUSTRATIVE EXAMPLE OF SCENARIOS, RISK, DISAPPOINTMENT, AND REGRET IN DECISION MAKING

Let us suppose that you are at an ice-cream stand. The stand sells only portions made of one brand. You are wondering whether to buy or not. We assume here that the decision to buy and the decision not to buy have the same apriori value, i.e. you have no pre-attitude towards either of them. This decision-making situation is described graphically in figure 3.1.

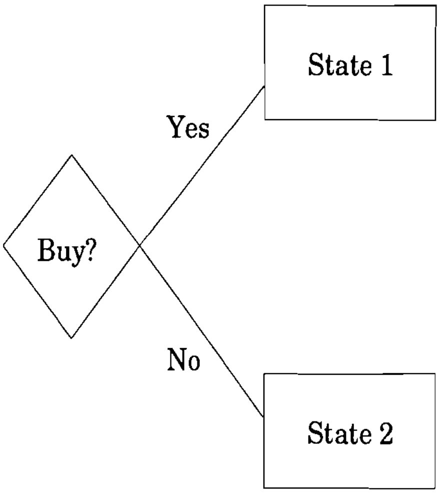

Figure 3.1 A decision and its possible consequences in the illustrative example.

You can describe your feelings in different end states the following way:

1: Depending among other things on the taste of the ice cream you are glad you decided to buy and you enjoy your ice cream portion or you do not enjoy your ice cream and regret your decision to buy in the first place.

2: Your feelings are a bit mixed, you might be a bit sorry not even trying but on the other hand you are glad you did not have to go through the possible disappointment.

You might be able to assign a real value to both end states, which reflects the amount of pleasure or displeasure you feel. This value is called an (expected modified) utility - expected because you are not sure of the outcome and modified because the element of regret or rejoicing is included. If your utility at end state 1 is greater than that at end state 2, you would decide to buy.

You still cannot decide and want to study the decision problem a little more. One way to proceed is to think about possible futures. Several come to your mind:

A. The ice cream tastes bad.

B. The ice cream is good but you are clumsy and drop it.

C. There are salmonella bacteria in the ice cream.

D. The ice cream is really delicious.

We can now write the decision problem into a matrix form (figure 3.2).

		Future				
		A	B	C	D	X
Decision	Yes	1A	1B	1C	1D	1X
	No	2A	2B	2C	2D	2X

Figure 3.2 The regret or scenario matrix of the illustrative example. X denotes all other possible futures besides A,B,C and D that can happen.

Besides explicitly defined futures there are also three other kinds of futures:

1) Futures, which you regard as impossible or having only very minimal probability. An example in our case would be a future, where somebody has deliberately put poison into the ice cream to kill you.

2) Futures, which you regard as having no effect on your current decision making process. An example in our case would be a scenario where the ice cream, which you thought was strawberry, is actually raspberry but you don't mind.

3) All the rest of all possible futures.

For the utility analysis you should now first be able to describe your feelings at each end state except 2B, which is an impossible end state. At the second step you should give a real value for your utility at that state without comparing where you would be if you had made a different choice. And at the third step you would compare utilities of end states of two decisions and one future. After these steps for each end state you have an array of modified utilities, one for each different decision you could have made. In the ice cream example there would be only one modified utility for each end state because there are only two decisions.

Let us now assume that there are two main criteria, on which you base your decision. So you would have to have done the analysis separately for both criteria. The first criteria is the monetary loss you experience if you buy and the second criteria is a more ambiguous one, your feeling of pleasure or displeasure. Let's also assume that you're not alone and you will hear comments how the ice cream tastes. The future C is a little problematic because it might take time before know it happened. If you have made the analysis described above, you can plot your utilities in a two-dimensional space as in figure 3.3. Let the dotted line in figure 3.3 to represent your target or satisfaction level between pleasure and monetary loss.

From the figure 3.3 you can deduct that end states 1D, 2A, and 2C are positive, you are glad if you end up at those states. Similarly end states 1A, 1B, 1C, and 2D are negative, you are not glad to end up at those states, you feel disappointment. If the future D happens and you have made the decision 2, you feel regret but if the future A happens and you have made the same decision you rejoice. From this you can deduct that both decisions, to buy or not to buy, contain risks. The decision to buy contains risks because you might end up in a situation where you have less money and you feel less pleasure. The decision not to buy contains risks because you might end up in a situation where you feel less pleasure.

You can make the decision solely based on the risks associated with the decisions. For instance if you assess the risk of getting salmonella being too high you would choose not to buy. And if you assess the risk of loosing the pleasure to enjoy good ice cream being too high you would choose to buy - of course in a real situation if you are with friends who buy, then you might choose to wait and hear their opinions.

If you can place (subjective) probabilities on different scenarios, you can calculate the expected (modified) utilities for each decision. Plural, because if there are more than two possible choices, then there are more than one expected modified utilities.

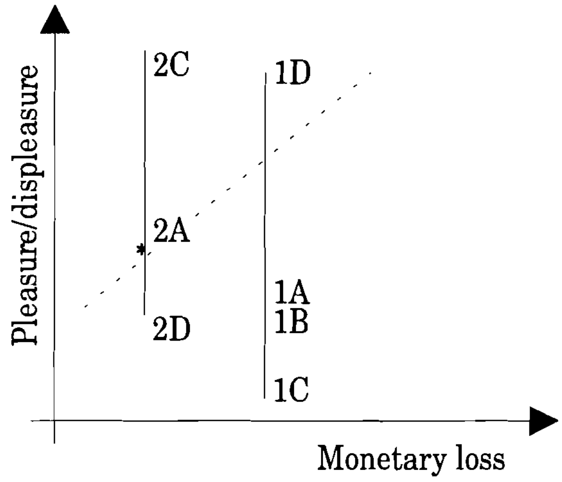

Figure 3.3 The utilities at the objective space. The axes don't cross at point (0,0), which is marked by a star (*). The point (0,0) -the star- is the point where the decision maker is before the decision. The dotted line represents the decision maker's target or satisfaction level between pleasure and monetary loss.

After having described the basic concepts of theory of regret in decision making we will proceed to the water quality management of rivers in the next chapter. Water quality management is decision making, where some management measures/projects are executed/started wholly, partially or not at all. The motivation behind the management is a need and quest for better ambient water quality. This motivation is many times a more or less vague pressure from the society but it can have more exact expressions in the form of water quality needs of water supply or industries which use water as raw material. Execution of water quality management measures requires effort and economic resources. The amount of "pleasure/displeasure" the society and/or users of the water resources feel is the measure against which the successfulness of the management is measured. When the results of the management actions are less what was expected there is regret and disappointment. The regret may have the form of feeling of overspenditure and the disappointment may have the form of feeling of not having tried hard enough.

4 IMPLEMENTATION OF REGRET ANALYSIS IN WATER QUALITY MANAGEMENT OF RIVERS

The water quality management problem is here defined/understood the following way. The wastewater treatment levels on point sources along a river must be upgraded to achieve an acceptable ambient water quality level. The upgrading requires money and in the return of this investment on the waste water treatment plants (WWTPs) the decision maker expects to see some improvements in the nature. So the decision making requires balancing between these two main criteria. In this analysis it is assumed that the decision maker is willing to participate in the process of reducing the dimensionality and then make his/hers final decision based on a small number of one- or two-dimensional information.

If the decision maker wants to make a robust decision and there are no legal or other kinds of imperatives to use only one design future, a scenario analysis might be appropriate. Also the complexity of the water quality model with many input variables and complex structure might make it unclear what is the critical future. Having many control points where the water quality is checked also make it unclear what is the critical future. A scenario analysis is also one way to express our uncertainty upon the system, the nature, and the future conditions.

The decision maker and the analyst together must decide upon a methodology how to generate or to use a certain set of design futures. This set must cover the whole range of possible futures, at least in a subjective way, for the scenario analysis to be a comprehensive one. This means that the sum of the probabilities of all futures should be one. If the values of the variables are picked up from an discrete probability distribution, it is possible to assign probabilities to each future. It might also be the case that the decision maker and analyst can only give a qualitative, or fuzzy, probability value (e.g. "very important", "important", etc.), or no probability value at all, to each future. In these cases the subsequent analysis must be an appropriate one.

The decision space is the set of all possible combinations of waste water treatment levels. In a quite small case where there are ten WWTPs each with ten possible upgrade levels the size of the decision space is 10^{10} vectors. In the decision making process one vector out of all possible vectors is selected. Considering scenario analysis there are two basic ways to carry out this process. 1) Produce a manageable set of candidate decisions and obtain their implications on the economics and on the environment under all different futures and then feed this data into the scenario/regret analysis. 2) Pick a decision randomly from the decision space, obtain its implications, use that information in the scenario analysis and let its results guide your next pick. The first one does not use feedback and the second one does. It is possible to use a pick size larger than one in order to not end in a local optimum. These algorithms are visualized on figure 4.1.

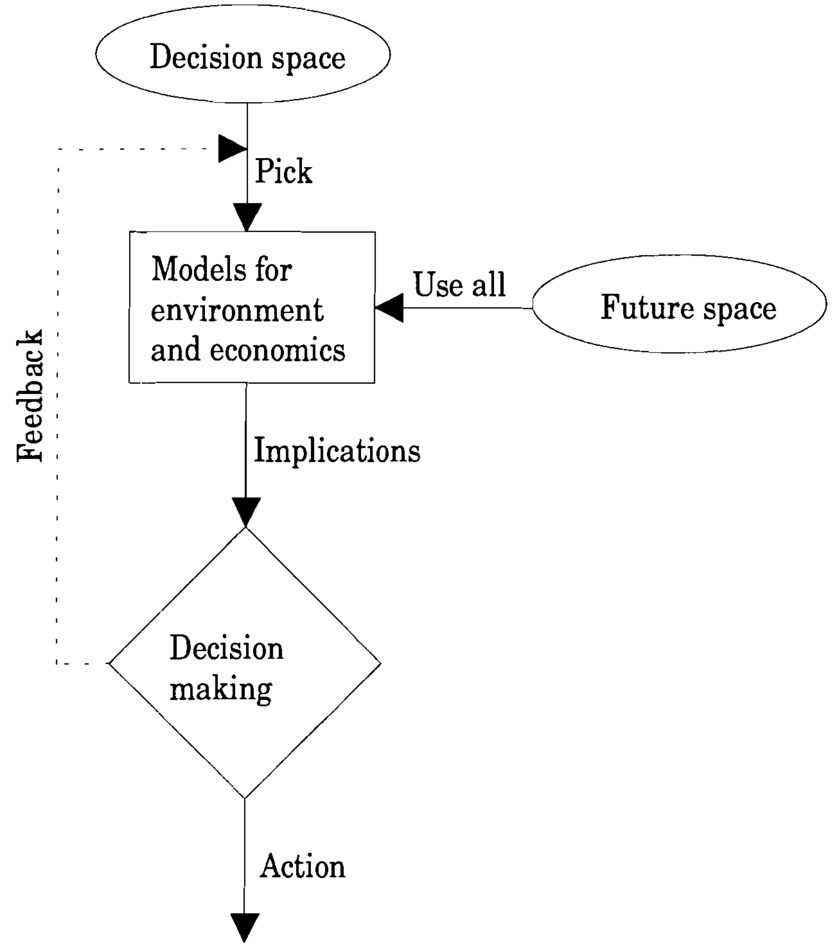

Figure 4.1 The decision making process in water quality management with scenarios

There are problems in both aforementioned basic approaches to decision making process utilizing scenario analysis. The problem in the first approach is, how to produce the initial candidate set. There are many methods one can use, one is to use a single criteria optimization algorithm to find the best decision assuming one specific future and certain objectives. Another method is to use an expert guess. Probably the best method is a mixture of many methods so that the candidate set is not biased. The problem in the second approach is, which method to use in the feedback/guiding of the pick. A mathematically sound method would be to estimate the multivariate cumulative density function of the probability of a decision being good and then use that in a traditional Monte Carlo way to produce random variables of desired probability distribution function. This could however prove to be difficult or even impossible in practice. One method, which has shown potential in problems like this, is the use of genetic algorithms.

The river water quality management can use models for simulating the water quality in the river given certain discharge conditions, temperature profiles, effluent discharges and contents, etc. It can also use models to calculate the effects of the water quality on the flora and fauna of the river, and models to calculate the investment costs and management costs of a given construction or upgrade plan. There is no end to variables describing different aspects of water quality management because it is always possible to calculate secondary variables from the ones already available. For practical reasons it is however many times necessary to boil down the problem to two criteria opposing each other. Many times these are water quality and economics. After a decision has been made using only these two criteria, it faces the tough challenge of other criteria as justness, possibilities/impossibilities of enforcement etc.

There is no unique function - even for one person - which could be used to calculate one number to represent the utility of water quality for the decision maker and which would also be at the same time independent of the economics. This is simply because different things become more important when you use more money. Instead of an utility function a preference structure can be defined. This can take the form of an ordered list of water quality conditions. A water quality condition can be e.g. "DO must be above 3 mg/l at all places" or "there must not be algae blooms". From the implications of a scenario it is then straightforward to tell which conditions are met and which are not met - of course only to the extent the model, which is used, tells us.

An one-dimensional water quality criteria can be obtained using the importance list defined above. The formulation is again a subjective matter and should be decided in accordance with the decision maker. One possibility is to use utilities assigned to meeting a condition and to not meeting a condition. The total utility of a certain water quality conditions would then be the sum of all of them.

The cost for realizing a river water quality management plan can be divided into two parts: the investment costs and the management costs. There is uncertainty in both of these and the actual costs depend on the future and also partly on some economics modeling decisions (e.g. economic life). A commonly used one-dimensional cost-variable is TAC (total annual cost), which is a sum of investment costs multiplied by capital recovery factor and maintenance costs.

4.1 The regret analysis

The first step in the analysis is the check for pareto-optimality of all the decisions in the regret matrix. This requires no special considerations because of the water quality management problem and can be done the way described in the chapter 2.

If the completeness condition for futures (see above) is fulfilled, the expected utilities (expected utility for water quality, expected costs) and other statistics of a decision can be calculated. Also, if we step a little backwards to expressing the water quality with more than one variable, numbers like the percentage of times (sum of probabilities of futures) a water quality condition is met can be calculated. The less numeric and more qualitative the different factors are the more ingenuity the analyst must have to sum up the information for the decision maker.

In the light of regret theory all utilities defined above should be treated in the decision making as "choiceless utilities". That is to say they are utilities the decision maker feels when he gets them *without making any decision*. If the decision maker makes his/hers decision solely based on expected utilities calculated from those, he/she has not taken into account risks involved in making that decision. In the following we will be considering how to formulate regret analysis in such a way that it gives information also on risks besides expectancies associated with decisions. The assumption is then, that the decision maker can make a good, rational, and maybe even optimal decision solely based on that knowledge (i.e. expectancies and risks).

In a decision making situation with multiple criteria the risks associated with a decision also have multiple dimensions. In principle all these figures should be calculated. This is because otherwise we would need a function to combine utilities of different criteria, which is generally not available and should be avoided.

The assessment of the risk of disappointment requires attaching a goal or satisfaction level for each decision. In water quality management a goal comprises of a water quality level and of a cost figure. In scenario analysis each decision is mapped to a set of points in objective space. In the case of a two dimensional objective space the goal is a curve, and points on the worse side of that curve mean disappointment, and points on the other side mean a positive surprise. This is shown for the ice-cream example in figure 3.3, where the dotted line divides the objective space into the disappointment space and positive surprise space.

The goal must be obtained from the decision maker interactively and probably as a continuation for the definition of the one-dimensional water quality criteria. A failure in the process of defining the goal strongly suggests that the water quality criteria is unusable and should be changed or that the decision maker needs to consider the problem using more than one water quality criteria. In the last case two, or more, parallel scenario analyses should be run.

The assessment for the regret in each scenario is quite straight forward if we assume that a person regrets only decisions, which would have resulted in scenarios where all criteria have better values than in the one which realized.

The choice of the weights for risk of disappointment and risk of regret is totally subjective and as such it should be controlled by the decision maker. The same should also apply to the choice of the statistic used to describe the probability/utility set of points of risk with one number.

5 IMPLEMENTATION OF SCENARIO ANALYSIS IN A RIVER WATER QUALITY MANAGEMENT DECISION SUPPORT SYSTEM

A DSS has generally three parts: a database, a model base, and an user interface. The database of a DSS consists of two kinds of data: data which tells about the structure of the problem domain (metadata) and "raw" data. Some examples of pieces of raw data are a real number 3.286, a bitmap image file shot.bmp, and a record ["Nitra";15;shot.bmp]. Raw data is completely useless without metadata, which tells what each piece of it signifies and how it is related to other pieces. A model is a result of modeling and it is used for some purpose. In the figure 5.1 there is a decomposition of the concept DSS.

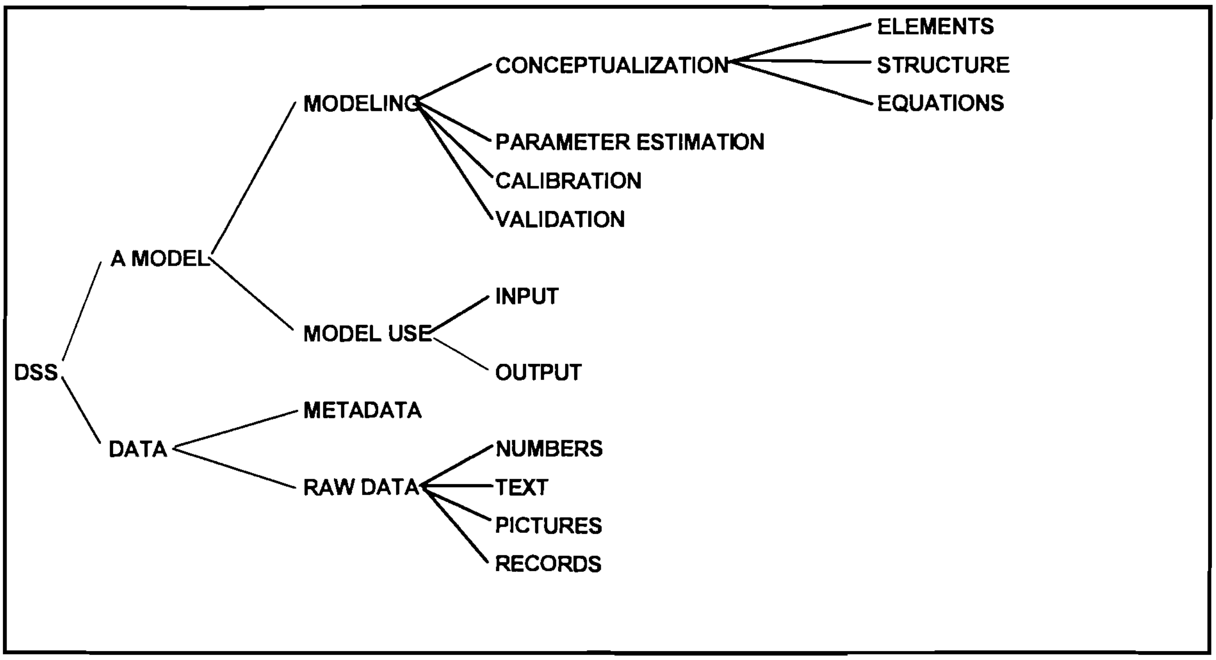

Figure 5.1 Decomposition of the concept DSS. The object of data modeling is to bring the roots of these two trees in contact in a meaningful way.

Data model for a river water quality management decision support system is based (among others) on the following concepts:

- *a river* is a spatial structure of riverbeds, conjunctions and bifurcations, their lengths and locations of different point sources etc.
- *a riverbed* has hydraulic characteristics (roughness, slope, cross-section parameters)

- *hydrological conditions* are a set of discharge and other values (temperature, etc.) with each value tied to a certain point along a riverbed
- *point load* is a vector of variables (discharge, dissolved oxygen concentration (DO), biochemical oxygen demand (BOD), concentrations of nitrogen fractions and phosphorus fractions, etc.), describing the effluent to a certain point along a riverbed, load may vary with time, a point load is itself an attribute of a point source
- *waste water treatment plant* is a entity tied to each point load with a treatment option or a vector of reduction factors (a reduction factor for each pollutant)
- etc.

The interface between scenario analysis and the rest of the system builds on the concepts *decision, future* and *scenario.* A decision is a set of subdecisions describing the projected new waste water treatment plants and the upgrading of the old ones. Each subdecision consists of design wastewater discharge and treatment technology. If the system includes non-point source pollution, a decision includes also planned water quality management measures in the watershed. A future is a realization of hydrological and other conditions. Different kinds of conditions can be generated as a combination of flow values, background pollution, and human behavior. Human behavior affects for example diurnal variation in wastewater flow. A scenario is a set of implications of a certain decision and of a certain future. The scenario analysis part consists of two modules: 1) future generation and database, 2) scenario matrix analysis. This system is displayed graphically in figure 5.2.

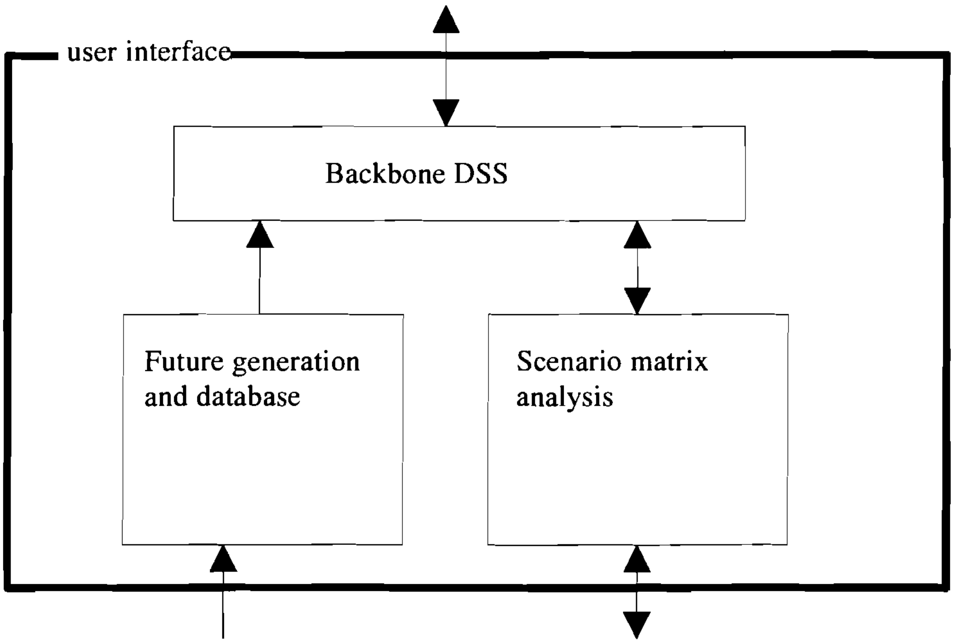

Figure 5.2 Scenario analysis modules in a DSS for regret analysis.

The backbone DSS needs to implement some commands for the user to be able to prepare his/hers work for scenario analysis, and to start a scenario analysis with all necessary information.

6. A DEMONSTRATION OF REGRET ANALYSIS IN RIVER WATER QUALITY MANAGEMENT

The regret analysis methodology described in this study was demonstrated in an experimental DSS. The DSS was applied to a hypothetical river and decision making situation, where the number of all possible decisions was very limited (64) in order to avoid the preprocessing of them. The decision making situation also had a hypothetical decision maker, who contributed by defining the water quality index (*wqi*) model (see below) and by giving her target and allowance curves for the relationships cost - water quality and future probability - water quality respectively (Figure 6.2). There was no need for a allowance curve for the relationship probability - cost because the cost of a decision was the same in all futures. The water quality index, which is defined below, is such that the bigger it is, the worse the forecasted water quality is.

The reasoning behind the target and allowance curves is the following. First, referring to earlier discussion in this paper, the criteria for which the decision maker is requested to define this kind of relationship, must be calculated by a manner he/she/they accepts. If the decision maker spends no money at all, then he, for all practical purposes, can't expect anything from the water quality. So, using criteria of this example, the point (cost=0, $wqi=\infty$) is on the target curve. If the decision maker spends an infinite amount of money he has all the right to expect the best possible water quality. So again, the point (cost=∞,$wqi=0$) is on the target curve. Between these two points the course of the curve is under the subjective judgment of the decision maker. Similarly for the probability - *wqi* curve, if the probability of a future goes to zero, the water quality requirement for that future goes to infinity. And if the probability of a future goes to one then the water quality target goes to zero i.e. is the most stringent.

The water quality index was made up of violations of DO standards in the following way. There were two DO-standards: the upper, which was 7 mg/l and the lower, which was 5 mg/l. At each time step a violation of the upper standard at one control point meant that 1 unit was added to the *wqi*. If the violation happened at two points at the same time., it meant an addition of 4 units, if at three points 9 units and so on. The violation of the lower standard was three times more costly than the violation of the upper standard.

6.1 The experimental DSS

The implementation schema of the DSS is shown in figure 6.1. The experimental DSS consists of three main parts: 1) the "backbone" DSS, 2) the water quality model StreamOx (Cook et al 1993), and 3) the scenario analysis and future database. StreamOx is a simple model to simulate DO-BOD concentrations in a river dynamically in time and space. It allows multiple point sources along one river reach. The DSS requires three input files: DDB, BODMODEL, and MODEL. The last two of these are in essence input files for StreamOx with future (conditions) and decision data left out. Data file DDB contains information from which the backbone creates the decision set. The DSS was implemented on the DOS/Windows™ platform.

The backbone of the experimental DSS acts as a surrogate for a real DSS. Its first purpose is to transmit information between input data files, StreamOx, and scenario analysis/future database. The second purpose of the backbone is to implement two loops, one within the other, to run the system for all decisions and for all futures.

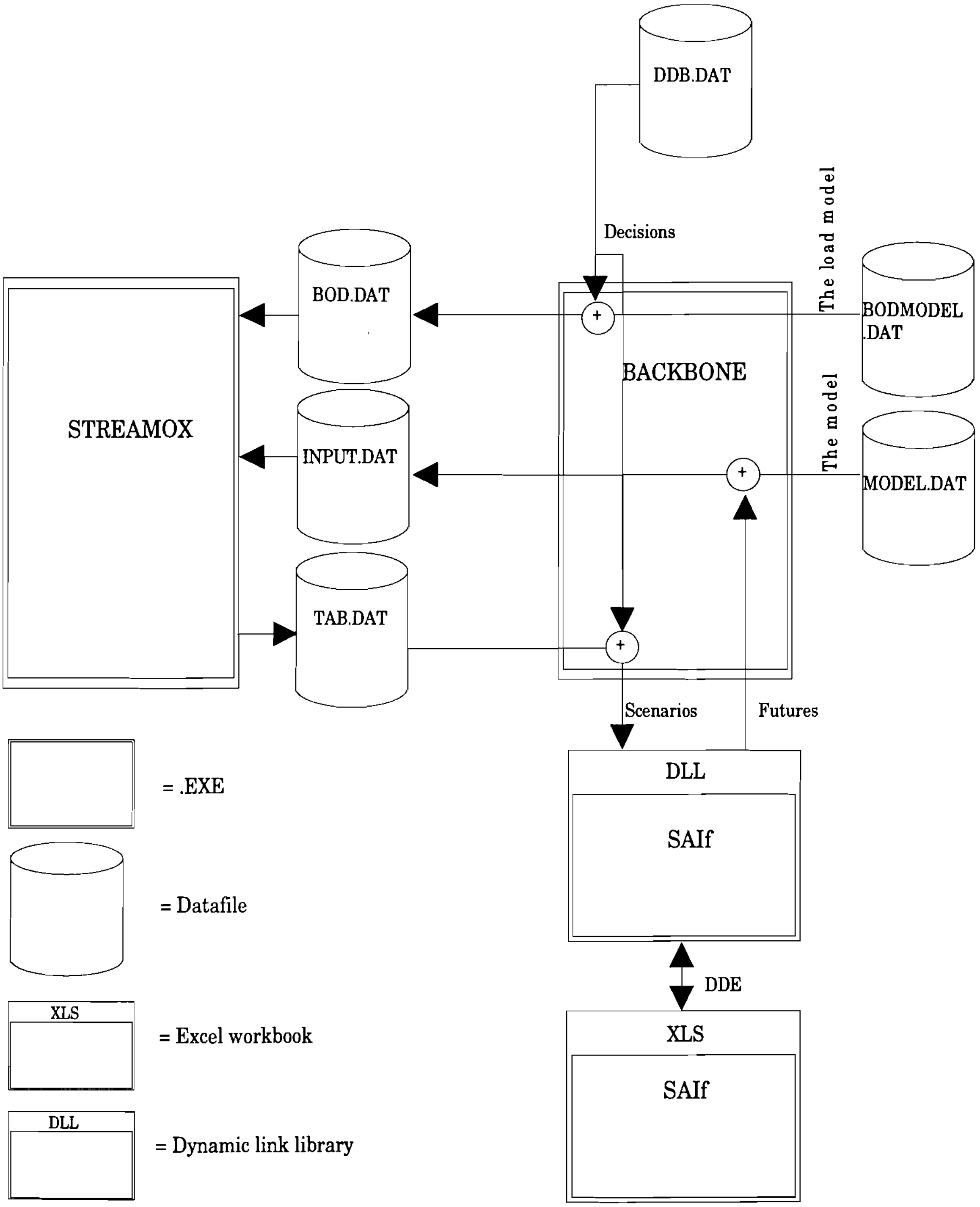

Figure 6.1 The implementation of the experimental DSS

StreamOx is a program written to simulate the interaction between dissolved oxygen and BOD in a stream with a known steady state flow regime (Cook et al 1993). StreamOx was chosen as the water quality model for the experimental DSS because of its availability, simplicity and good documentation. The underlying model is somewhat limited with its applicability. A technical problem in the utilization of the model was its interface based on data files. A functional interface would have been preferred. A functional interface can be implemented in Windows with dynamic link libraries (DLLs).

The regret analysis/future database part of the experimental DSS was implemented as an spreadsheet document and an functional interface to it was written. Excel^{TM} (version 5.0) was the spreadsheet program used. The spreadsheet workbook consisted of several work- and codesheets. These are listed in the table 6.1.

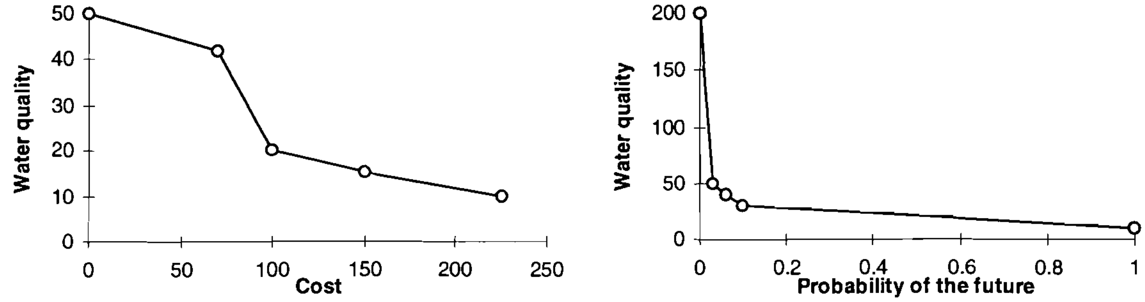

Figure 6.2 The target and allowance curves of the hypothetical decision maker.

The limitations of the chosen implementation for the scenario analysis are evident. The spreadsheet cannot effectively handle large scenario databases. This limitation could be eased if the scenario analysis database was calculated from the scenario database without intermediate steps. The elimination of the non-pareto-optimal decisions is not implemented. The reduction of the multidimensionality of the utility, which is also to a large degree implemented in the scenario model of the backbone, is very crude. The analysis could be elaborated much further than to the scenario analysis database.

A strong point of the spreadsheet implementation is its openness. All data is all the time readily available and many ad hoc plottings and analyses are easy to perform.

A future consisted of four variables which are listed in table 6.3. A scenario consisted of only two variables (Cost and WaterQuality) because there was a scenario model in the backbone-DSS which reduced the multicriteria information into only two-criteria information.

Table 6.1 Work- and macrosheets of the regret analysis and future database workbook.

Name of the sheet	type	purpose
Interface	worksheet	Slots for incoming and outgoing data
InterfaceCode	macrosheet	Implementation of the workbook part of the interface functions
Fsource	worksheet	Tables of (discrete) probability distributions of the elements of the future vector
Fgenerator	macrosheet	Macro, which creates the future database from tables of probability distributions
FutureDB	worksheet	The future database
ScenarioDB	worksheet	The scenario database
DB->matrixes	macrosheet	Macro, which creates the cost and the water quality matrix from the scenario database
Cost matrix	worksheet	The regret matrix for cost
WQ matrix	worksheet	The regret matrix for water quality
Satisfaction	macrosheet	Cost-water quality and probability-water quality target curves.
CM	worksheet	Success/failure to meet the cost-water quality target
WQM	worksheet	Success/failure to meet the probability-water quality target
matrix->SA	macrosheet	Macros to create success/failure matrixes and scenario analysis database from them
SA	worksheet	A database for probabilities of failures to meet the targets and probabilities of regrets if the target is not met

6.2 The hypothetical model

The hypothetical river was a reach with five dischargers. The setup of the system, the futures, and considered upgrading possibilities of WWTPs of the dischargers is briefly described below. The description is only for the sake of completeness.

The length of the river stretch is 59 km. It has a constant slope of 1 % and constant Manning roughness 0.03. The model (StreamOx) was run with a spatial resolution of one kilometer and with a time resolution of four hours for five days. Checkpoints were set 10 kilometers apart starting 9 km downstream from the beginning of the stretch. The upstream conditions were assumed constant with BOD concentration of 15 mg/l and oxygen concentration of 8.4 mg/l. The initial conditions for the river run are also 15 mg/l and 8.4 mg/l BOD and O_2 concentrations respectively. The model includes five BOD dischargers. Their data is shown in the table 6.2.

Table 6.2 The BOD dischargers in the model.

Name of the discharger	**Distance from the start of the reach**	**BOD 5 discharge**
Z1	0 km	constant 40 g/s
Z2	2 km	variable 0 g/s - 100 g/s
Z3	5 km	constant 20 g/s
Z4	13 km	constant 30 g/s
Z5	40 km	variable 0 g/s - 50 g/s

Variable BOD 5 dischargers follow a diurnal pattern which is shown in figure 6.3. The diurnal pattern simulates a diurnal pattern of an industrial wastewater source.

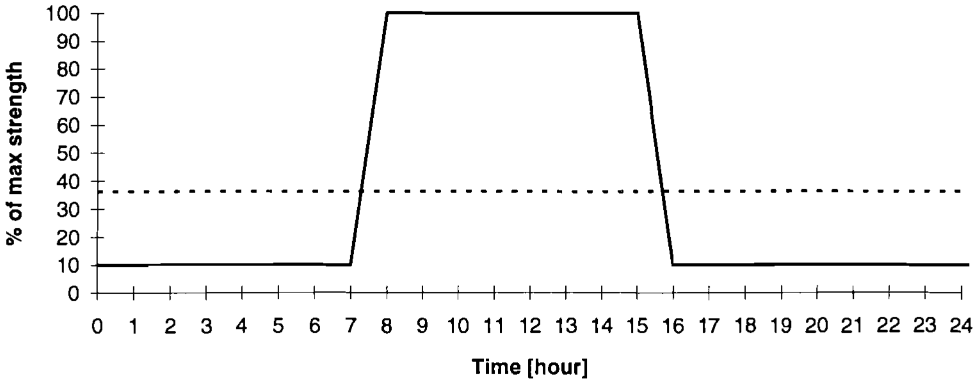

Figure 6.3 The diurnal pattern of the variable BOD-dischargers (the solid line) and a constant loading with the same total mass (dotted line).

A future is one set of values for the parameters which are listed in the table 6.3. Re-aeration rate in the reach was set constant in all futures. All considered values, i.e. the distributions are on the tables 6.4-6.7. These make up the 3*2*3*2=36 different considered futures. All distributions were assumed independent of each other.

Table 6.3 The parameters which make up a future.

	Parameter	**unit**
Q	Steady state flow	m^3/s
DISP	Dispersion coeff.	m^2/s
XK1	BOD decay rate	1/day
OSAT	Oxygen satur. conc.	g/m^3

Table 6.4 The distribution of the steady state flow.

The value	**The probability**
5 m^3/s	30 %
10 m^3/s	40 %

15 m^3/s	30 %

Table 6.5 The distribution of the dispersion coefficient

The value	**The probability**
500 m^2/s	80 %
1200 m^2/s	20 %

Table 6.6 The distribution of the BOD decay rate

The value	**The probability**
0.3 1/day	50 %
0.4 1/day	40 %
0.5 1/day	10 %

Table 6.7 The distribution of the oxygen saturation concentration DO_{SAT}. DO_{SAT} is a function of the temperature of the river water (in clean water). The use of different values for DO_{SAT} was used to not only simulate different temperature conditions but also different background pollution loading conditions.

The value	**The probability**
8.0 mg/l	60 %
9.0 mg/l	40 %

Decisions were made up from reducing the BOD load at different locations and/or removing the diurnal variation of the load (so that the total load remained the same). All possible decisions and their costs are shown in table 6.8. Different rows in this table make up all the 2*4*2*2*2=64 possible decisions.

Table 6.8 All upgrading options considered in the study. The cost figures are hypothetical.

Name of the discharger	**original BOD 5 discharge**	**discharge change options**	**total costs (investment plus variable) of changes [10^6 currency units]**
Z1	constant 40 g/s	no change / reduction to 20 g/s	0 / 10
Z2	variable 0 g/s - 100 g/s	no change / change to constant / 50 % reduction / comb. of change to constant + 50 % reduction	0 / 50 / 100 / 130
Z3	constant 20 g/s	no change / reduction to 10 g/s	0 / 10
Z4	constant 30 g/s	no change / reduction to 20 g/s	0 / 15
Z5	variable 0 g/s - 50 g/s	no change / 50 % reduction	0 / 60

6.3 The results of the experiment

On figure 6.4 there is an example of the behavior of the oxygen household of the river stretch according to the model StreamOx. The effect of daily variation is clearly visible as the formation of the steady state conditions (no day-to-day variation).

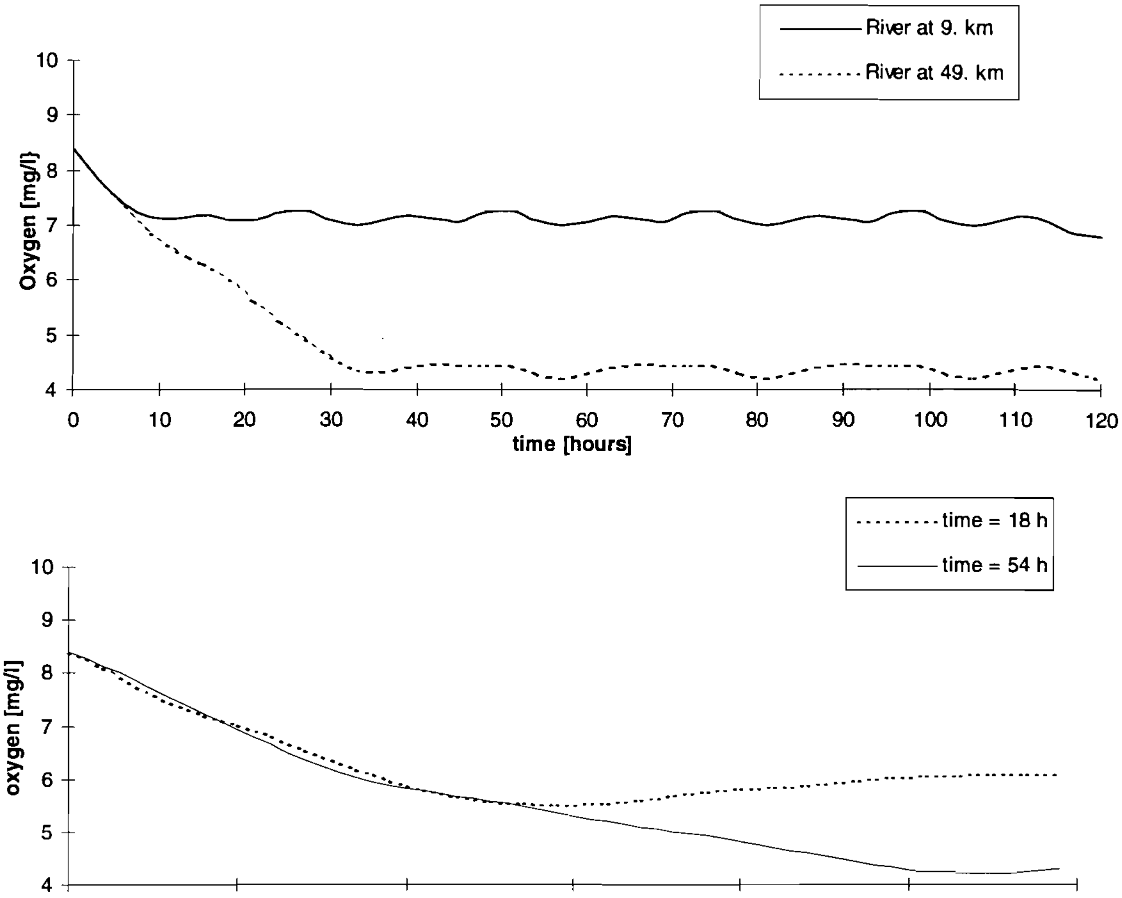

Figure 6.4 The behavior of the oxygen household of the hypothetical river stretch with decision 4 and future 1. Decision 4 means that there is reduction at sites Z1 and Z4 but not elsewhere. Future 1 means flow = 5 m^3/s, dispersion coefficient = 500 m^2/s, BOD decay rate = 0.3 1/day, and oxygen saturation concentration = 8 mg/l.

For the regret analysis a (64x36) Q matrix was calculated. The elements of Q were vectors of two Boolean values. The first element of each of these vectors indicated whether the water quality index obtained and cost generated at each scenario was above (failure) or below (success) the target level of water quality index for the respective cost defined by the decision maker (figure 6.2). The second one indicated whether the water quality index obtained at each scenario was above (failure) or below (success) the allowance level defined by the decision maker (figure 6.2) for the respective probability of the scenario.

From this Boolean information four risk probabilities were calculated for each decision:

1. probability of failure to meet at least the target level of water quality set by the decision maker for the costs generated
2. probability of regret caused by the fact that at least the target level of water quality for the costs generated was not met while it could have been met by making another decision (all 64 decisions were treated here as being equally probable)
3. probability of failure to meet at least the target level of water quality set for the probability of the scenario

4. probability of regret caused by the fact that at least the target level of water quality set for the probability of the scenario was not met while it could have been met by making another decision (all 64 decisions were treated here as being equally probable)

The ranking of the decisions based on risk probabilities 1. and 2. were almost identical with only small changes, which were not on the top of the list. Risk probabilities 3 and 4 had only about three different values and in about half of the cases they were zero. They followed each other tightly. An example plotting of all risk probabilities is in the figure 6.5.

The regret analysis was done for four cases which were identical except some changes on the cost-*wqi* target curve. The numerical changes to the (discrete) target functions, which make up the different cases are listed in table 6.10. The changes on that curve were made to investigate the sensibility of the optimal solution to this element. In three cases both decisions 15 and 16 (their *wqi*s were identical) got all smallest risk probabilities. Risk probabilities of decisions 11 and 12 (also their *wqi*s were identical) were among the smallest in these cases and their risk probabilities were smallest in the fourth case. Risk probabilities of decisions 15 and 16 were among the smallest in the fourth case.

Table 6.10 Cost vs. *wqi* target function in different cases.

Cost:	***wqi* target**			
	case 1.	**case 2.**	**case 3.**	**case 4.**
0	50	40	40	40
70	42	35	35	28
100	20	28	20	20
150	15	15	15	15
225	10	10	10	10
Optimal decision:	15,16	15,16	15,16	11,12

Decisions 11, 12, 15, and 16 were similar (Z2=change to constant, Z3=reduction to 10 g/s, and Z5=no change) except for the case of discharger Z1, reduction at which had no effect on *wqi* in these decisions, and for the case of Z4, discharge of which was reduced to 20 g/s in the decisions 15 and 16.

The case, where decisions 11 and 12 were deemed better, had *more* stringent water quality target for the cost than the cases, where decisions 15 and 16 were deemed better. This might seem to be wrong since decisions 15 and 16 include more BOD reductions. But at the same time decisions 15 and 16 are more expensive, which might lead to failure more often in the sense of cost/achieved water quality.

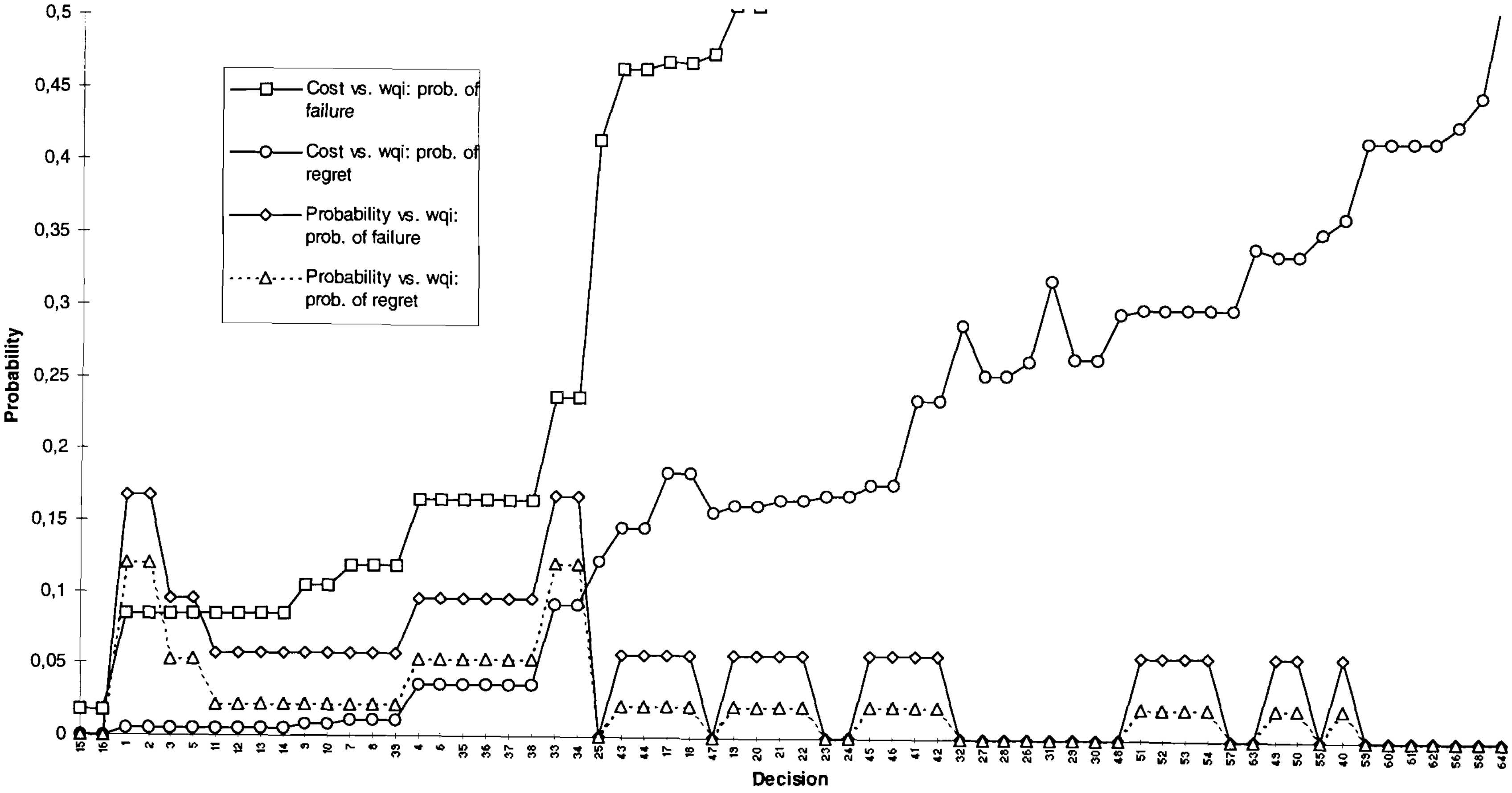

Figure 6.5 An example plotting of the risk figures (see text) vs. different decisions. The plotting order is determined by risk probability nr. 1.

7. DISCUSSION AND CONCLUSIONS

Regret analysis as described in this paper makes up a tool or procedure, which offers an interesting view into the uncertainties and risks associated with decisions. It is a tool which should be used after a candidate set of decisions is generated with single criteria optimization, simulation, or by some other means. The optimal decision in the problem described in previous chapter was decided upon solely based on risk information. This is probably the most useful information in a real decision making situation. However, regret analysis is only one of all analyses which a decision maker should utilize.

One of the fundamental ideas of decision support with systems like the experimental one described here is the idea of a simulation model of the real system as a tool. A "tool" in the word's most every day sense: totally under the control of the user, a means to gain insight into the real system, its dynamics and sensitivities. The regret analysis approach described in this paper combines the two aspects of stochasticity: 1) uncertainty of the right structure and parameter values of the model and 2) randomness of the nature. If there were more than one simulation model in the DSS, the conceptual difference between these two aspects should be reflected in the system design. The "tool" approach enforces this.

The water quality management problem - like many other management problems - has basically two dimensions: 1) the economic criteria and 2) the environmental criteria. The means to reduce the dimensionality of the problem can be called as a decision model. The decision model must be built with the input from the decision maker(s). A key issue in this is the interaction between the decision maker and the analyst who builds the decision model into the analysis. The decision model is not a single module in the system, decision maker's objectives and preferences should be reflected in the whole system, in all its modules.

Decisions should be made on the basis of information about expected value and risk associated with the decision. The notion of this paper is that there are two aspects within risk: the risk of failure and risk of regret. The exact qualitative and qualitative meaning (definition) of these is case specific, an example of this is given in the chapter 6. These definitions should be written with the information obtained from the decision maker(s) and they should be worked to the point where they can be used to calculate a small amount of numerical information of each considered decision.

The data of the real system, whose management is the decision problem, is by nature case specific, scattered and therefore unorganized. Models, on the other hand, are generic. Data modeling closes this gap. There are two main tasks in data modeling: 1) finding and defining of the key concepts and 2) organization of the data according to the conceptual structure. In the light of the experimental DSS described in this paper these tasks are highly significant.

8 REFERENCES

Burn, D.H. and Lence, B.J. 1992. Comparison of Optimization Formulations for Waste-Load Allocations. *Journal of Environmental Engineering*, **118**, 597-612.

Cardwell, H. and Ellis, H. 1993. Stochastic Dynamic Programming Models for Water Quality Management. *Water Resour. Res.*, **29**, 803-813.

Cook, C., Copeland, S., and Yokoyama, K. 1993. STREAMOX: Program to compute dissolved oxygen concentration and biochemical oxygen demand in river reaches. Version 1.0 6/14/93. Program users manual. University of California, Davis. 53 p.

Crousillat, E.O., Dörfner, P., Alvarado, P., and Merrill, H.M. 1993. Conflicting objectives and risk in power system planning. *IEEE Transactions on Power Systems*, **8**, 887-893.

Loomes, G., and Sugden, R. 1982. Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, **92**, 805-824.

Somlyódy, L. 1993. Quo Vadis Water Quality Management in Central and Eastern Europe. WP-93-68. International Institute for Applied Systems Analysis. 12 s.

Somlyódy, L., Masliev, I., Petrovic, P., and Kularathna, M. 1994. Water Quality Management in the Nitra River basin. Collaborative Paper. International Institute for Applied Systems Analysis. 205 p.