Working Paper

A Prototype of a Decision Support System for River Basin Water Quality Management in Central and Eastern Europe

Rainer Berkemer, Marek Makowski, David Watkins

WP-93-049
September 1993

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

A Prototype of a Decision Support System for River Basin Water Quality Management in Central and Eastern Europe

Rainer Berkemer, Marek Makowski, David Watkins

WP-93-049
September 1993

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Foreword

The research described in this Working Paper has been performed by two participants of the Young Scientists' Summer Program 1993 and one researcher of the Methodology of Decision Analysis (MDA) Project, in collaboration with researchers and YSSPers of the Water Resources Project. Although the three-month duration of the YSSP limited the scope of the research reported here, there are two main reasons for publishing the outcome of this research: to document the work that has been done, and to provide a basis for discussing the possibility of further research in this area at IIASA.

The data used in this research are preliminary, since data analysis and verification of data are on-going tasks of the Water Resources Project. Thus, the results discussed in this paper serve only to illustrate the capabilities of the applied methodology and the developed tools. However, well documented formulation of the underlying mathematical programming model and of the data used are essential for using the reported results at IIASA and for discussing the possible continuation of this activity.

Abstract

This Working Paper documents the implementation of a prototype of a Decision Support System (DSS) for regional water quality management applied to a case study of the Nitra River in Slovakia. With the goals of flexibility and simplicity in mind, two different approaches and tools have been implemented and tested. First, the object-oriented development tool ORVAN was used for fast prototyping of the mathematical programming model and for scenario analysis. Second, a problem-specific generator was implemented to generate various single criterion and multiple criteria optimization problems useful in examining the water quality problem. The resulting mixed-integer optimization problems were solved by the MOMIP package.

Provided in the paper are the following: a complete formulation of the mathematical model, a detailed discussion of the data used, documentation of the developed software, an overview of interesting results, and recommendations for future work. Since only preliminary data were available at the time of performing the reported research, results are given merely as illustration of the methodology and software and should not be considered policy recommendations. For the latter task a verified data set and water quality model will be required.

Contents

A Prototype of a Decision Support System for River Basin Water Quality Management in Central and Eastern Europe

Rainer Berkemer[*], *Marek Makowski*, *David Watkins*[**]

1 Introduction

Surface water quality in many Central and Eastern European countries is generally quite poor, and the cost of cleaning up the rivers in this region is estimated to be enormous. In light of increasing municipal discharges and severe financial constraints in these countries, imposing "best available technology" standards or ambient water quality standards typical of Western Europe and North America is not feasible in the near future. However, there is a need to set strategies which are feasible in the short term and also consistent with long-term planning (when implementation of higher quality standards may be feasible). Thus, decision-makers need to evaluate the trade-offs among capital investment, treatment costs, and ambient water quality with respect to a number of constituents.

The activity reported by this paper is just one element of a research effort undertaken by the Water Project at IIASA. Description of this research is beyond the scope of this paper, and the reader is advised to consult [26, 27, 28]. However, we would like to stress that results reported in this paper have been obtained in close cooperation with, and using results of, our colleagues in the Water Project.

The scope of our problem is a river basin or a larger region composed of several basins where the water quality is extremely poor. We consider also a set of waste-water emission points, at which a waste-water treatment plant either exists or could be constructed or upgraded. At each emission point, one technology (to be selected out of the given set of possible technologies) can be implemented in order to improve the water quality in a region. The traditional approach to solving such a problem consists of looking for a set of plants and technologies whose implementation would result in maintaining prescribed water quality standards at the minimum costs. However, such an approach would likely result in an infeasible solution because of the costs involved. Therefore, another approach to the design of a decision-support system (DSS) for this purpose must be taken.

The models and software tools being developed by the Water Resources Project are envisioned to serve two purposes (cf [27]):

[*]Participant of the Young Scientists' Summer Program 1993 at IIASA, current affiliation: Institut für Textil- und Verfahrenstechnik, (ITV), Gruppe Textil-Management, D-73770 Denkendorf, Germany.

[**]Participant of the Young Scientists' Summer Program 1993 at IIASA, current affiliation: Department of Civil Engineering, University of Texas, Austin, TX 78712.

- As a decision support tool for high-level decision makers to establish the effluent and/or ambient water quality standards and the associated appropriate economic instruments that can be enforced to control the waste-water discharges,
- To aid in the evaluation of alternative treatment strategies (technologies in treatment plants) and/or in selecting the most appropriate strategy based on water quality standards and costs (capital investment and operational).

With these purposes in mind, we propose a prototype decision-support system consisting of an object-oriented development tool and a problem-specific generator which allows for the generation (and solution by a general purpose modular MIP[1] solver) of relevant single and multiple criteria optimization problems. Such tools could be highly beneficial to water quality decision-makers in Central and Eastern Europe by enabling scenario analysis, the evaluation of trade-offs among several criteria, and heuristics implementation.

This prototype differs from traditional single-criterion optimization approaches, and also from more recent expert system/AI approaches, though it has the potential to incorporate expert knowledge as the water quality management problem becomes better defined. Attractive traits of object-oriented programming (OOP) and multicriteria decision aid (MCDA) are simplicity and flexibility, which allow the user to learn about the decision situation during the process of decision-making. One possible disadvantage of the implemented approach is that simple, linear models are used rather than more physically-based (i.e., non-linear) ones. However, the usefulness of complex models for management may be limited by uncertainty in system identification and a lack of high quality field data (cf [8, 27, 28]).

Using OOP and MCDA, the aim of the research reported in this paper is twofold:

- To implement one of several possible mathematical programming models and to provide a means for comparison of the results obtained from this model with results from other models implemented at IIASA for the same case study (cf e.g. [24, 26, 27, 28]).
- To compare two parallel implementations (i.e., with the same model formulation and data): (1) using the object-oriented development tool, ORVAN, for fast prototyping of the model formulation, data handling, and comparative scenario analysis; and (2) the problem-specific generator and MIP solver to evaluate trade-offs in single criterion and multiple criteria analyses.

In Section 2 of the paper, a mathematical formulation of the problem is detailed, along with model assumptions. This formulation is in a form most suitable for checking the correctness of the model. However, this form is different from the standard formulation of optimization problems, so a reformulation which corresponds to an equivalent MPS standard form is provided in Section 9.

In Section 3 a method of examining the model is presented, followed by discussions of the application of ORVAN and optimization in Sections 4 and 5, respectively. Section 6 contains a description of the data used. In Section 7 the results of the applications are discussed, and conclusions and recommendations for future work are given in Section 8. Finally, Section 9 contains the reformulation of the model.

2 Model formulation

There are many ways of formulating a water quality management problem (cf [14, 24, 26]). They can be grouped into two basic approaches: (1) to use scenario analysis (i.e. simulation of results of assumed decisions), or (2) to optimize with respect to a selected criterion

[1] Mixed-Integer Mathematical Programming Problem.

while meeting prescribed constraints. For instance, one can minimize costs subject to water quality constraints, maximize water quality subject to a budget constraint, or find a Pareto efficient solution using a multiple objective programming method for different combinations of criteria (performance indices) and values of reference points. There are also a number of variations for each of these three alternatives, including the minimization of investment and annual costs, as well as the maximization of several water quality criteria.

However, both approaches (scenario analysis and optimization) require the definition of a *core* model which relates waste water emissions, treatment decisions, and the resulting water quality. Specification of the *core* model and the data used for it are obviously of critical importance for the quality of any model-based DSS. Therefore, to provide a basis for possible future work, we specify in detail both the mathematical formulation and the data used in this model.

2.1 Assumptions

After discussions with our colleagues from the Water Project, examination of available data and of alternative model formulations, and consideration of the time available and the purpose of the research, we have adopted the following assumptions in the model presented:

1. Our water quality simulation model is very simple. Regarding the hydraulic of the river, we employ a steady-state formulation, using a "critical" low-flow measured in August 1992. We assume complete mixing at each emission point and tributary confluence, and we assume plug flow along the river. For the mass balances of each constituent, we basically follow Thomann and Mueller [31]. We assume first-order decay and apply the extended Streeter-Phelps equation for dissolved oxygen. In this model we neglect BOD (Biological Oxygen Demand) settling and use a simplified nitrogen model. In order to maintain a linear system, we use a simple dilution model for dissolved phosphorous, and we neglect the effects of algae and phytoplankton. Finally, we neglect evaporation and sediment nutrient exchange.
2. We consider six water quality constituents. The following values of subscript l are used for the respective constituents:
   0. DO, dissolved oxygen
   1. CBOD, carbonaceous oxygen demand
   2. NBOD, nitrogeneous oxygen demand
   3. NH4, ammonia
   4. P, dissolved phosphorous
   5. SOD, sediment oxygen demand.
3. For a given river system there is a set of points (given implicitly by the provided data), each of which is at least one of the following:
   - Emission point: waste water is discharged at this point. The amount of discharged waste depends on the treatment technology chosen in the decision process. These include industrial discharges and junctions with minor tributaries, where only the "no treatment" option is currently considered. These also include points where water is apparently being extracted from the river. At these points one can consider a "negative" emission, whereby the constituent loads are reduced proportionally to the reduction in flow.
   - Monitoring point: water quality is compared to given standards at this point.

- Confluence point: junction point of two rivers. Constituent loads are the sum of loads from both rivers.
- Other points: points for which hydrological data exist and therefore new transfer coefficients are calculated. The load of waste does not change at these points.

Each of these points is called a *node*, denoted by the subscript j. In every node the equations that define water quality are defined. This enables us to employ a formulation for the mass balances of constituents at each of these points on the river.

4. For every emission point, one emission source is assumed[2]. At each emission node there exist a number of technology options, denoted by the subscript k. Included in each of the k treatment technologies are the option of no treatment (with raw waste concentrations and no cost), as well as the option of maintaining the existing technology (with the operating cost but no investment cost).
5. For monitoring points, standards for constituents 0, 1, 2, 3, and 4 (above listed) may be set (if no standard is given for a particular node then a default standard is applied). For each monitoring node, a variable which corresponds to a relative violation of the standard is defined. A matrix of such variables (rows corresponding to nodes, columns to water quality constituents) may then be used for various analyses of the resulting water quality.

2.2 Decision variables

The decision variables (which are set in scenario analysis via simulation and are computed in optimization) are the treatment technologies to be implemented at the j-th node where waste-water emissions occur. Let these be denoted by x_{jk}, where j is the index of an emission node and k is the technology choice. Since only one technology can be implemented at each point, we impose the following constraint:

$$\sum_{k \in K(j)} x_{jk} = 1 \qquad x_{jk} \in \{0,1\}, \qquad j \in E \tag{1}$$

where $K(j)$ is a set of technologies considered for the emission node j, and E is a set of nodes where emissions occur.

2.3 Endogenous decision variables

The following quantities are assumed to be given:

1. $\overline{TAC}$ - maximum value of the total annual cost
2. $\overline{IC}$ - maximum value of the total investment cost
3. $\overline{g_{all}}$ - maximum value of the water quality index
4. $\overline{g_l}$ - maximum value of the water quality indices for the l-th type of waste. In the current implementation such maxima can be set for $l = 0, 3$ (DO and NH4).

Such values are used as "hard" constraints for corresponding quantities in single criterion optimization. Although formally these constraints are assumed given (as in a corresponding classical single-criterion formulation), they are in fact decision variables, and their values are of critical importance for the existence of a feasible solution. Note however, that these values are not used in multicriteria optimization, in which corresponding *aspiration levels* (reference points) are used instead.

[2]This has been done to simplify the description. Actual implementation can easily be modified to accommodate any number of sources in a single point.

2.4 Auxiliary variables

Auxiliary variables are the model quantities whose values depend on the values of decision variables. Other model quantities (called parameters and defined in Section 6) are assumed to be given or to be calculated from the provided data. This distinction is necessary since parameters are actually computed before scenario analysis starts.

Auxiliary variables are defined to ease both the problem formulation and the interpretation of results. Not all of the auxiliary variables will be used in every possible formulation of the model (cf Section 3), but for consistency all variables that might be used are defined below. They are divided into two groups: variables related to water quality and variables related to costs.

2.4.1 Water quality variables

- At each monitoring point a vector wq_j of water quality indices is defined as:

$$wq_{j0} = (aqs_{j0} - aq_{j0})/aqs_{j0} \qquad j \in M \tag{2}$$

$$wq_{jl} = (aq_{jl} - aqs_{jl})/aqs_{jl} \qquad l \in [1,4] \qquad j \in M \tag{3}$$

  where aq_{jl} (defined by (6) or by (7)) is the ambient concentration of the l-th constituent at node j, and aqs_{jl} is the corresponding water quality standard; and set M contains indices of monitoring nodes. Note that the water quality index for DO (dissolved oxygen) is defined by eq. (2) in a different way than indices for the other constituents defined by eq. (3). Since DO should maximized while other ambient concentrations should be minimized, such an approach allows for minimization of all water quality indices, thus simplifying various model formulations.
- One can consider a maximum violation of a standard for a particular constituent, namely:

$$g_l = \max_{j \in M}(wq_{jl}) \qquad l \in [0,4] \tag{4}$$

- Additionally, one may wish to consider an aggregate index of regional water quality which can be defined as

$$g_{all} = \max_{l \in [0,4]}(g_l) \tag{5}$$

  Indices defined by (4) and (5) - if positive - show a maximum relative violation of water quality standards for the l-th constituent or for all constituents, respectively. A negative value of g_l or g_{all} indicates that water quality standards are observed at every monitoring location, and the corresponding absolute value in such a case shows the relative "margin of safety" of water quality at the worst monitoring location.
- The ambient concentration of DO (denoted for j-th node by aq_{j0}) is affected by several constituents, as well as by the saturated dissolved oxygen concentration. DO is given by

$$aq_{j0} = DOsat_j - 1./Q_j * \sum_{i \in I(j)} Q_i * \left(TC_{i0}(DOsat_i - aq_{i0}) + \sum_{l \in \{1,2,5\}} TCp_{il} aq_{il} \right) \tag{6}$$

  where aq_{il} is defined by eq. (7), and the remaining right hand side quantities are given (or computed from given data - cf Section 6): $DOsat_j$ is DO saturation level at j-th node, TC_{il} and TCp_{il} are transfer coefficients, and Q_j is a river flow at j-th node. The set $I(j)$ contains indices of nodes located immediately up-stream of j-th node (this set contains two elements for confluence nodes and one element otherwise). Note that indices 1, 2, and 5 correspond to CBOD, NBOD and SOD, respectively.

- Ambient concentrations of other constituents (denoted by aq_{jl}) are defined by:

$$aq_{jl} = b_{jl} + (\sum_{i \in I(j)} TC_{il} aq_{il} Q_i + e_{jl})/Q_j \qquad l \in [1,4] \tag{7}$$

In this equation, the first term represents the background concentration of constituent l at node j, which accounts for non-point or non-controllable source pollution. The second term represents the load of the constituent l coming from upstream, so i is a member of the set of upstream points $I(j)$. Thus, Q_i is the flow at point i, and TC_{il} is a dimensionless transfer coefficient for constituent l in a segment from node i to the nearest node downstream, j. This formulation is needed to accommodate flow from tributaries. Note that the eq. (7) is formulated with the assumption that Q_j accounts for the waste flow[3]. If this assumption is not true then the equation (7) will be replaced by:

$$aq_{jl} = b_{jl} + (\sum_{i \in I(j)} TC_{il} aq_{il} Q_i + e_{jl})/(Q_j + q_j) \qquad l \in [1,4] \tag{8}$$

- At each emission node j there is a given waste flow rate, q_j [m^3/day], and the water quality constituent concentrations resulting from the implementation of the k-th technology, em_{jkl} [mg/l]. The emission of the l-th constituent at the j-th node is denoted by e_{jl} and is defined by:

$$e_{jl} = q_j \sum_{k \in K(j)} x_{jk} em_{jlk} \qquad l \in [1,4] \qquad j \in E \tag{9}$$

Note that – due to the eq. (1) – for each j exactly one out of $K(j)$ binary variables x_{jk} will be equal to one while the others will be equal to zero.

2.4.2 Cost variables

Corresponding to the k-th treatment technology implemented at the j-th node are the investment cost IC_{jk} and the operating and maintenance cost OMC_{jk}. Included in the technologies are the option of no treatment (with raw waste concentrations and no cost) and the option of maintaining the existing technology (with O&M cost but no investment cost).

- The investment costs Inv_j for the j-th emission point are defined by

$$Inv_j = \sum_{k \in K(j)} x_{jk} IC_{jk} \qquad j \in E \tag{10}$$

- The O&M costs OM_j are given by:

$$OM_j = \sum_{k \in K(j)} x_{jk} OMC_{jk} \qquad j \in E \tag{11}$$

- The total annual cost (TAC) of each technology is defined as

$$TAC_j = [r(r+1)^n/((r+1)^n - 1)] Inv_j + OM_j \qquad j \in E \tag{12}$$

where r is a given interest rate, and n is a given capital recovery period.

[3]This assumption has been adopted for the data currently used.

• Finally, one may want to consider the sums of respective costs for the whole region:

$$Tot\_Inv = \sum_{j \in E} Inv_j \tag{13}$$

$$Tot\_OM = \sum_{j \in E} OM_j \tag{14}$$

$$Tot\_TAC = \sum_{j \in E} TAC_j \tag{15}$$

3 Examination of the model

3.1 General remarks

The model defined in Section 2 can be used for several purposes. The ultimate goal of this research is to create a DSS which is helpful to the decision-maker, but at the current stage the model should be used only for the analysis and data verification. For these purposes, available software (ORVAN, MOMIP, and GNUPLOT) has been used, and new software has been written (problem generator, postprocessor, and an interface for making plots) in order to provide a flexible framework for scenario analysis.

Scenarios can be generated either by a "manual" selection of decision variables or by solving an optimization problem. For manual selections, the implementation of heuristics can be helpful. Examples of possible approaches are discussed in Section 4.

Different formulations for optimization problems (both single- and multiple-objective) are generated by a problem-specific generator. We want to stress that optimization is not considered as a tool which provides *"one best solution"*. On the contrary, the generation, solution, and examination of various formulations help to identify factors and scenarios which should be considered in the decision process.

Tools for examining results are currently very simple. One can obviously examine complete solutions (i.e. values of all variables listed in Sections 2.2 and 2.4). Additionally, a simple tool has been developed for plotting the resulting ambient concentrations at each node and for each constituent. Examples of such plots (which can be examined on an X11 terminal and stored in the Postscript format) are provided in Section 7.

3.2 Organization of software and data

3.2.1 Model modification and scenario analysis by ORVAN

ORVAN can be used for three main purposes:

1. Modification and formal analysis of the data. As mentioned, ORVAN makes a simplified formal analysis of the provided data (e.g., it checks for the uniqueness of the node identification numbers and for the presence of flow data at each node) and provides a simple way to modify data. ORVAN also generates a free-format data file that is easier to read and modify by an editor.
2. Fast prototyping of the model formulation and its modification. ORVAN was used for fast implementation of the equations formulated in Sections 2, for data conversion (cf Section 6) and for checking the consistency of the model with the observed data.
3. Scenario analysis using simulation. One can easily select a set of technologies at different locations and examine the resulting water quality and costs. For such an analysis, performance indices (cf Section 3.3) are helpful.

3.2.2 Optimization

Though one can consider many different optimization problems (cf Section 5), each boils down to a mixed integer programming problem. This is because the decision variables are 0-1 variables – either a particular treatment technology is selected or it is not. Therefore, a problem-specific model generator (subsequently referred to as *the generator*) has been implemented. The generator allows for the formulation of either a single criterion problem or a multicriteria problem (using the reference point method). In the latter case the generator facilitates the conversion of a selected multicriteria problem to an equivalent single-criterion problem with the help of the scalarizing function (cf e.g.[15]). A general purpose modular MIP solver MOMIP (cf [21]) is being used to solve the resulting MIP problem.

3.2.3 Data

The data has been assembled from different sources (cf Section 6) and has been combined in one free–format ASCII file. The data file is composed of several segments containing groups of related data and a description of data items. The organization of the data file is flexible and provides adequate documentation so that its organization is to easy to modify. Moreover, both ORVAN and the problem generator can be easily adapted for such modifications.

Upon being read by ORVAN, a formal diagnostic of the provided data is made, and a processed file (which is also a free-format ASCII file with a description of records) is generated. ORVAN also provides an easy way for modifying data, which can be done with any text processor. In the future, a user-friendly interface may be developed for modifying the data file.

This data file is also being used by the problem-specific generator to generate (for a selected type of optimization problem – cf Section 5) either an MPS file or an equivalent binary communication file, which contains data corresponding to a mixed integer programming (MIP) problem.

3.3 Performance indices

Performance indices serve for comparing the solutions obtained for different formulations of scenarios (i.e., different selections of treatment technologies) examined with ORVAN or obtained as a result of solving an optimization problem. For a single criterion optimization, one such index is selected to be the objective function, while constraints are usually set for some of the other indices (cf Section 5.2).

Possible performance indices are the following:

- Violation of standards set for water quality constituents (such as DO, CBOD, NH4, and P) among the set of monitoring points; this is equivalent to minimizing $wq_{jl}, j \in M$ (eq. 4) for the l-th constituent,
- The regional water quality index g_{all} (eq. 5)
- The total annual cost for the whole region Tot_TAC (eq. 15).
- The total investment cost for the whole region Tot_Inv (eq. 13).

In order to facilitate both the formulation and analysis of the model, all indices have been defined in such a way that a smaller value is preferred to a larger one.

4 Application of OOP

4.1 General remarks

The motivation for using Object-Oriented Programming (OOP) is threefold. First, OOP provides a developer of a DSS tool with a natural representation of an application area. Second, OOP enables the "rapid prototyping" of such DSS tools. Third, it is much easier – compared with conventional programming languages – to modify data structures and to consider modifications of the problem formulation.

In order to obtain a natural representation of an application area, it is necessary to assign an adequate object to each relevant element of the problem area. Accordingly, we have defined objects for each river, for each emission, monitoring and tributary point, for the technologies available at each emission point, and for the relevant water quality constituents.

Classes, which are generalizations of single objects, allow the ordering of the elements of the model. The result is an understandable object structure and a model which is similar to the way the user is thinking. This helps to close the gap between user and developer, supporting necessary discussions about assumptions and other decision-making issues.

The ability to define classes and to apply the related *"inheritance principle"* accounts for most of the capability to develop prototypes in a relatively short time. For instance, one can provide all of the necessary variables (according to the problem formulation Section 2) for a class called "points". The equations (6) and (7) for calculating concentrations are implemented for this class as well. Furthermore, one can define a subclass "monitoring-points" for which the equations (4) and (5) for calculating the water quality indices are implemented. Then, for a subclass "emission-points" another method may be implemented which asks the user to choose a technology, and for another subclass "tributary-points" a method can check the consistency of flow data. This kind of approach is also called *"programming by difference"*. It means that most of the program can be written generally, and then modified or extended in only a few locations of the code.

In each program some general functions are needed, such as importing files, viewing results, and printing lists to file or printer. Since these functions are implemented in a general way (for both classes and subclasses), it is probable that one can find helpful classes in already existing OOP applications, which can be considered *"software ics"*. Therefore, it is possible to speed up the development of a DSS tool even more if some of these *"software ics"* can be imported from other applications.

Obviously, this kind of programming leads to a modular software design in which redundancy of program code is minimized. Thus, making changes or extensions is less complicated. Changes are normally only necessary at a few key points, and extensions are often possible just by exchanging one module for a more powerful one. Of course, a good programmer is able to develop a modular program with any structured programming language, but in OOP even a poor programmer must make efforts to destroy the modular structure.

4.2 Functional description of ORVAN

The object-oriented development tool ORVAN (cf e.g. [22]) was used for the reported research. ORVAN was developed at the ITV Denkendorf (Germany) by Christoph Plapp from 1987-1990. The name "ORVAN" is derived from **O**bjects **R**elations **I**nheritance (in

German: Vererbung) **A**ttributes **M**essages (in German: Nachrichten).

This tool has been used predominantly to develop several DSS prototypes in the area of production planning. We believe it can be useful for water quality management as well, though there are several disadvantages. First of all, ORVAN runs on the operating system "OS/2", which is currently not available at IIASA[4]. Also, the full capability of ORVAN could not be exploited since fewer "software ics" were available for this new application area. Finally, ORVAN is not yet connected with a suitable graphic tool.

Despite these limitations, ORVAN is quite a powerful tool, especially when one considers that it has just 3 man-years of development. One powerful feature of ORVAN is that it can handle multiple inheritance, which means that an object can be assigned to several classes and can inherit attributes (variables) accordingly. This is very convenient for our application because a point might have several "class" properties (e.g., a point in our system might have properties of both an emission and a monitoring point). ORVAN supports not only structural inheritance but value inheritance as well. This is useful for providing default values for missing data.

Another powerful feature of ORVAN is that relation names are free definable. For instance, one can define relations like:

- tech_0 is_available_at H_V3975DVA
- tech_1 is_available_at H_V3975DVA
- tech_2 is_available_at H_V3975DVA
- tech_1 is_chosen_at H_V3975DVA
- H_V3975DVA is_controlled Emission_Point

Hence, ORVAN provides the capability of "semantic nets," which make the object structure even more understandable. Such a structure is demonstrated in Figure 1.

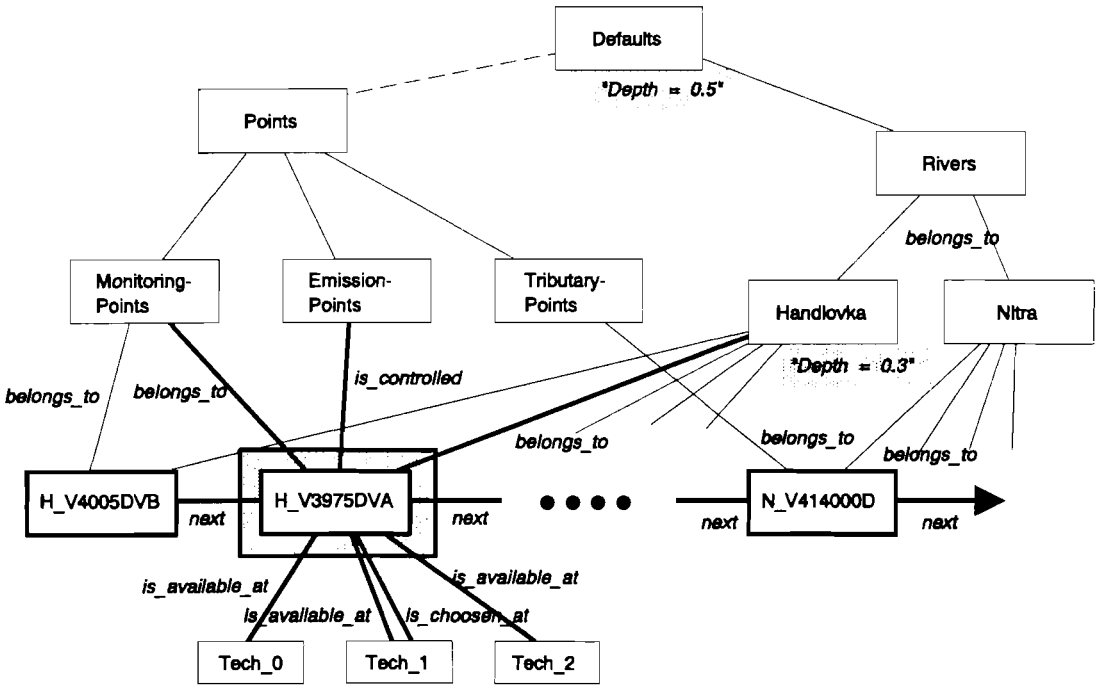

Figure 1: Semantic nets implementation in ORVAN.

[4]The computing environment at IIASA is based on the network of Sun Workstations and IBM compatible MS-DOS personal computers.

In Figure 1 objects are indicated by boxes and relations are indicated by lines. The object "H_V3975DVA", which is a point on the river system, and its relations (thick lines) are highlighted. At the same level, one can see other points which are connected by a relation named "next", which models the topology of the river system (the arrow at the right side may indicate the flow direction). At the level above, one can see the connections to the relevant subclasses of "Points", which enable the inheritance of different properties. Furthermore, one can easily see which river "H_V3975DVA" is on, which technologies are available at this location, and which of them is chosen.

It is also easy to provide river specific default values for depth or velocity. For instance, the general default value for depth may be 0.5 m. It is defined at an object "Defaults", can be inherited through the "river hierarchy", and can be redefined at the object "Handlovka" with a more specific value (say, 0.3 m). Alternately, the developer may want other default values, which are not "river specific", to be inherited through the "point hierarchy". This is indicated by the dashed line connecting "Points" and "Defaults". Thus, the inheritance capabilities of ORVAN are quite powerful and flexible, primarily due to multiple value inheritance and the opportunity for flexible exclusion of attributes from inheritance.

ORVAN has several other features which support the developer and the user. For instance, ORVAN provides the developer with powerful tracing and debugging opportunities. ORVAN supports the user with an interface consisting of menus, in which one can choose among various options, and masks, in which one can edit information. Specific help messages can be displayed in each menu or mask, making ORVAN a highly interactive system.

4.3 Scenario Analysis

Scenario analysis is understood here as the examination of consequences (represented by performance indices (cf Section 3.3)) of different selections of decision variables. There are several reasons why a decision maker should have the opportunity for scenario analysis.

First, it is necessary to verify the implemented model. For this reason, it is helpful to create a "baseline" scenario, in which the existing technologies are considered. We have done this and have compared the result with the expectations of the Water Project researchers.

Second, one should check the qualitative behavior of the model by performing a sensitivity analysis. For instance, one should expect that modifications of the reaeration coefficient have significant effects on the calculated DO concentrations. Similar effects (but perhaps less substantial) should be obtained by changing other parameters, such as the CBOD removal rate or flow rates. We have done this as well, and the qualitative behavior is correct (cf Section 7).

Third, scenario analysis provides a rational way of using any DSS. Though optimization results can be a good starting point for scenario analysis, it is not reasonable to provide a decision maker with one solution and to tell him/her that this is the "optimal" decision. If the decision maker has preferences as to what should be done at certain nodes, it would be very difficult to explain to him/her the advantage of another solution. This is particularly true if he/she is aware of the uncertainties involved, and/or if there are no obvious choices of optimization criteria or values of constraints (cf Section 5). In such a situation it is helpful to view the consequences of each acceptable decision combination. Therefore, we have also provided the capability for pairwise comparison of solutions.

4.4 Heuristics

Assume a decision maker is provided with an "optimal" solution or, more precisely, with a solution calculated by an optimization tool (e.g., with respect to one of the formulations in Section 5). For any number of reasons – which cannot be incorporated into the model – the decision maker might not agree with the selected technology at a specific emission point and would like to choose a cheaper technology instead. In order to avoid the violation of ambient water quality standards, an adequate improvement at one or several other emission points upstream would be necessary. If there are only a few emission points to consider, a rational solution could be found by "trial and error". But if there are many alternatives available, the problem becomes much more complicated, and a support tool for selecting other technologies is necessary.

For this reason, heuristics implementation in ORVAN might be very helpful. Outlined below are three possible heuristics for cost minimization with water quality constraints.

4.4.1 "Dynamic" heuristic

The first heuristic exploits an idea similar to the dynamic programming approach used for the same case study (cf [24, 28, 29]). Though dynamic programming could be very useful for this problem (since water quality at a point is obviously not affected by any emissions downstream), the "curse of dimensionality" must be considered (cf [29]). For instance, with 10 emission points and 4 decision alternatives at each point, there are already about 1 million decision combinations to consider. This problem becomes even worse when multiple water quality constituents are considered. Therefore, it is necessary to make some "pre-decisions" while proceeding downstream from the first (furthest upstream) emission point.

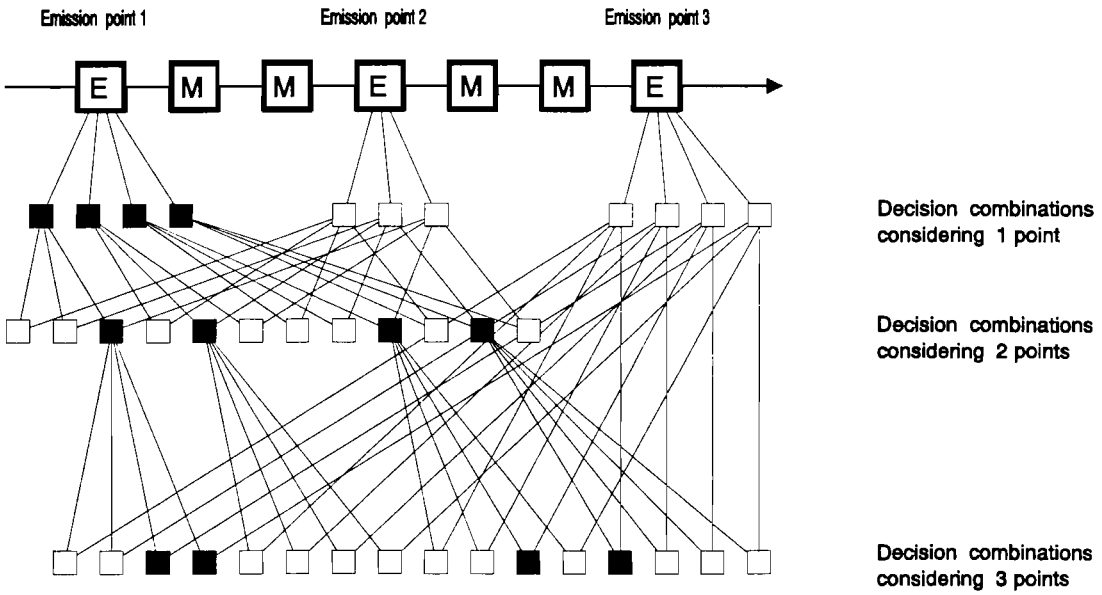

Figure 2: Dynamic heuristics.

The idea of making such pre-decisions (which follows the implementation described in [29]) is demonstrated in Figure 2. in which emission points are denoted by (E) and monitoring points by (M). At emission point 1 there are 4 acceptable decision alternatives

with respect to constraints at monitoring points just downstream. Therefore, we must consider all 4 for further calculations, as indicated by the filled boxes. If we combine these 4 with 3 possible alternatives at emission point 2, we have already 12 combinations to consider. At this point "pre-decisions" are incorporated in order to select only 4 of the 12 as "reasonable" solutions. One way to make this selection is to divide the admissible range of concentrations downstream of emission point 2 into 4 intervals (ranges) and to choose only the minimum cost solution within each interval. With these 4 possible solutions (filled boxes at second level) one can combine the 4 alternatives at emission point 3 and then repeat the interval selection process. After repeating this procedure for all emission points, the "best" solution is selected, and the technology chosen at each site is found by tracing back up the river.

This procedure could easily be supported by an object-oriented tool like ORVAN. It is only necessary to create objects for each considered decision combination and then store concentrations and a cost summation at each one. Even for larger problems (20 emission points and 6 to 8 alternatives), the problem does not seem intractable because there would be no more than 1000 objects.

Obviously, more sophisticated methods are possible in order to choose the solutions at each level. Particularly, the "pre-decisions" should not be made automatically, but instead the user should be allowed to interact with the selection process at each level.

4.4.2 "Giving incentives" heuristic

In the second heuristic we decompose the problem into N subproblems, assuming that decisions can be made at each emission point. We refer to decision "units"[5] which communicate in order to make rational decisions on a regional scale.

The procedure starts at emission point N (decision unit N) which is the furthest downstream. Decision unit N assumes first the worst situation: at the monitoring points upstream the concentrations of constituents are just at the ambient standards. Under this assumption, the unit N chooses the cheapest technology which fulfills all the standards downstream. Furthermore, this unit determines which conditions (concentrations upstream) must be provided so that it is possible to choose a cheaper technology. The cost difference between this cheaper technology and the technology which must otherwise be chosen is the amount unit N may offer to decision unit N-1 as an incentive.

The problem for decision unit N-1 is then slightly different from unit N's problem, since unit N-1 has to consider both the obligatory ambient standards and the standards desired by decision unit downstream. Again, under the "worst case" assumption, the cost effective solution is chosen, taking into account the incentives which have been offered. New "desired" standards and the incentives for the unit N-2 may then be calculated. This procedure is continued until decision unit 1 is faced with a problem consisting of the obligatory standards and a set of "desired" standards.

To implement this heuristic in a real case, additional coordination would be necessary as soon as a tributary has to be considered. Fortunately, though, the equations for calculating concentrations need not be linear. It is only required that one can invert the relevant functions in order to calculate the "desired" standards.

[5]Decision "units" might be individuals or groups. Of course, a similar approach can be applied if decisions for all units, or groups of units, are made by one person or by one group.

4.4.3 "Decoupling" heuristic

In the third approach we consider a larger problem (20 or more controlled emission points) and seek "robust decisions" at some points which allow the problem to be decomposed into smaller problems.

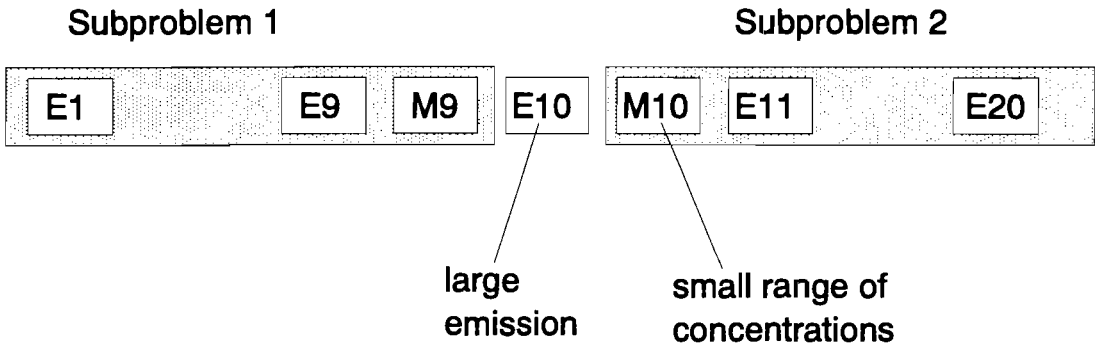

Figure 3: Decoupling heuristics.

For instance, one might be faced with the situation shown in Figure 3. At the emission point E10 there might be such a large emission that only the most expensive alternative is admissible at this point. The concentrations at the monitoring point M10 just downstream might be almost completely dependent on this large emission. Since the technology chosen at E10 is fixed (or nearly fixed), there is a small range of possible concentrations at M10. The subproblem, concerning all decisions from E11 to E20, might be not very sensitive to variations within this small range. One could then check to see if decisions at some other points are fixed, which would allow the problem to be decomposed further.

Some possible advantages of this procedure are as follows:

— Subproblems are easier to examine in detail, which is more useful in the decision-making process.

— Some subproblems might even be solved by exact methods like complete enumeration.

— One can apply this kind of heuristic to other problem formulations (such as maximization of ambient quality with cost constraints).

The subproblems illustrated in Figure 3 need not be completely independent. For instance, the interval of possible concentration levels at monitoring point M10 might be small yet significant. Then, different solutions could be obtained for subproblem 2 at both ends of this interval, so that the decisions in subproblem 1 are not independent. In this case, additional coordination methods might be considered.

5 Optimization

5.1 General remarks

Traditional single-criterion optimization for water quality management usually entails one of three basic formulations:

1. cost is minimized subject to water quality constraints,
2. water quality is maximized subject to a cost constraint, or
3. cost is minimized with the costs of water quality violations included in the form of penalty functions (cf e.g.[14])

Each of these formulations has severe limitations, especially for Eastern and Central Europe. With respect to the first two formulations, the constraints may be considered "soft" in that the decision-makers must also decide how much funding to allocate for wastewater treatment and at what levels to set water quality standards. Of course, in Central and Eastern Europe, finances are extremely scarce, and there is no good reason to impose "western-style" water quality standards (cf [26]). In analyzing the problem, decision-makers might wish to run many simulations, varying the cost/water quality constraints each time. Regarding the third formulation, placing monetary values on environmental quality is controversial and requires much case-specific research (cf [3, 25]).

Thus, multiple criteria decision analysis (MCDA) can be a valuable way to evaluate the trade-offs among costs and water quality. Rather than imposing constraints (such as water quality or cost constraints) which may not be identifiable in the real world, the decision-maker can choose among different constraints and objectives, and thus view the problem in a more flexible and realistic manner.

In this section, the methodology of both single criterion and multiple criteria analyses are presented, and the benefits and limitations of each are discussed.

5.2 Single criterion optimization

For single criterion optimization one of the performance indices must be chosen as a goal function. In order to avoid trivial and unacceptable solutions, it will also be necessary to introduce a constraint for another criterion. For the sake of illustrating this approach, we use two of the well known approaches, namely:

1. Maximization of the environmental quality under given financial resources. For our model this is equivalent to the minimization of the regional water quality index g_{all} (where g_{all} is equal to the maximum violation of the given standards) under a given constraint on total annual cost, which takes the following form:

$$\min g_{all}$$

$$\text{subject to:} \qquad Tot\_TAC \leq \overline{TAC} \tag{16}$$

   where g_{all} and Tot_TAC are defined by equations (5) and (15), respectively.
2. Minimization of costs for achieving a given water quality standard. Such a standard might be represented by g_{all} so that the problem has the following form:

$$\min Tot\_TAC$$

$$\text{subject to:} \qquad \min g_{all} \leq \overline{g_{all}} \tag{17}$$

   Note that $g_{all} \leq 0$ implies that water quality standards are held at every node. Therefore, selecting $\overline{g_{all}} = 0$ is equivalent to setting hard constraints for the water quality standards. Solution of the problem (17) for such hard constraints could result in costs which are not be acceptable. Therefore, one may want to consider *"soft"* constraints for the water quality standards, which can be achieved by trying different values of $\overline{g_{all}}$.

Of course, the practical use of any single criterion problem formulation would seldom result in one acceptable solution. On the contrary, it would require the generation and solution of many problems for different values of constraints (for $\overline{TAC}$ and $\overline{g_{all}}$, respectively, in our example). Many of these problems will not have feasible solutions and, without a good *a priori* understanding of the solution characteristics, it is generally not easy to find a set of acceptable solutions. Therefore, the application of multicriteria optimization techniques might be helpful in our case study.

5.3 Multiple criteria optimization

5.3.1 Selected methodology

An adequate discussion of different approaches to Multiple Criteria Decision Analysis (MCDA) is beyond the scope of this paper, so we will present only a short summary of the approach that we have implemented, namely the reference point approach. A reader interested in the methodological foundations and details related to different MCDA approaches is advised to consult one of many publications which provide a bibliography of this subject. Another important issue for the successful implementation of a DSS is the paradigm of *rational decision-making.* A good overview of the different concepts and issues related to real-life applications of DSS and rational decisions is given by Keeney in [6, 7], by Rapoport in [23], by Wierzbicki in [35], by Yu in [38], and by Zeleny in [39]. Several different methodological approaches to multi-objective decision support and associated techniques are presented in [9, 11, 15, 20, 37], and the software which supports them are described in [9, 37]. Two recent critical surveys of both theoretical and practical aspects of multi-objective decision making can be found in [10, 30].

From the user's point of view, the critical step of MCDA is generating a part of the Pareto-optimal solution set[6]. Generating the entire Pareto-set is practically impossible and – even if done – would result in a vast amount of useless information. Therefore, most MCDA methods generate a very limited number of Pareto-solutions[7] and then provide a tool for the analysis of these solutions and for generating another set of Pareto-optimal solutions based on these results. One danger of these methods is that they could be used as "black boxes." We believe that for a water quality DSS, the best method would be an interactive one which enhances learning[8] during the decision process. For this reason, we have implemented the *reference point* (RFP) method.

The RFP method is based on the concept of *satisficing behavior* (also called *bounded rationality*), in which the decision maker attempts first to improve the criterion which shows the worst performance (cf e.g.[18]). This method has two noteworthy advantages over other MCDA methods. First, the RFP method does not rely on explicit weighting of criteria, which can often be difficult for the decision-maker and lead to confusing results (cf e.g.[34]). Second, in contrast to multi-attribute utility theory, the RFP method does not require the identification of a *utility function*, which requires that all decision outcomes be aggregated to a single value. To estimate such a function, if even possible, the decision maker must answer many questions and make many pairwise comparisons. Many times, this does not remarkably increase the usefulness of optimization in DSS.

In practice, the RFP method may be summarized in the following stages:

1. The decision maker (DM) specifies a number of criteria (objectives). In typical applications there are 2–7 criteria. For an LP problem a criterion is often a linear combination of variables, but criteria may have other forms for specific applications (cf e.g. [17]).
2. The DM specifies an aspiration level $\bar{q} = \{\bar{q}_1, \ldots, \bar{q}_{NC}\}$, where $\bar{q}_i$ are the desired values for each criterion, and NC is a number of criteria. In some applications the DM may

[6]Efficient, or Pareto-optimal, solutions are those for which an improvement in the value of one criterion cannot be attained without worsening the value of at least one other criterion.

[7]We do not discuss here approaches based on the idea of converting a multi-criteria problem into a single-criterion one by summing up weighted criteria. The primary reasons for avoiding this approach are summarized in [34].

[8]That is, understanding the correspondence between the aspirations (represented by desired values of criteria) of a user and the attainability of such aspirations. Since aspirations are usually not attainable, a user has to learn (using the mechanisms of the RFP method) how to adjust aspirations in order to find a feasible solution which best meets his/her expectations.

additionally specify a reservation level, which is composed of the worst values of criteria that a DM would like to accept.

3. The underlying formulation of the problem is the minimization[9] of an (piece-wise linear) *achievement function* which can be interpreted as an ad-hoc non-stationary approximation of the DM's value function dependent on the currently selected aspiration levels. Then, the problem is transformed by the DSS into an auxiliary parametric single-objective problem, the solution of which gives a Pareto-optimal point[10]. If a specified aspiration level $\bar{q}$ is not attainable, then the Pareto-optimal point is the nearest (in the sense of a Chebyshev weighted norm) to the aspiration level. If the aspiration level is attainable, then the Pareto-optimal point is uniformly better than $\bar{q}$. Therefore, this approach may be considered an extension of goal programming. Properties of the Pareto-optimal point depend on the localization of the reference point (aspiration level) and on optional weights[11] associated with the criteria. Some applications offer the option of computing weights based on the utopia and nadir points[12], which usually provide for proper scaling in the criteria space. It is reasonable to expect the values of each criterion to lie between the utopia and nadir points.
4. The DM explores various Pareto-optimal points by changing either the aspiration level $\bar{q}$, the weights attached to criteria, and/or other parameters related to the definition of the multicriteria problem.
5. The procedure described in points 2, 3 and 4 is repeated until a satisfactory solution is found. Additionally, the user can temporarily remove a criterion (or a number of criteria) from the analysis. This results in the computation of a Pareto optimal point with respect to the remaining "active" criteria, but values of criteria that are not active are still available for review.

Thus, multiple criteria optimization with the reference point method can be thought of as inverse simulation: rather than repeatedly adjusting the decision variables to determine acceptable states, the user chooses desired states and determines the resulting values of the decision variables. This provides a useful complement to scenario analysis.

A more formal presentation of the technique outlined above can be found in [11], and for methodological foundations of the RFP method, one should consult [13, 33, 36]. This is just an example of a possible approach which has proven to be useful in many application areas (cf [15] for a summary). The software packages described in [16] implement this methodology for different types of mathematical models.

5.3.2 Formulation of multicriteria problems

In order to illustrate an application of the RFP method to the considered problem, we examine two formulations of the multicriteria problem. The first corresponds to the single-criterion problems formulated in Section 5.2, and the second to a more realistic five-criteria problem.

---

[9] It can be also formulated as maximization problem, depending on the interpretation of the achievement function.

[10] For the sake of brevity we will refer to properly Pareto-optimal solutions as Pareto solutions (unless otherwise mentioned). A pareto-optimal point is composed of values of all criteria for a corresponding Pareto-optimal solution.

[11] Here, weights should not be confused with weights used to convert a multi-criteria optimization problem into a single-objective problem (cf e.g. [17, 34] for the discussion of related problems).

[12] Utopia and nadir point are composed of the best and worst values for each criterion out of the set of all Pareto-solutions. One should note that finding a utopia point can be done by computing "selfish solutions" (creating a pay-off table) but finding a nadir point may be difficult (cf e.g. [5]).

For the two-criteria problem, the following criteria are used:
- minimize TAC
- minimize g_{all}

whereas for the five-criteria problem the criteria are as follows:
- minimize Tot_IC
- minimize Tot_TAC
- minimize g_{all}
- minimize g_0
- minimize g_3

Admittedly, the two-criteria problem is of limited usefulness for making a final decision. The use of g_{all} as a "universal" water quality index is questionable. For instance, because of the data and model assumptions we have used, phosphorous practically controls the value of g_{all} so that there is no mechanism for increasing DO concentrations. However, we have chosen this formulation only to demonstrate how MCDA can be applied for examining the overall trade-offs between cost and water quality. The two criteria formulation is also useful for solving the problem of minimization of costs with soft constraints on water quality.

Once experience is gained with the model and with the RFP method applied for the two criteria case study, though, one should consider also the five-criteria problem. In this formulation, two key constituents (DO and NH4) are considered as separate criteria, g_{all} is included to allow consideration of the other constituents, and the investment cost tot_Inv is a criterion in addition to the total annual cost tot_TAC.

Both multicriteria problem formulations are composed of the above objectives, equations (1) through (7), and equations (9) through (15). Note that these equations involve only the constraints related to the transport of constituents in a river and to the definitions of variables. Therefore, none of the decision variables is constrained by a quantity which is in fact an edogenous decision variable (such as a maximum available budget or an acceptable constituent concentration). However, due to the nature of the RFP method, one can examine various Pareto-optimal solutions which represent compromises between costs and water quality. The basic advantage of this approach is that it provides a natural way to generate a new scenario (by setting a new RFP) without facing the risk of infeasibility.

5.3.3 Solving the multicriteria problem

A Pareto-optimal solution can be found by the minimization of the achievement scalarizing function in the form

$$\max_{i=1,\ldots,NC}(w_i(q_i - \bar{q}_i)) + \varepsilon_m \sum_{i=1}^{NC} w_i q_i \rightarrow \min \tag{18}$$

where NC is number of considered criteria[13]. This form of the achievement function is a slight modification of the form suggested by Wierzbicki [33]. Note that for $\varepsilon_m = 0$ only weakly Pareto-optimal points can be guaranteed as minimal points of this function. Therefore, the use of a very small ε_m results (except in situations in which reference point has some specific properties) in properly Pareto-optimal solution with trade-off coefficients bounded approximately by $\varepsilon_m NC$ and $1/\varepsilon_m NC$. If ε_m is very small, these properly efficient solutions might not differ significantly from weakly efficient (Pareto

[13]Note that the constant term $(-w_i\bar{q}_i)$ was dropped from summation in this equation.

optimal) solutions. On the other hand, too large a value of ε_m could drastically change the properties associated with the first part of the scalarizing function. For details of the conversion of the multicriteria problem with the help of the achievement scalarizing function, consult e.g. [17].

The first step in the analysis is to determine the utopia point and an approximation of the nadir point. These are determined by finding the "selfish solution" for each criterion. For instance, the utopia point value for Tot_TAC can be found by solving the single criterion problem of minimizing Tot_TAC without constraints on the water quality.

Preliminary computations of the utopia and nadir points show that the problem, as formulated, is already well scaled. Therefore, we implement a very simple scheme for determining the weights used in the scalarizing function (18), using only the reference point values to automatically scale these weights as follows:

$$w_i = \text{amax1}(rfp_i) * NC / \sum_{k=0}^{NC} \text{amax1}(rfp_k) \tag{19}$$

where rfp_i and w_i are values of i-th reference point component, and amax1() is defined as:

$$\text{amax1}(x) = \max(1., \text{fabs}(x)) \tag{20}$$

where fabs(x) is a standard library function which returns the absolute value of its argument.

6 Data used for the prototype

All of the primary data used in our simulations have been obtained from the Water Resources Project at IIASA. The data are preliminary and are used exclusively for testing the model formulation and the software being developed.

6.1 Uncertainty

Uncertainty in data and system identification are major concerns in any water quality modeling effort, and we can identify many sources of error/uncertainty in this study. Though using the best available data, we are aware that great deal of effort is being made by colleagues in the Water project to collect and verify more complete, coherent data.

The uncertainty stems from at least several factors, including the following: natural variability, measurement and analysis error, lack of knowledge of water quality processes, and the failure to identify non-point or small point source emissions in the river, as well as extractions from the river (cf [1, 8]). The first two sources of error (measurement/analysis error and natural variability) are apparent in the data received from a Slovakia agency (cf [28, 29] for details). Analyses were performed in two different labs, and results were often quite different. Even for analyses performed in the same lab, different values were obtained for samples taken at different times of the day. For example, at one monitoring point measured DO concentrations (analyzed in the same lab) ranged from 3.1 to 9.2 mg/l over the course of 12 hours. Also, for samples taken at the same time, the two labs reported DO concentrations of 9.2 and 7.9 mg/l, respectively. The second two sources of error (system and source identification) are discussed in a paper by Koivusalo [8]. In this paper, the authors defined "error loads" to account for unknown discharges or tributaries, and in some cases these loads were greater than any single known discharges.

The effects of this uncertainty can be seen in model calibration. Koivusalo [8] calibrated a simple Streeter-Phelps model and determined a value for the deoxygenation coefficient (CBOD decay rate coefficient) of 0.0488/day for the entire river. Masliev [19] calibrated an extended Streeter-Phelps model and determined a value of 0.72/day, more than an order of magnitude greater. Also, both parameter estimates are far from the literature values of 0.2-0.3/day (cf [31]).

Obviously, the uncertainties are being considered during the on-going process of data verification. We mention them here as a basic argument for our statement that the results discussed in this Working Paper are solely for illustrating the developed software and suggested methodology, and by no means should they be interpreted as recommendations for the Nitra River Case Study.

6.2 List of Data

The following list is provided to summarize our data sources and corresponding model variables. From a survey performed by Slovakian agencies on 25-26 August 1992:

KM_j – distance of each node from the mouth of the river on which it is located, km

TR_j – travel time to each node from the first surveyed point upstream, $hours$

Q_j – river flow at each node, m^3/day

q_j – waste flow at each emission point, m^3/day

T_j – temperature of the river at each node, C

$CBOD5_j$ – 5-day CBOD concentration at each node, mg/l

$NH4_j$ – ammonia concentration at each node, mg/l

P_j – dissolved phosphorous concentration at each node, mg/l

Note: CBOD5, NH4, and P are measured in the river at monitoring and confluence points, and in the waste stream at emission points.

From a QUAL2E modeling study (cf [2]):

v_j – velocity at each node j, m/s

h_j – depth of water at each node, m

From the calibration of an extended Streeter-Phelps model (cf [19]):

K_0 – average reaeration rate coefficient for the entire Nitra River, $/day$

K_1 – average CBOD removal rate for the entire river, $/day$

K_2 – average NBOD removal rate for the entire river, $/day$

And, from expert estimates of treatment costs and efficiencies (cf [28]):

em_{jl} – effluent concentration of each constituent resulting from each technology, mg/l

IC_{jk} – investment cost of each treatment technology at each emission point, (in 10^6 US$)

OMC_{jk} – operating cost of each treatment technology at each emission point, (in 10^6 US$)

6.3 General comments

Measured flows (emissions and in the river), travel times, temperatures, and pollutant concentrations (CBOD-5, NO2, NO3, NH4, P, PO4) were obtained from a survey done by the Slovakian agency on 25-26 August 1992 (cf [28, 29] for details). CBOD-5 and water balances were done on the data for Nitra river kilometers 6.5 to 132.5, but not its important tributaries (the Handlovka, Zitava, and Stara Nitra Rivers). Even so, there is at least one stretch of the Nitra where an unexplained loss of water occurs, and travel time is not available.

For this stretch of the Nitra (river kilometer 65.25 to 91.1) we relied on velocity estimates from a hydraulic modeling exercise using QUAL2E. QUAL2E estimates of depth and velocity were also used to calculate the reaeration rate K_0 (cf [2]). Along the tributaries, where morphometry data is not available so no estimates of depth and velocity were made, we simply assumed $K_0 = 4.0$ /day.

The reaeration coefficient (K_0) and the CBOD and NBOD decay rates (K_1 and K_2, respectively) used in our model were obtained from the calibration of an extended Streeter-Phelps model. The methodology of calibration was the Hornberger-Spear-Young (HSY) approach, based upon the Monte-Carlo technique (cf [1]). The result was a single value for the coefficients for upper part of the Nitra River, kilometer 91.1 to 132.5 (cf [19]). Again, at the time of our analysis, calibration for the rest of the Nitra and its tributaries was not possible, so we simply used these same rate estimates along these reaches.

Site-specific data for different treatment technologies was also compiled by the Water Resources Project. The data included waste flow capacity, investment and operating costs, and the constituent concentrations resulting from each technology (cf [28]). We obtained data for ten municipal discharge sites, but only considered nine sites in our model due to a lack of flow data.

6.4 Implemented data conversion

Several data items are provided in a form that requires simple conversions before using them in the model. Such conversions are listed below for the sake of documentation.

- The water quality data includes values of 5-day BOD (CBOD5) and ammonia (NH4), from which we can calculate ultimate CBOD emissions em_{j1}:

$$em_{j1} = CBOD5_j/(1 - exp(-5 * K_{1j})) \qquad j \in E \tag{21}$$

where K_{1j} is the CBOD removal rate [/day] at emission point j (cf eq. (26), and NBOD emissions em_{j2}:

$$em_{j2} = \alpha * em_{j3} \qquad j \in E \tag{22}$$

where the coefficient α is estimated from the stoichiometry of the nitrogen-oxygen cycle and is equal to a number between 0.0 and 4.57. Using $\alpha = 4.57$ would represent a "worst-case" assumption (i.e., all of the ammonia immediately consumes oxygen). We assumed $\alpha = 2.0$, but this value should really be determined through model calibration.

- Travel times TR_j, in hours, are provided for each node from the farthest upstream measurement point. For each river and for all but the last nodes, a difference of those quantities is computed, converted into days and stored as TR_j. For the last node (on each river) the travel time is computed from the velocity at this node and the distance of this node from the confluence node.
- We assume that the sediment oxygen demand [g/m^2-day] is constant along the river

$$SOD_j = 1.5 \tag{23}$$

- For several data items, we have provided for the possibility of using default values whenever a value is not available. In the presented implementation the following defaults are used:

  Q_j – a sum of upstream values at a confluence node, otherwise a weighted (by distances) average of the nearest upstream and downstream values (if one of them is not available, then the other one is assumed instead of the average),

  T_j - 22°C

  v_j - 0.5 m/s

  h_j - 0.3 m

  Thanks to the organization of the software, the list of data items for which defaults is used can be easily modified.

6.5 Calculated parameters

Several parameters used in the model formulation are calculated in advance from the provided data. These calculations are as follows:

- The saturated DO concentration, $DOsat_j$, is empirically defined as

$$DOsat_j = exp(-139.344 + 1.5757e5/T_j - 6.6423e7/T_j^2 + 1.2438e10/T_j^3 - 8.622e11/T_j^4) \quad (24)$$

  where T_j is the temperature (converted to Kelvin scale) at the j-th node.

- Assuming first-order decay, the transfer coefficients in equations (6, 7, 8) are defined as follows:

$$TC_{jl} = exp(-K_{lj}TR_j) \qquad l \in [0,4] \quad (25)$$

  where K_{lj} is a temperature-dependent decay rate [/day], and TR_j is the travel time (in days) in river segment that starts at node j.

- We treat ammonia (NH4) and NBOD synonymously. Their concentrations differ by a factor of α (cf eq (22)) and their decay rates are the same ($K_{2j} = K_{3j}$). Also, we will treat phosphorous as a conservative nutrient ($K_{4j} = 0$).

- Each of the decay rates in the transfer coefficients are temperature-dependent. The rate coefficients (for 20°C) and temperature corrections are given as follows:

$$K_{ij} = K_{ij}^{20}\beta_i^{(T_j-20)} \qquad i = 1,2 \quad (26)$$

  where $\beta_1 = 1.04$ and $\beta_2 = 1.08$. The reaeration rate (K_0) is also dependent on the hydraulic of the river so that (cf [19]):

$$K_{0j} = K_{0j}^{20}(1.024)^{(T_j-20)}\sqrt{v_j/h_j} \quad (27)$$

  where v_j is the average velocity [m/s], h_j is the depth [m], and T_j is the temperature [C]. The values of coefficients K_{0j}^{20}, K_{1j}^{20}, and K_{2j}^{20} provided in the data file (cf Section 6.2) are for 20°C.

- The effects of the oxygen-demanding constituents on the DO concentration are represented by different transfer coefficients than calculated in (25). We denote these by TCp_{jl} and define them as follows:

$$TCp_{j1} = [K_{1j}/(K_{0j} - K_{1j})][exp(-K_{1j}TRj) - exp(-K_{0j}TR_j)] \quad (28)$$

$$TCp_{j2} = [K_{2j}/(K_{0j} - K_{2j})][exp(-K_{2j}TR_j) - exp(-K_{0j}TR_j)] \quad (29)$$

$$TCp_{j5} = (K_{0j}/h_j)[1 - exp(-K_{0j}TR_j)] \quad (30)$$

  where h_j is the depth [m] of the river.

6.6 Availability of data

The data used in the reported research is not listed here because the data set is quite large and it will soon be modified substantially. A reader interested in the data used in this research should contact M. Makowski by e-mail (`marek@iiasa.ac.at`).

7 Discussion of results

In this section we present some results from scenario analysis with ORVAN and from optimization with the problem generator and MOMIP. In no sense are these results "complete", nor should they be used for policy recommendations. They merely demonstrate the capabilities of the decision support tools investigated.

7.1 Scenario Analysis with ORVAN

The plot shown in Figure 4 shows DO concentrations for four of the scenarios generated by ORVAN. The labels (1, 2, 3, 4) for the curves represent the following four scenarios:

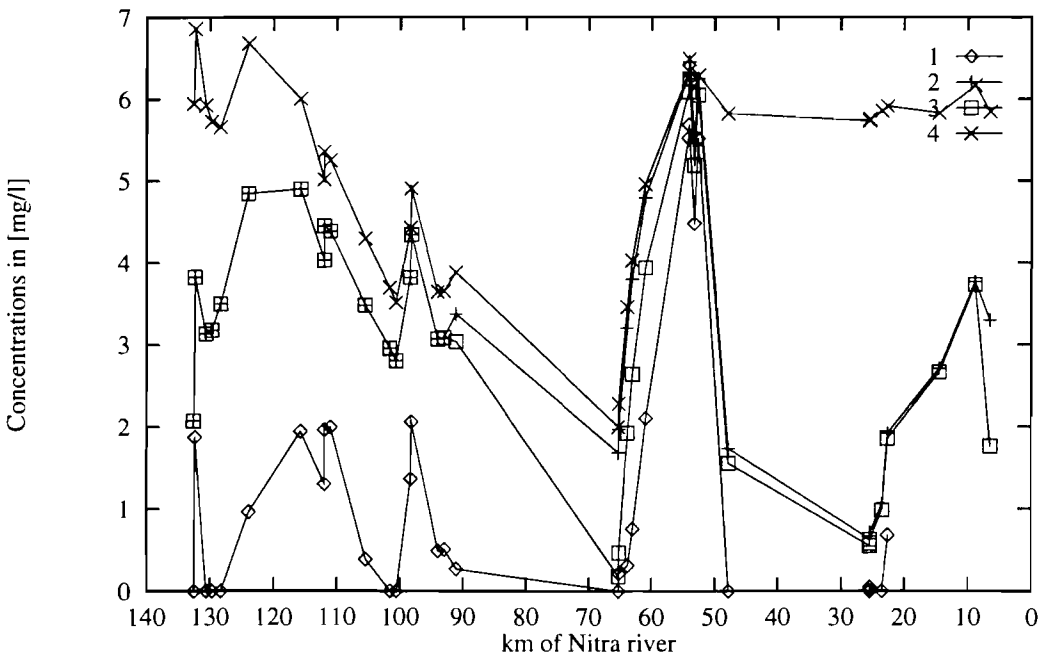

Figure 4: DO concentrations for 4 scenarios.

1. Do Nothing: No treatment occurs along the river, representing the worst-case condition. Total annual cost and investment cost are zero.
2. Baseline: The existing treatment technologies are used at each emission point. Total annual cost is 5.39 [10^6 US$], and investment cost is zero.
3. Do Something: Some new treatment technologies are implemented, requiring a total annual cost of 12.5 [10^6 US$], and an investment cost of 37.4 [10^6 US$].
4. Do All: The best available treatment technologies are implemented at each point. This is very expensive, requiring a total annual cost of 21.74 [10^6 US$] and an investment cost of 86.3 [10^6 US$]

Each scenario was generated using the same reaeration and deoxygenation coefficients ($K_0 = 2.7$ /day and $Kr = 0.72$ /day, respectively).

Several interesting things are shown in Figure 4. First, we note that the baseline scenario results in much better water quality than the "do nothing" case. Thus, although financial resources are scarce in Slovakia, the existing treatment is highly beneficial. We also see an appreciable improvement in water quality when some new treatment technologies are implemented, but (for most of the river) little more is gained by implementing the best available treatment. Also of great interest are reaches along the river where water quality is apparently quite unaffected by the technology chosen. This is to be expected downstream of large, uncontrolled emissions, providing incentive for a closer look at decomposition heuristics as presented in Section 4.4.

Figure 5 again shows the results of four scenarios, but these differ only by the value of the reaeration coefficient. The same technologies are implemented for each scenario, but values for the reaeration coefficient ranged as indicated in the caption.

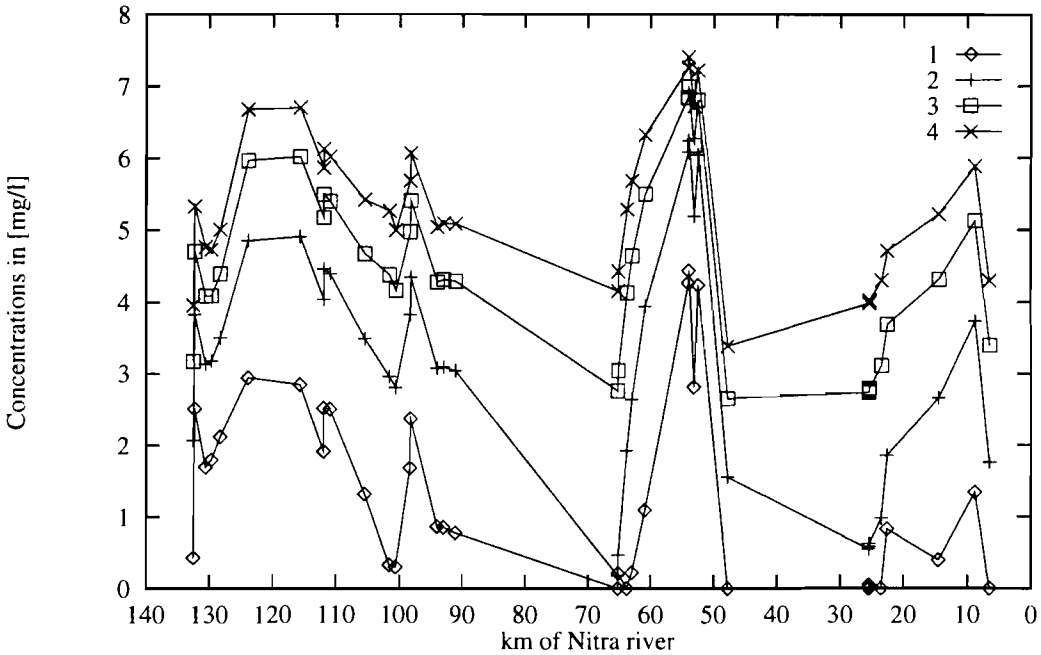

Figure 5: DO concentrations for different values of the reaeration coefficient: [1]- 1.7, [2]- 2.7, [3]- 3.7, [4]- 4.7, per day, respectively.

Though we have already discussed the uncertainties involved in the data used in our model (cf Section 6), Figure 5 allows us to quantify this uncertainty to some degree. We see that relatively small variations in the reaeration coefficient result in dramatic changes in simulated DO concentrations, especially where waste loads are high. One can recognize the overall uncertainty in the model results when sensitivity analysis is performed on other parameters. However, DO concentration is most dependent on the reaeration coefficient, so we show only the dependence on K_0 here. In any case, this uncertainty should be made evident in the decision-making process, and scenario analysis is perhaps the method most understandable to the decision-maker.

7.2 Optimization

Figures 6 and 7 show the results (for DO and NH4 concentrations, respectively) of a single criterion optimization in which g_{all} is minimized with a constraint on TAC, ranged as indicated in the captions. The value of g_{all} is calculated using the following standards:
$aqs_0 = 6.0$ mg/l
$aqs_1 = 20$ mg/l
$aqs_2 = 30$ mg/l
$aqs_3 = 6.0$ mg/l
$aqs_4 = 1.0$ mg/l

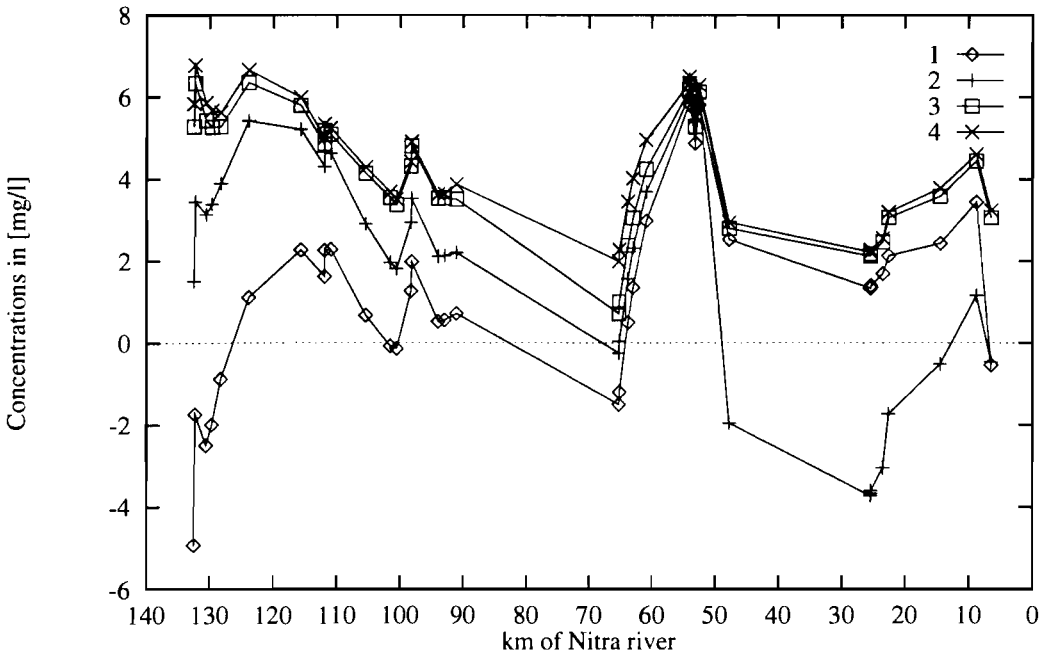

Figure 6: DO concentrations for different constraint values for TAC (in [10^6 US$]): [1] - 3.0, [2] - 5.0, [3] - 8.0, [4] - 12.0

Figure 6 illustrates some results from optimization which are comparable to simulation results presented in Figure 4. Again, we see a great improvement in water quality for the first increment of financial allocations, but diminishing returns for higher spending. The same is true for Figure 7, which shows only a minor decrease in NH4 concentrations with spending increases above 3.0 10^6 US$/year. Of course, small changes could be significant if there exists, for instance, a "threshold" which is essential for protecting human health or the habitat of aquatic species.

A notable difference in Figure 6 is that DO is below zero in some places. This occurred because constraining DO to a non-negative value would have required a substantial reformulation of our model and extensive changes to the generator. In a physical sense, these negative DO concentrations show where complete deoxygenation (anoxic conditions) occur along the river. Such conditions may well exist in the Handlovka River, which flows into the Nitra River at the point where our analysis of the Nitra starts. Of course, an alternative to imposing a non-negativity constraint on DO in the model is to examine the Handlovka River separately and to assume initial point conditions for the Nitra using the results of this examination.

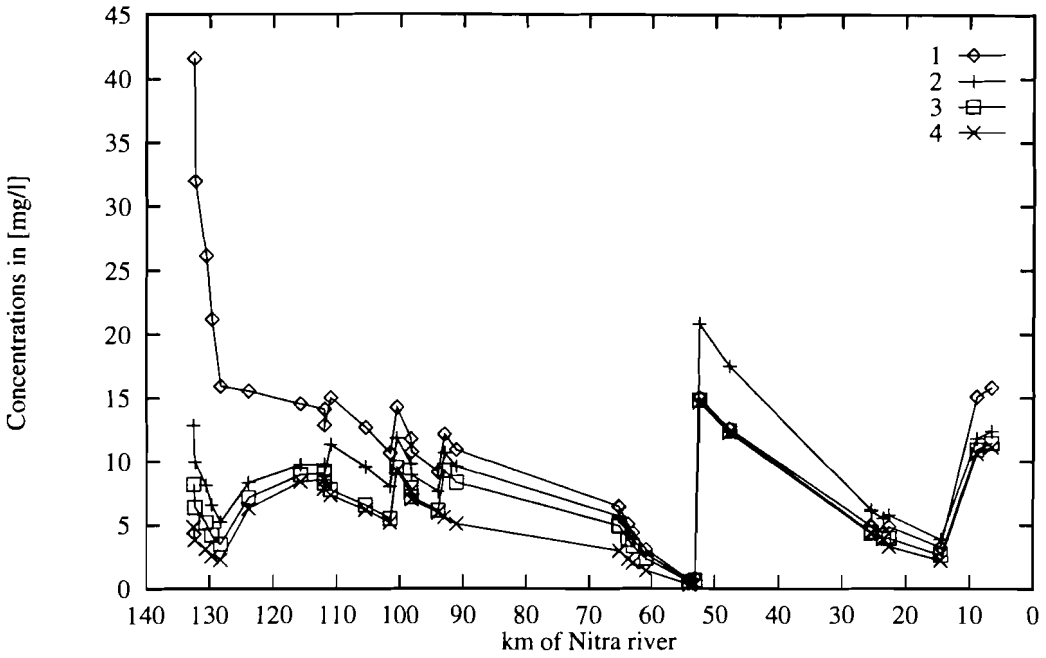

Figure 7: NH4 concentrations for different constraint values for TAC (in [10^6 US$]): [1] - 3.0, [2] - 5.0, [3] - 8.0, [4] - 12.0

We could show similar plots for the second formulation of the single criterion problem (17) in which TAC is minimized with given constraints on g_{all}. However, such an analysis would provide basically the same information.

Of greater interest is the two-criteria formulation, in which the trade-off between TAC and g_{all} can be analyzed with the reference point method. We start by calculating the utopia and nadir points, which are found to be ($TAC = 0$, $g_{all} = 0.67$) and ($TAC = 18.26$, $g_{all} = 14.4$), respectively. Note that TAC is given in [10^6 US$], and g_{all} represents the maximum (over all monitoring points and for all types of constituents) relative violation of a standard[14]

The pay-off table for 2 criteria problem		
Criterion	Criterion value	
minimized	TAC	g_all
TAC	0.0	14.4
g_all	18.26	0.67

Using these values to guide our selection of reference points, we have computed several Pareto-optimal points to illustrate the Pareto set, as shown in Figure 8 (for a real application more solutions should be used for such plot). One should note that for a MIP problem a Pareto set is composed of points. Therefore, there is actually no continuous Pareto surface as might be implied by the plot in Fig. 8. However, this plot provides not only insight to the characteristics of the Pareto set, but also illustrates an example of a surface spanning a Pareto set which is not convex. This is important because none of the Pareto-optimal solutions located on the concave portion of the plot (for

[14]Therefore, $g_{all} = 14.4$ means that at some node a violation of the standard of one constituent was equal to 14.4 times the given standard value.

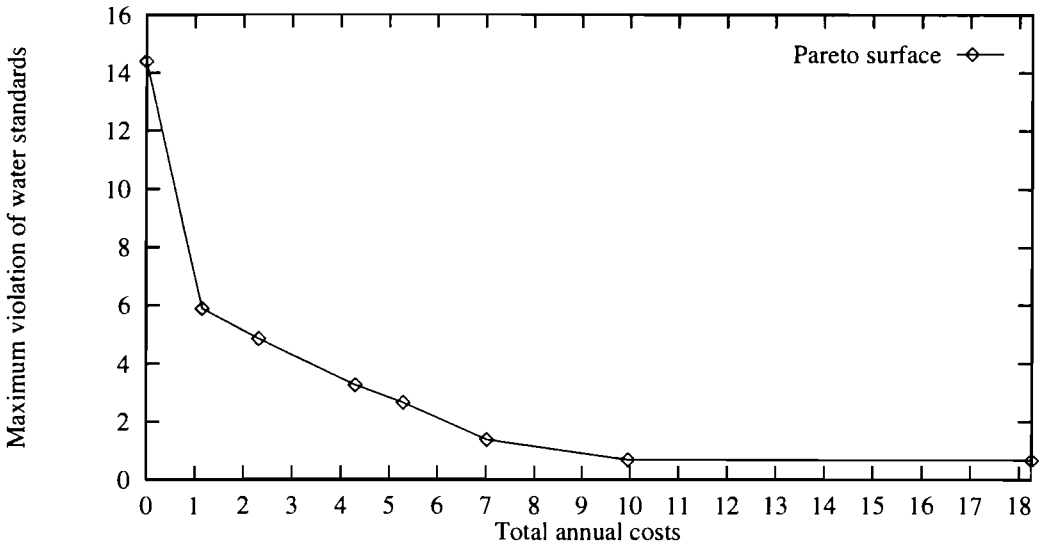

Figure 8: An illustration of the Pareto set for the two-criteria problem

TAC $\in$ [4.1, 6.9]) can be found if the multicriteria problem is aggregated into a single-criterion problem with weights. Similarly, it would be difficult to find Pareto-optimal solution for TAC $\in$ [1.14, 9.95]) which is (cf the following arguments) the most interesting range of TAC that should be examined.

One can draw some conclusions even upon a rough analysis of the Pareto set illustrated by Figure 8. To facilitate the discussion, let us consider the plot on this Figure as an approximation of a function f such that $TAC = f(g_{all})$. We expect that such a function would be decreasing, but it is interesting to note domains of the function in which the absolute value of the derivative of the function (which illustrates the "speed" of its decreasing) is large, moderate, and small. For the two extreme segments of the function, the values of the derivative are 7.43 and 0.005, respectively. One can expect rational solutions for TAC to belong to the sub-domain of moderate values of the function derivative. In a physical sense, this means that one would likely choose to increase spending only as long as significant pollution reductions continue to be seen. Another significant result is that for a given set of considered technologies (and assuming correctness of data), even the most expensive solution would result in a substantial (67% of the standard value) violation of a standard.

Finally, we present some more detailed results from the analysis of the 5-criteria problem.

An illustration of the Pareto set for the 5-criteria problem is given in Figure 9, which illustrates a projection of the Pareto set (based on 10 Pareto-optimal solutions) for 2 selected criteria (violation of DO and NH4 standards) plotted against minimization of TAC criterion. Also plotted is the maximum violation of standards by other constituents (CBOD, NBOD, and P).

One can observe that the range of TAC for which one should expect interesting trade-offs is narrower than for the two-criterion problem. Analysis of results for the IC criterion would probably make this range even smaller. Again, the non-convexity of the Pareto "surface" is apparent. This further illustrates (see also Figure 8) our reluctance to use any approach based on the idea of aggregation of criteria into one criterion with the help of weights. More detailed arguments that support our point of view can be found in [34].

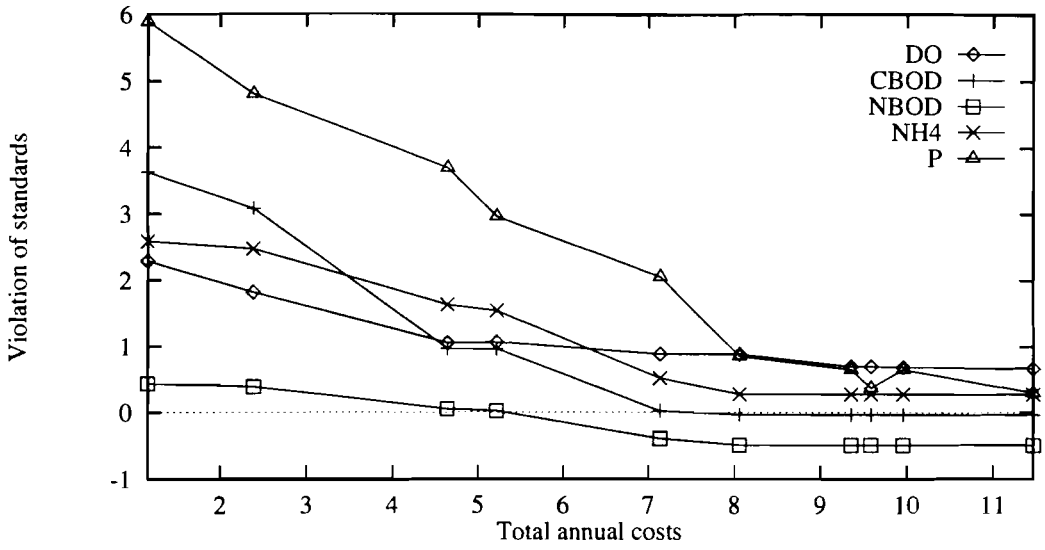

Figure 9: An illustration of the Pareto set for three criteria (TAC, DO, NH4). Violations of standards for other constituents are added for comparison. Units are 10^6 US$ and (dimensionless) relative violations of standards.

With 5 criteria, it useful first to compute a "pay-off table" which shows the criteria values obtained from selfish solutions. This is shown in the following table.

The pay-off table for 5 criteria problem					
Criterion	Criteria value				
minimized	TAC	IC	g_all	g_0	g_3
TAC	0.0	0.0	6.72	3.65	2.73
IC	4.5	0.0	5.07	1.10	1.63
g_all	11.2	25.4	0.67	0.67	0.27
g_0	10.9	25.2	0.74	0.67	0.27
g_3	7.1	9.0	1.10	1.01	0.27
Nadir	11.2	25.4	6.72	3.65	2.73
Utopia	0.0	0.0	0.67	0.67	0.27

Note, that for a two-criteria problem it is possible to calculate a true Nadir point. However, this is generally not possible for a problem with more than two criteria (cf e.g. [5] for more details). Therefore, the row labeled *Nadir* in the above table does not provide a true Nadir point, but instead it is based on the worst values of the respective criteria obtained for the Pareto-solutions computed. This is why, for example, the TAC component is equal to 11.2 in the above pay-off table, while its true value should be at least 18.26 (cf the pay-off table for the two criteria problem).

A payoff-table can then be used as a guide for the evaluation of the trade-offs among a number of criteria, which is done by a series of reference point selections. A typical starting point is to choose the utopia point as the first reference point, representing a compromise solution for all the criteria. Doing so results in the DO and NH4 concentrations (labeled 1 and 2, respectively) shown in Figure 10.

From here, it might be reasonable to compare the trade-off between total annual cost and investment cost. Selecting the appropriate reference points to favor one of these at a time provides the results shown in Figure 11. Though the DO levels are quite similar

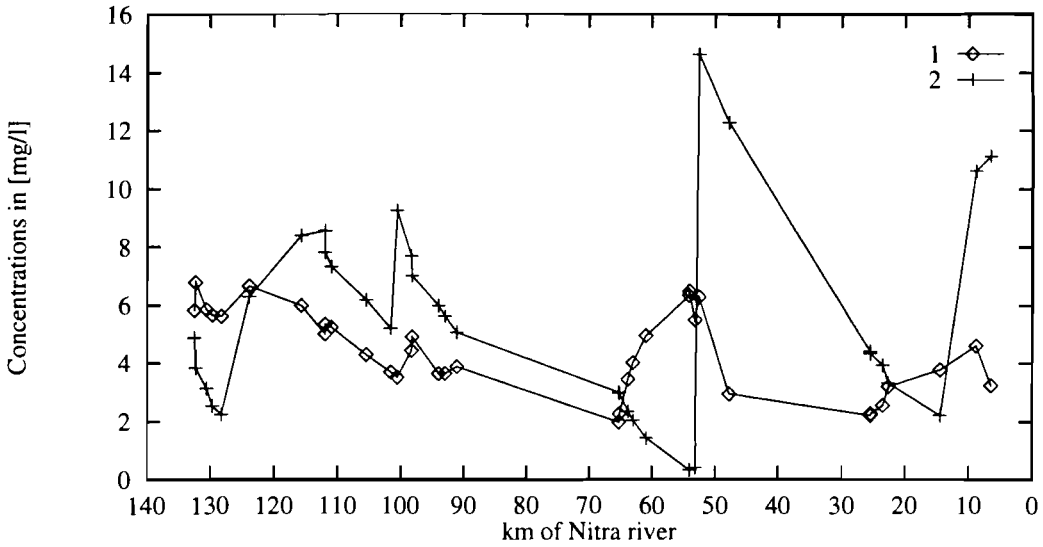

Figure 10: DO[1] and NH4[2] concentrations for first compromise solution. (TAC = 4.47, IC = 4.5)

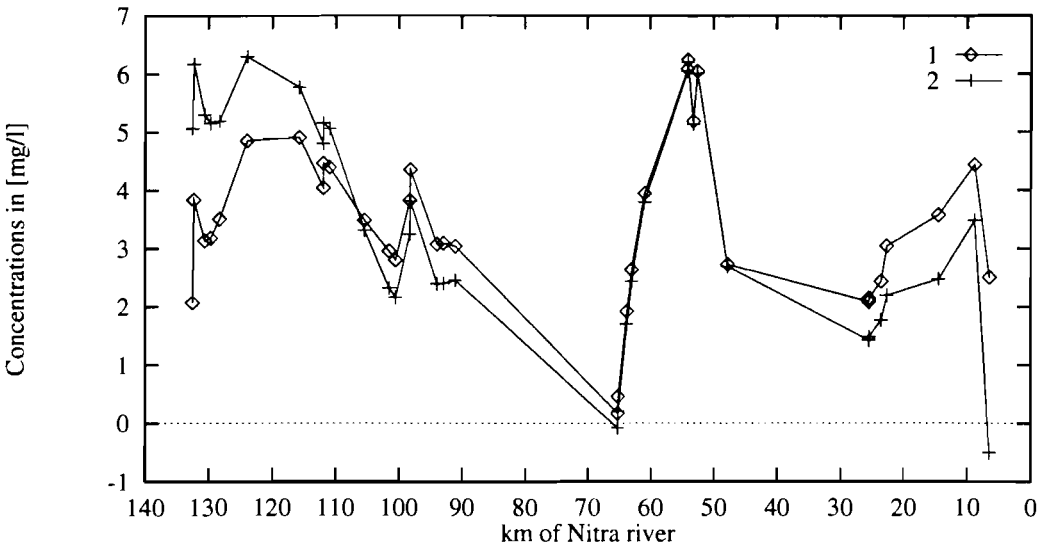

Figure 11: DO concentrations for [1] TAC = 6.03, IC = 2.5 and [2] TAC = 5.16, IC = 6.7

along the river, it is interesting to note that the results were obtained with significantly different levels of investment cost.

Similar trade-off analyses with respect to the different water quality criteria or the water quality in different stretches of the river would also be interesting, but for the sake of brevity will be omitted here. For the same reason the analyses presented here are done only for the Nitra river despite the fact that the actual computations have been made for the basin composed of four rivers.

The results presented in this Section illustrate the potential application of both scenario analysis and MCDA methodologies and tools for waste water management in the Nitra Basin and for similar case studies.

8 Conclusion

The research reported in this paper was performed during the three-month YSSP 1993. However, a substantial part of this research was completed in September after the YSSP was finished, with the authors located in different countries cooperating intensively via internet.

Critical problems for any DSS applied to regional water quality management are the quality of the data and the formulation of the mathematical model. To check our data and model formulation, we compared the results from our model with the results obtained from other models developed by our colleagues from the WAT Project. Although our model is much simpler, and no formal model calibration has been performed, the key results (with the same data and assumptions) were judged reasonable upon comparison with the more complex model[15]. This statement does not imply that our simpler model is better. On the contrary, important data used in our model has been provided by more complex models. However, our simple model-based DSS can complement more complex models by helping identify dominated solutions and by providing insight to the sensitivity of decision outcomes to various model parameters. This shows the value of the approach developed by WAT Project: developing a "family" of models of varying complexity can make it possible to rely on a relatively simple, yet adequate, management model in a DSS.

The results discussed in Section 7 illustrate the potential of two types of tools for decision support:

- First, the object-oriented development tool ORVAN, which was programmed in a short time for a new application area. Despite a lack of experience in water quality modeling, we were able to use ORVAN for efficient data processing, fast prototyping of the model formulation, and scenario analysis.
- Second, a problem specific generator for generating selected types of optimization problems. A modular MIP solver MOMIP was used to solve the corresponding optimization problem. The public domain software GNUPLOT was used for creating plots of results. In order to make this possible, a postprocessor was written for extracting the necessary data from the MOMIP output file. So far, four types of problems have been implemented (two single-criterion and two multi-criteria), but other types can easily be added.

We think that this research has resulted in some interesting results which can provide a basis for continuing this activity at IIASA. Obviously, verification of the model used in our research and comparison with the results obtained from other models developed in the WAT Project (e.g. [28, 29]) will be required. Most likely, we will need to run both the simulation and optimization with a new set of data which is being compiled by the WAT Project.

If the tools and methodology developed in this research are recognized as useful in the activities of the WAT Project, we suggest the following topics for further work:

- Improvement of the functionality and efficiency of the ORVAN application and implementation of heuristics supporting scenario analysis. In particular it might be interesting to deal with issues of coordination problems (cf Section 4.4). Increasing user-friendliness, and developing more flexible capabilities for scenario generation and analysis should obviously be also done.

- Development of a user-friendly, stand-alone tool for data modification.

[15]We cannot offer any quantified comparisons because, at the time this research was performed, the WAT Project models and data were still under development.

- Development of a user-friendly interface for the optimization problem generator.
- Implementation of a tool supporting interactive multicriteria analysis. This can be done on the basis of the tool presented in [4]. Such a tool should support basic functions for MCDA, including the selection of reference points upon analysis of previous results, the choice of an achievement function from several alternatives, and the selection of criteria to be considered.
- Assessment of the value of incorporating new criteria into the multiple criteria decision problem. Some ideas include the following:
  1. Grouping nodes on the river according to the desired water use and considering the water quality of each group to be an individual criterion.
  2. Including economic incentives in the problem formulation.
  3. Adding a "risk" criterion (such as minimization of the derivative of DO with respect to the reaeration coefficient) in order to deal with model uncertainty more directly in the decision-making process.
  4. Perhaps meeting with decision-makers to design a DSS in better accordance with their backgrounds and needs.
- All of the optimization problems generated during this research have been solved efficiently (we have not recorded times, but each problem was solved in less than one minute) by the current version of MOMIP. However, for more computationally intensive problems, special ordered sets (SOS) should be used. This technique is already implemented in MOMIP, but modification of the generator and additional tests of this option may be necessary.

The above list is based on the experience gained so far and can be easily adapted to the actual needs of the WAT Project.

Acknowledgments

The authors thank Laszlo Somlyody, Maddumage Kularathna, Ilia Masliev, and Steve Breithaupt for their help in understanding many different and difficult aspects of the modeled problem and also for providing data used in the reported research. Thanks are also extended to Włodek Ogryczak who has made the necessary improvements on the MOMIP package upon short notice of problems and to Antoni Stam for his cooperation in MCDA.

The research would not have been possible without the involvement of our colleagues in this activity. However, the authors assume full responsibility for any errors and faulty assumptions that might remain in this report.

References

[1] Beck, M.B. (1987) Water Quality Modeling: A Review of the Analysis of Uncertainty, Water Resources Research, vol. 23(8), pp. 1393-1442.

[2] Breithaupt, Steve (1993), private communication.

[3] Carmichael, J. (1993), private communication, possibly to be published as IIASA Working Paper.

[4] Granat, J. and A.P. Wierzbicki (1993) Interactive Specification of DSS User Preferences in Terms of Fuzzy Sets, WP-93-xx, International Institute for Applied Systems Analysis, Laxenburg, Austria, (to appear).

[5] Isermann H. and R.E. Steuer (1987) Computational Experience Concerning Payoff Tables and Minimum Criterion Values over the Efficient Set, European Journal of Operational Research, vol. 33.

[6] Keeney, R.L. and H. Raiffa (1976) Decisions with Multiple Objectives: Preferences and Value Tradeoffs, John Wiley and Sons, New York.

[7] Keeney, R.L. (1992) Value Focused Thinking, A Path to Creative Decisionmaking, Harvard University Press.

[8] Koivusalo, H., Varis, O., and Somlyody, L. (1992) Water Quality of Nitra River, Slovakia–Analysis of Organic Material Pollution, WP-92-084, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[9] Korhonen P. and A. Lewandowski, J. Wallenius, Eds. (1991) Multiple Criteria Decision Support, Lecture Notes in Economics and Mathematical Systems, vol. 356, Springer Verlag.

[10] Korhonen P. and H. Moskowitz, J. Wallenius (1992) Multiple criteria decision support – A review, European Journal of Operational Research 63, 361-375.

[11] Lewandowski, A., and Wierzbicki, A. (1988) Aspiration Based Decision Analysis and Support, Part I: Theoretical and Methodological Backgrounds, WP-88-03, IIASA, Laxenburg, Austria.

[12] Lewandowski, A. and A.P. Wierzbick, Eds. (1989) Aspiration Based Decision Support Systems. Theory, Software and Applications. Lecture Notes in Economics and Mathematical Systems, vol. 331, Springer Verlag.

[13] Lewandowski, A. and A.P. Wierzbicki, Decision Support Systems Using Reference Point Optimization. In [12].

[14] Loucks, D.P., Stedinger, J.R., and Haith, D.A. (1981) Water Resource Systems Planning and Analysis, Prentice-Hall, Englewood Cliffs, New Jersey.

[15] Makowski M. (1991) Selected Issues of Design and Implementation of Decision Support Systems,WP-91-16, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[16] Makowski M. and T. Rogowski (1991) Short Software Descriptions, WP-91-18, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[17] Makowski M. and J.S. Sosnowski, HYBRID: Multicriteria Linear Programming System for Computers under DOS and Unix, in [32].

[18] March, J., and Simon, H.A. (1958) Organizations, John Wiley and Sons, New York.

[19] Masliev, I., and Somlyody, L. (1994). Probabilistic Methods for Uncertainty Analysis and Parameter Estimation for Dissolved Oxygen Models. Paper submitted to the 1994 Budapest Biennial conference of the International Association of Water Quality (IAWQ).

[20] Michalowski W. and T. Szapiro (1992) A Bi-reference Procedure for Interactive Multiple Criteria Programming, Operations Research, Vol. 40, No. 2.

[21] Ogryczak W. and K. Zorychta (1993) Modular Optimizer for Mixed Integer Programming - MOMIP, WP-93-xx, International Institute for Applied Systems Analysis, Laxenburg, Austria, (to appear).

[22] Plapp C. (1990) ORVAN für OS/2, Referenzhandbuch, Institut für Textil- und Verfahrenstechnik, Denkendorf, Germany.

[23] Rapoport A., (1989) Decision Theory and Decision Behaviour, Normative and Descriptive Approaches, Theory and Decision Library, Kluwer.

[24] Ruszczyński A. (1993) Water quality management: problem formulations and solution methods, WP-93-36, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[25] Stam, A., Cesar, H., and M. Kuula (1989) Transboundary Air Pollution in Europe: An Interactive Multiple Criteria Tradeoff Analysis, WP-89-61, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[26] Somlydoy, L., and Paulsen, C.M. (1992) Cost-Effective Water Quality Management in Central and Eastern Europe, WP-92-091, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[27] Somlyody L. and M. Kularathna (1993), Decision Support System for River Basin Water Quality Management, WP-93-xx, International Institute for Applied Systems Analysis, Laxenburg, Austria, (to appear).

[28] Somlyody L., M. Kularathna and I. Masliev (1993) Water Quality Management of Nitra River Basin, Slovakia: Alternative Control Strategies and the Role of Uncertain Parameters. WP-93-xx, International Institute for Applied Systems Analysis, Laxenburg, Austria, (to appear).

[29] Somlyody L., M. Kularathna and I. Masliev (1994) Development of Least-Cost Water Quality Control Policies for the Nitra River Basin in Slovakia, Paper submitted to the 1994 Budapest Biennial conference of the International Association of Water Quality (IAWQ).

[30] Stewart, T.J. (1992) A Critical Survey on the Status of Multiple Criteria Decision Making Theory and Practice, International Journal of Management Science, 20(5/6), 569-586.

[31] Thomann, R.V. and Mueller, J.A. (1987) Principles of Surface Water Quality Modeling and Control, Harper and Row, New York.

[32] Wessels J. and A.P. Wierzbicki, Eds. (1992) User-Oriented Methodology and Techniques of Decision Analysis and Support, Lecture Notes in Economics and Mathematical Systems, vol. 397, Springer Verlag.

[33] Wierzbicki, A.P. (1980) The use of reference objectives in multiobjective optimization. In G. Fandel and T. Gal, eds., Multiple Criteria Decision Making, Theory and Applications, Springer Verlag, Heidelberg.

[34] Wierzbicki, A.P. and M. Makowski (1992) Multi-Objective Optimization in Negotiation Support, WP-92-07, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[35] Wierzbicki, A.P. (1992) The Role of Intuition and Creativity in Decision Making, WP-92-78, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[36] Wierzbicki, A.P. (1992) Multi-Objective Modeling and Simulation for Decision Support, WP-92-80, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[37] Vincke P. (1989) Multicriteria Decision-aid, John Wiley and Sons, New York.

[38] Yu, P.L., (1990) Forming Winning Strategies, An integrated Theory of Habitual Domains, Springer Verlag.

[39] Zeleny, M. (1992) An Essay into Philosophy of MCDM: A Way of Thinking or Another Algorithm, Computers Operations Research, 19(7), 563-566.

9 Appendix: Mathematical Programming Problem

This Appendix contains a formulation of the mathematical programming problem equivalent to the formulation in Section 2. Section 9.1 contains equations converted to the form corresponding to the standard formulation of an LP (Linear Programming) problem. Section 9.2 contains definitions of names used for rows and columns in the LP problem formulation. Finally, Section 9.3 contains definitions of goal functions used for different optimization problems.

9.1 Conversion of equations

The equations defined so far have been defined in a way that that is easy to interpret. However, most of them had to be converted to a form that is accepted for the formulation of an optimization problem. In order to document those conversions, all of the equations used for in the definition of optimization problems are listed below in this new form.

To remind the reader, the following definition of sets is being used in the equations:
M – is set of indices of all monitoring nodes
E – is set of indices of all emissions nodes
J – is set of indices of all nodes
L – is set of indices of all water quality constituents
$K(j)$ – is set of indices of all technologies considered at the emissions node j
$I(j)$ – is set of indices of all nodes located immediately up-stream of node j

The indices of water quality, l, are as follows:
0. DO, dissolved oxygen
1. CBOD, carbonaceous biological oxygen demand
2. NBOD, nitrogeneous biological oxygen demand
3. NH4, ammonia
4. P, dissolved phosphorous
5. SOD, sediment oxygen demand.

The only equation without changes is eq (1):

$$\sum_{k \in K(j)} x_{jk} = 1 \qquad x_{jk} \in \{0,1\}, \qquad j \in E$$

Equation (2), $wq_{j0} = (aqs_{j0} - aq_{j0})/aqs_{j0}$ is converted to:

$$wq_{j0} + aq_{j0}/aqs_{j0} = 1. \qquad j \in M \tag{31}$$

Equation (3), $wq_{jl} = (aq_{jl} - aqs_{jl})/aqs_{jl}$ is converted to:

$$wq_{jl} - aq_{jl}/aqs_{jl} = -1. \qquad l \in [1,4] \qquad j \in M \tag{32}$$

Equation (4), $g_l = \max_{j \in M}(wq_{jl})$ is converted to:

$$wq_{jl} - g_l \leq 0. \qquad l \in [0,4] \qquad j \in M \tag{33}$$

and equation (5) $g_{all} = \max_{l \in [0,4]}(g_l)$ is converted to:

$$g_l - g_{all} \leq 0. \qquad l \in [0,4] \tag{34}$$

Consult Section 9.3 for additional information about the last two conversions.

Equation (6):

$$aq_{j0} = DOsat_j - 1./Q_j * \sum_{i \in I(j)} Q_i * (TC_{i0}(DOsat_i - aq_{i0}) + \sum_{l \in \{1,2,5\}} TCp_{il} aq_{il})$$

now has the form:

$$aq_{j0} + 1./Q_j * \sum_{i \in I(j)} Q_i * (-TC_{i0} aq_{i0} + \sum_{l \in \{1,2\}} TCp_{il} aq_{il}) = \tag{35}$$

$$= DOsat_j - 1./Q_j * \sum_{i \in I(j)} Q_i * (TC_{i0} * DOsat_i + TCp_{i5} SOD_i) \qquad j \in J$$

where SOD_i is a given SOD at i-th node.

Equation (7), $aq_{jl} = b_{jl} + (\sum_{i \in I(j)} TC_{il} aq_{il} Q_i + e_{jl})/Q_j$ is converted to:

$$aq_{jl} - (e_{jl} + \sum_{i \in I(j)} TC_{il} aq_{il} Q_i)/Q_j = b_{jl} \qquad l \in [1,4] \qquad j \in J \tag{36}$$

and it is assumed that $TC_{i3} = TC_{i2}$ and $TC_{i4} = 1$.

Equation (9), $e_{jl} = q_j \sum_{k \in K(j)} x_{jk} em_{jlk}$ is converted to:

$$e_{jl} - q_j \sum_{k \in K(j)} x_{jk} em_{jlk} = 0. \qquad l \in [1,4] \qquad j \in E \tag{37}$$

Equation (10), $Inv_j = \sum_{k \in K(j)} x_{jk} IC_{jk}$ takes the form:

$$Inv_j - \sum_{k \in K(j)} x_{jk} IC_{jk} = 0. \qquad j \in E \tag{38}$$

Equation (11), $OM_j = \sum_{k \in K(j)} x_{jk} OMC_{jk}$ is now:

$$OM_j - \sum_{k \in K(j)} x_{jk} OMC_{jk} = 0. \qquad j \in E \tag{39}$$

Equation (12), $TAC_j = [r(r+1)^n/((r+1)^n - 1)]Inv_j + OM_j$ is now:

$$TAC_j - [r(r+1)^n/((r+1)^n - 1)]Inv_j - OM_j = 0. \qquad j \in E \tag{40}$$

Equation (13), $Tot\_Inv = \sum_{j \in E} Inv_j$ is converted to:

$$Tot\_Inv - \sum_{j \in E} Inv_j = 0. \tag{41}$$

Equation (14), $Tot\_OM = \sum_{j \in E} OM_j$ is now:

$$Tot\_OM - \sum_{j \in E} OM_j = 0. \tag{42}$$

Equation (15), $Tot\_TAC = \sum_{j \in E} TAC_j$ becomes:

$$Tot\_TAC - \sum_{j \in E} TAC_j = 0. \tag{43}$$

9.2 Names in the MPS formulation

9.2.1 General remarks

All names are generated automatically by the problem generator, and their compositions are currently pre-specified according to the following rules:

- Names are composed of an id followed by a node number (j) and a constituent number (l). For rows related to a multicriteria problem, a criterion number is used. Whenever any of these numbers is not relevant, it is skipped.
- Numbers are counted from 0.
- Name's id and numbers are separated by _ (an underscore).
- Every name is exactly 8 characters long, so dots are appended to any name which is shorter than 8 characters.

9.2.2 Rows

The following id's are used for the rows, which are generated in the sequence given below. The variables *n_nodes*, *n_mon*, and *n_emm* denote the the number of all nodes, monitoring, and emission nodes, respectively. The variables *n_waste* and *n_crit* are the numbers of water quality constituents and criteria, respectively.

goal – one goal function row,

x – n_emm rows for eq. (1),

<MCR – n_crit rows for conversion of multicriteria problem into an equivalent single criterion (cf e.g. [17] for the description of the applied conversion method),

g_all – n_waste rows for eq. (34),

g – n_mon * n_waste rows for eq. (33),

wq – n_mon * n_waste rows for eq. (32),

aq – n_nodes * n_waste rows for eq. (35) and (36),

e – n_emm * n_waste rows for eq. (37),

inv – n_emm rows for eq. (38),

om – n_emm rows for eq. (39),

tac – n_emm rows for eq. (40),

tot_inv – n_emm rows for eq. (41),

tot_om – n_emm rows for eq. (42),

tot_tac – n_emm rows for eq. (43).

9.2.3 Columns

The following id's are used for the columns, which are generated in the sequence given below. In addition to the numbers defined in the previous section, *n_tech* denotes the number of all technologies considered at all emission points.

x – n_emm * n_waste columns for decision variables

<M_CRIT – one column for conversion of multicriteria problem into an equivalent single criterion,

g_all – one columns for the regional water quality index g_{all}

g – n_waste columns for water quality indices (for each type of waste),

wq – n_mon * n_waste columns for water quality indices,

aq – n_nodes * n_waste columns for ambient water quality concentrations,

e – n_emm * (n_waste - 1) columns[16] for amounts of discharged constituents,

inv – n_emm columns for investment costs,

om – n_emm columns for O&M costs,

tac – n_emm columns for TAC,

tot_inv – one column for total investment costs,

tot_om – one column for total O&M costs,

tot_tac – one column for total TAC.

[16]No columns are generated for DO.

9.3 Goal functions in the MPS formulation

There is a predefined coefficient ϵ (currently set to 0.001) which is used for conversion of equations (4), (5) into equations (33), (34) respectively. Therefore, unless the respective term enters the goal function because of the definition of a single criterion problem or of the achievement scalarizing function, for every formulation the goal function is augmented by the term:

$$\epsilon * (g_{all} + \sum_{i=0}^{n\_waste} g_i) \tag{44}$$

For single criterion problems, the goal functions contain either tot_tac or g_{all}, depending on whether the minimization of costs or of the regional water quality index is performed. For the definition of a goal function for the multicriteria problems, consult [17].