STOCHASTIC OPTIMIZATION MODELS FOR LAKE EUTROPHICATION MANAGEMENT

László Somlyódy*
Roger J-B Wets**

April 1985
CP-85-16

* Research Center for Water Resources Development, VITUKI, Budapest, Hungary
** IIASA and University of California
(Supported in part by a grant of the National Science Foundation)

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

FOREWORD

The development of stochastic optimization techniques, related software, and its application, was a major part among the activities of the System and Decision Sciences Program during the past few years.

This paper describes results of the application of stochastic programming to water quality management. It provides an example of both important issues for investigation: a realistic problem with inherent stochasticity, and a valuable test problem for the algorithms under development. It also gives some insights into the nature of solutions of certain classes of stochastic programming problems, and the justification for the consideration of randomness in decision models.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences
Program

ABSTRACT

We develop a general framework for the study and the control of the eutrophication process of (shallow) lakes. The randomness of the environment (variability in hydrological and meteorological conditions) is an intrisic characteristic of such systems that cannot be ignored in the analysis of the process or by management in the design of control measures. The models that we suggest take into account the stochastic aspects of the eutrophication process. An algorithm, designed to handle the resulting stochastic optimization problem, is described and its implementation is outlined. A second model, based on expectation-variance considerations, that approximates the "full" stochastic model can be handled by standard linear or nonlinear programming packages. A case study illustrates the approach; we compare the solutions of the stochastic models, and examine the effect that randomness has on the design of good management programs.

Key Words: Lake eutrophication, eutrophication management, stochastic programming, recourse model, water quality management, probabilistic constraints.

CONTENTS

STOCHASTIC OPTIMIZATION MODELS FOR LAKE EUTROPHICATION MANAGEMENT

by

László Somlyódy and Roger J-B Wets

INTRODUCTION

Man-made (or artificial) eutrophication has been considered as one of the most serious water quality problems of lakes during the last 10-20 years. Increasing discharges of domestic and industrial waste water and the intensive use of crop fertilizers -- all leading to growing nutrient loads of the recipients -- can be mentioned among the major causes of this undesirable phenomenon. The typical symptoms of eutrophication are among others sudden algal blooms, water coloration, floating water plants and debris, excreation of toxic substances causing taste and odor problems of drinking water and fish kills. These symptoms can easily result in limitations of water use for domestic, agricultural, industrial or recreational purposes.

One of the major features of artificial eutrophication is that although the consequences appear within the lake, the cause -- the gradual increase of nutrients (various phosphorous and nitrogen compounds) reaching the lake — and most of the possible control measures lie in the region. Consequently, eutrophication management requires analysis of complex interactions between the water body and its surrounding region. In the lake, different biological, chemical and hydrophysical processes -- all being time and space dependent, furthermore non-linear — are important, while in the region one must take into account human activities generating nutrient residuals and control measures determining that portion of the emission which reaches the water body.

Eutrophication management requires a sound understanding of all these processes and activities which, in fact, belong to quite diverse disciplines. Additionally, various uncertainties and stochastic features of the problem have to be also taken into account, for example, the estimation of loads from infrequent observations and the dependence of water quality on hydrologic and meteorologic factors, respectively. The fact that we are dealing with a stochastic environment, is especially important for shallow lakes, due primarily to the absence of thermal stratification which predicates a much more definite response to randomness as would be the case for deep lakes.

Models developed with the aim of analyzing eutrophication can be classified into two broad groups:

(a) Dynamic simulation models (ecological models) which intend to describe the temporal and spatial changes of various groups of species and nutrient fractions (algae, zooplankton, phosphorus, and nitrogen components, etc.). These models are being designed primarily for research purposes.

(b) Management optimization models being more applications oriented which have as objective of the determination of the best or "optimal" combination of alternative control measures.

There is a large literature devoted to descriptive models (see for example, Canale, 1976; Scavia and Robertson, 1979; Dubois, 1981; Orlob, 1983), but only relatively few management models have been proposed (Thomann, 1972; Biswas, 1981; Loucks et al. 1981; Bogárdi et al., 1983). Although changes in the ambient lake water quality should be an important element in decision making, to our knowledge, no lake is mentioned in the literature for which a proper description of the in-lake processes has been taken into account in the management model; i.e., the interaction between the models of type (a) and (b) is missing. The reason for this is at least threefold:

(i) Models (a) are too complex to be involved directly in optimization and no well-based methodology exists in this respect.

(ii) In general, for management purposes, it suffices to use some aggregated features of water quality. As will be seen, microscopic details offered by models (a) can be often ruled out, but it is not easy to determine that portion of information offered by the eutrophication model (a) which should be maintained in order to arrive at a scientifically well-based management model.

(iii) In most of the cases, not enough time and money are available to perform a systematic analysis including the joint development of both models.

Consequently, there is a gap between the application of descriptive and management models; the objectives of decision models are formulated in terms of nutrient loads rather than of lake water quality, which is only acceptable for relatively simple situations. The present paper — after considering management goals in Section 1 — offers an approach which allows the combined use of descriptive, simulation and management optimization models. This is discussed in Section 2. The derivation of the aggregated lake and planning type nutrient load models to be used in the management model is the subject of Section 3. Both models are stochastic, special emphasis is given to shallow lakes. Section 4 discusses further elements of the management model (cost functions, constraints, etc.). Alternative management models are formulated in Sections 5 and 6. Two of them were implemented: a "true" stochastic model (which uses as the starting point of the iterative solution procedure the corresponding deterministic model) and a linear programming approach capturing stochastic features of the problem through expectation and variance. Sections 6 and 7 give details of these models. Finally, in Section 8 the methods developed are used primarily for Lake Balaton that plays the role of a major case study.

1. MANAGEMENT GOALS AND OBJECTIVES

As mentioned in the Introduction, artificial eutrophication leads to water quality changes which then restricts the use of water. The objective of a manager when considering an eutrophication problem very much depends on the features of the particular system. In most of the cases, however, the wish of managers can be formulated in quite general terms. The basis for this is the definition of trophic classes.

Trophic state classification is the subject of limnology. In practice four major classes: oligotrophic, mesotrophic, eutrophic and hypertrophic categories are used (OECD, 1982), where oligotrophic indicates relatively "clean" water, while hypertrophic refers to a lake rich in nutrients and an advanced stage of eutrophication. To specify trophic classes, water quality components (concentrations of phosphorus and nitrogen fractions, oxygen content, algal biomass, etc., see for example, OECD, 1982) are applied as indicators. Based on observations and studies performed on many shallow and deep lakes trophic classes are specified quantitatively by certain ranges of the indicators. One of the most widely used indicator is the annual mean or annual peak chlorophyll-a concentration, (Chl-a), a measure of algal biomass. The chlorophyll content affects the color of water and thus it characterizes for example the recreational value of the lake. In terms of concentration $(\mathrm{Chl-a})_{max}$ the ranges of classes oligotrophic...hypertrophic are as follows: 0-15, 10-25, 25-75 and 75- (see OECD, 1982).

As these classes are closely related to the use of water, decision makers' objective is often to shift a lake, say, from hypertrophic to an oligotrophic state and water quality goals are specified accordingly. Note that the definition of trophic classes (with fixed boundaries) is not unambiguous, which is certainly not a surprise as the trophic changes are caused by complex ecological processes. Still, however, decision makers require guidelines easy to understand and apply, and in this sense the use of trophic state classification is inevitable in eutrophication management.

For many lakes there is a spatial variation in the water quality. The type of water usage can also be different in one area of the lake than in others, e.g., agricultural, recreational or industrial. For this reasons the objective of management can be different for different segments or basins of the water body. Thus spatial segmentation is a major component of management. Similarly, the decision makers should decide how important the random fluctuations in water quality (from year to year) are as compared to expectations. Would he for instance be satisfied with drastic reduction in the "average" water quality without excluding the occurence of extreme situations or rather would he prefer to achieve a modest improvement in the mean provided that he is now able to limit the range of possible fluctuations? Moreover, how would he judge the situation if fluctuations also strongly vary in space?

Simultaneously with developing his judgement on the lake basins and management goals, the decision maker has to perform a careful analysis as to where (and when) he can control nutrient loads in the watershed that determine the trophic state of the lake? How effective are the control measures, what are the costs, benefits and associated constraints? How would a control measure taken in a subwatershed of a basin influence the water quality of all the basins (including both expectations and variances)? Finally, how would he select among alternative combinations of projects? The methodology developed in the next sections is aimed at answering such questions, and thus provide technical support to the decision process.

2. THE APPROACH

The approach to eutrophication and eutrophication management is based on the idea of decomposition and aggregation (Somlyódy, 1982 and 1983a). The first step is to *decompose* the complex system into smaller, tractable units forming a hierarchy of issues (and models), such as biological and chemical processes in the lake, sediment-water interaction, water circulation and mass exchange, nutrient loads, watershed processes and possible control measures; influence of natural, noncontrollable meteorological factors, etc. One can make detailed investigations of each of these issues. This step is followed by *aggregation*, the aim of which is to preserve and integrate only the issues that are essential for the higher level of the analysis and hierarchy, ruling out the unnecessary details. In this way a sequences of corresponding detailed and aggregated (mainly descriptive type) models are developed. Only aggregated models are coupled in an on-line fashion (the approach is off-line for the detailed models) thus allowing their incorporation in a management optimization model at the highest stratum of the hierarchy. The procedure for deriving the eutrophication management optimization model (EMOM) is illustrated in Figure 1 (for further details, see Somlyódy, 1983b; and Somlyódy and van Straten, 1985) and consists of four stages:

Phase 1

This is the development phase of the dynamic, descriptive lake eutrophication model (LEM) which has two sets of inputs: controllable inputs (mainly artificial nutrient loads) and noncontrollable inputs (meteorological factors, such as temperature, solar radiation, wind, precipitation). The output of the model is the concentrations vector y of a number of water quality components as a function of time (on a daily basis) t, and space r: $y(t,r)$. LEM is calibrated and validated by relying on historical data; the inputs of LEM are the recorded observations. Because of methodological and computational difficulties on one hand, and uncertainties in knowledge

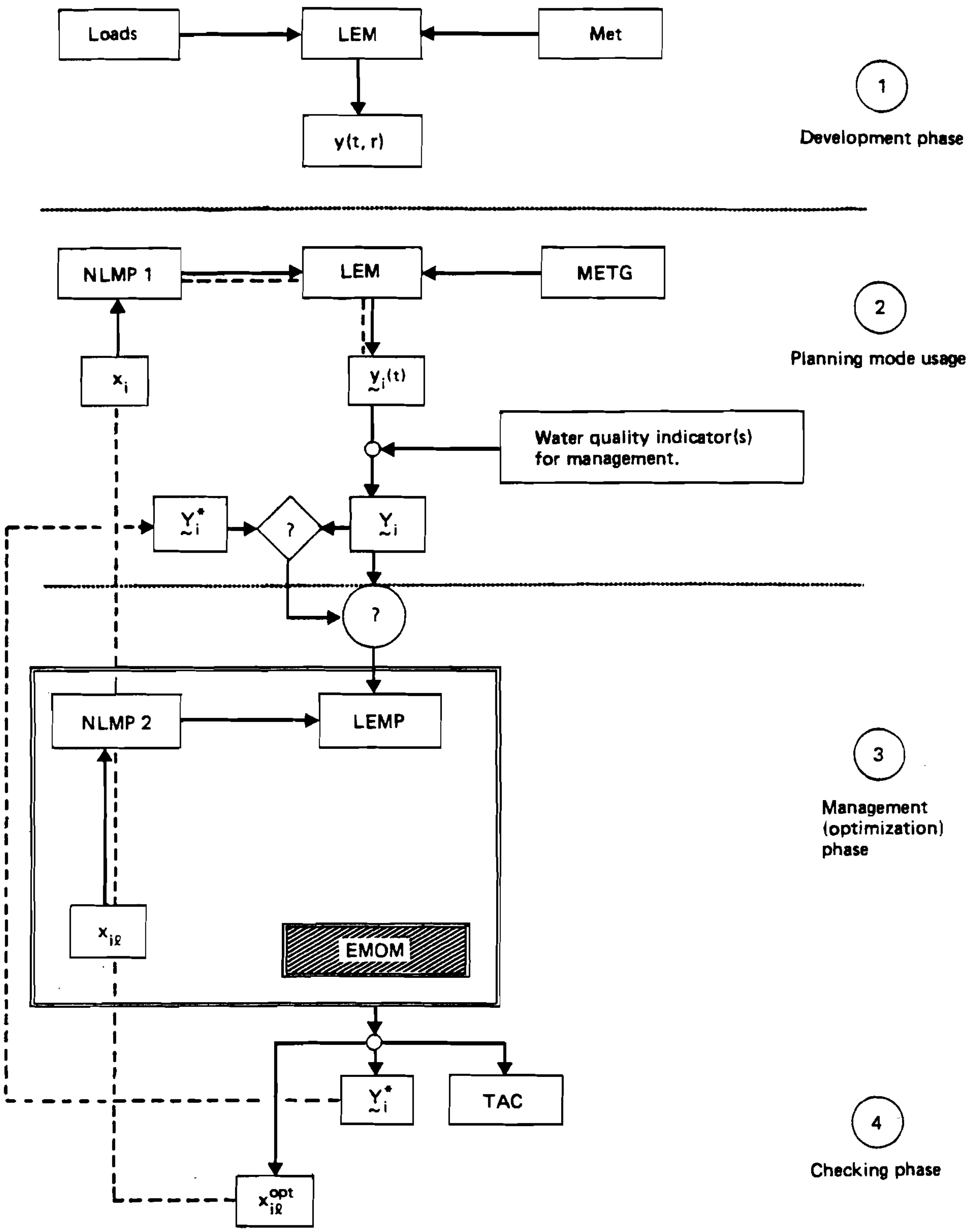


FIGURE 1. Structure of the Analysis.

and observational data on the other hand, most of the models are not "continuous" in space, but rather they describe the average water quality of lake segments (*basins*) having differing characteristics. For details, readers are referred to the literature on ecological modeling (e.g, Somlyódy and van Straten, 1985).

Phase 2

In order to apply LEM at a later stage for solving the management problem, two important steps must be taken:

(i) A decision has to be made as to the kinds of loads and meteorological inputs to be used in planning scenarios. For river water quality planning problems deterministic "critical" scenarios can be generally used, but for lake eutrophication problems (especially for shallow water bodies), the inputs should be considered stochastic functions. In Figure 1 NLMP1 and METG indicate models which generate the loads and meteorological factors in a random fashion based on the analysis of historical data. The proper time scale of NLMP1 and METG should be established in Phase 1; it generally suffices to use a month as a basis (Somlyódy, 1983b).

(ii) The planning type nutrient load model (NLMP1) incorporates aggregated control variables, for deriving the loads of individual lake basins ($1 \leq i \leq N$). They are used to determine classes of aggregate load scenarios with different expectations and fluctuations for each individual basin, but they do not express the way in which a scenario is actually realized in the watershed, The LEM can then be run systematically under stochastic inputs with different control variable vectors x_i yielding the stochastic vectors $\underset{\sim}{y}_i(t)$ that describe the temporal changes in various water quality components for the basins ($1 \leq i \leq N$) (in contrast to the $y_i(t)$ of Phase 1 which are deterministic and noncontrolled). These changes are of interest for understanding the eutrophication process, but for decision making purposes, various indicators that express the global behavior of the system can be used (see Section 1).

Experience has shown (OECD, 1982; Somlyódy, 1983b) that because of the cumulative nature of eutrophication, the indicators primarily depend on the annual average nutrient load with the (annual) dynamics of secondary importance, at least for lakes whose retention time is not too small. This means that time can be disregarded and we can work with (vector-) indicators, $\underset{\sim}{Y}_i$*, that are derived from the concentration vectors ($\underset{\sim}{y}_i(t)$, $1 \leq i \leq N$). This also implies that the LEM can be used in an off-line fashion.

Phase 3

There are various ways in which the LEM can be involved in the optimization model, and implications and conclusions requires careful analysis. For example, the $\underset{\sim}{Y}_i$ indicators obtained from systematic computer experiments can be stored as a function of the corresponding aggregated control vectors. These form a "surface" on which the "optimal" solution is built later. The other possibility, that we follow here, is to parametrize the

* Note that the number of indicators is generally less than that of the water quality components, and often a single (scalar) indicator is employed (see later).

outputs of LEM (obtained under various control vectors X_i) and to arrive at an analytical expression which then can be included in the optimization model. This is the LEMP model, cf. Figure 1.

As indicated in Figure 1, the success of this procedure is not *a priori* obvious and depends mainly on the complexity and major features of the system. Still, observations and model results have shown for several lakes that a linear relationship holds between the indicator (e.g., $((\mathrm{Chl-a})_{max})$ and the annual average phosphorus load (OECD, 1982; Somlyódy, 1983b; Lam and Somlyódy, 1983). Van Straten (1983) has proven analytically by using a simplified LEM, that for nutrient limited, turbid shallow lakes the maximum algal level is practically proportional to the phosphorus load. Based on these findings, we assume that an analytically expression, moreover a linear model, can be used as LEMP for a large variety of water bodies.

LEMP is an aggregated version of the LEM that can be employed directly for planning purposes. It describes approximately the indicators $\underset{\sim}{Y}_i$ as a function of the nutrient load, including stochastic variability. LEMP is connected to a nutrient load model, NLMP2, which includes the control variables, x_{il} $(1 \leq l \leq N_{Li};\ 1 \leq i \leq N)$. Compared with NLMP1, the NLMP2 model exhibits three significant differences:

(i) it covers only that part of the load that is thought to be controllable on the basis of available, realistic measures;

(ii) it contains more details about the subwatersheds of each basin (location of pollution discharges and control measures, etc.); and

(iii) it is aggregated with respect to time; as for LEMP the annual average load can generally be used.

The coupled NLMP2-LEMP models are then put through an optimization procedure, which generates (depending on the formulation of EMOM, the Eutrophication Management-Optimization Model) for example, "optimal" values for the control variables, x_{il}, the associated indicators $\underset{\sim}{Y}_i^*$, and—say—the total annual costs, TAC, needed for carrying out the project or fixed by budgetary considerations.

Phase 4

In the course of this procedure various simplifications and aggregations are made without a quantitative knowledge of the associated errors. Accordingly, the last step in the analysis is validation. That is, the LEM can be run with the "optimal" load scenario as indicated in Figure 1 by the dashed line, and the $\underset{\sim}{Y}_i$ "accurate" and $\underset{\sim}{Y}_i^*$ "approximate" solutions can be compared.

3. FORMULATION OF THE STOCHASTIC MODELS LEMP AND NLMP2

The modeling procedure was outlined in Section 2 in general terms. Here we continue the discussion detailing assumptions and limitations. The reason for these is threefold:

- All the models used in the field of eutrophication are system specific;
- Consequently our objective cannot be more than to capture some of the major features of eutrophication management problems; and
- We wish to avoid generalization beyond our experiences.

The major assumptions we make are as follows:

(i) the lake is shallow, the water quality of which is vertically uniform;

(ii) the lake can be subdivided into basins which are sequentially connected, see Figure 2;

(iii) the lake is phosphorus limited (like most water bodies) and thus nutrients other than P (phosphorous) are not involved in the analysis;

(iv) a single water quality indicator, the $(\mathrm{Chl-a})_{max}$ concentration is used for defining trophic state and the goal of management (see Section 1);

(v) as discussed before linear relationship holds for individual basins between $(\mathrm{Chl-a})_{max}$ and P load;

(vi) short term (a few years) management is considered, that is the renewal processes in the lake and its sediment layer following external load reduction, and the scheduling of the investments are out of the scope of the present effort;

(vii) only certain types of P sources and associated control alternatives are taken into account.

At the end of this section, we will show that some of these conditions can be relaxed if needed, and in fact the models to be introduced have a broader range of applicability than suggested by the above assumptions.

3.1. Aggregated Lake Eutrophication model (LEMP)

Based on Section 2, the assumptions made and the insight gained from the study on Lake Balaton (Section 8), the short term response of water quality to load reduction can be written as follows (Somlyódy, 1983b):

$$\underset{\sim}{Y} = E\{\underset{\sim}{Y}_o\} + \underset{\sim}{\mathbf{w}} - (D + \mathbf{d}\underset{\sim}{\mathbf{w}})\, \Delta\underset{\sim}{L}\,, \tag{3.1}$$

where the elements of the N-vector $\underset{\sim}{Y}$ represent the (to be controlled) water quality in the N basins — in terms of $(\text{Chl-a})_{\max}$ [mg/m^3] — and $\underset{\sim}{Y}_o$ refers to the (noncontrolled) nominal state; E is expectation. In the equation, the N-vector $\Delta\underset{\sim}{L}$ expresses the change in load due to control

$$\Delta\underset{\sim}{L} = E\{\underset{\sim}{L}_o\} - \underset{\sim}{L} \tag{3.2}$$

where the elements of $\underset{\sim}{L}$ are the annual mean volumetric biologically available P load, BAP [mg/m^{3}d] in each basins ($\underset{\sim}{L}_i = \underset{\sim}{L}_i^a / V_i$, here $\underset{\sim}{L}_i^a$ is the "absolute" load [mg/d] and V_i is the volume [m^3]). The BAP load covers the P fractions that can be taken up directly by algae and thus determine the short term response of the water body. Stochastic variables and stochastic parameters are bold faced. The random N-vector $\underset{\sim}{\mathbf{w}}$ represents the random changes of water quality caused by noncontrollable meteorological factors. Finally the elements of the square $N\times N$-matrix D and the vector $\mathbf{d}$ are derived from the analysis and simulation in Phase 2 (see Figure 1).

The elements of matrix D are the reciprocals of lumped reaction rates. The main diagonal comprises primarily the effect of biological and biochemical processes, while the other elements refer to those of interbasin exchange due to hydrological throughflow and mixing. These elements express that due to water motion a control measure taken on subwatershed of basin i will affect the water quality of other basins. As will be seen in Section 8, the diagonal elements d_{ii} of the matrix D (the slopes of linear load response relationships) are such that $(\mathrm{Chl-a})_{\max}$ does not necessarily approach zero (or a relatively small value) if $\underset{\sim}{L}_i$ goes to 0. The reason is that $(\mathrm{Chl-a})_{\max}$ is linearly related to the sum of external and internal loads. The internal load is the P release of the sediment (a consequence of P accumulation during preceding years and decades) which practically cannot be controlled. Since external and internal loads are coupled (a reduction in external load generates a time-lagged reduction in the internal load) the long-term improvement of water quality is generally larger than given by Equation (3.1). The memory effect and renewal of sediment are however poorly understood (see for example, Lijklema et al, 1983), this is one of the reasons why we concentrate on short-term control. Note that due to the definition of $\Delta \underset{\sim}{L}$, (3.1) also yields the random variations of water quality in the (noncontrolled) nominal state. The character of slopes d_i is similar to that of the diagonal elements d_{ii}, the term $d_i \, \underset{\sim}{w}_i \, \Delta \underset{\sim}{L}_i$ expresses a change, linear in $\Delta \underset{\sim}{L}_i$, in the random component of the water quality indicator in basin i. Of course, the effect of the random fluctuations $\underset{\sim}{w}_i$ caused by meteorology decreases if the sum of external and internal loads diminishes. This also means that with new control measures the water quality of a lake may approach a "new equilibrium" via major fluctuations, as observed in nature.

In view of (3.1) water quality varies randomly for three reasons:

(i) random changes in meteorological factors (primarily solar radiation and temperature) (the distributions of the $\underset{\sim}{w}_i$, typically skewed, are obtained in Phase 2 of the procedure (Figure 1) and can generally be approached by three-parameter gamma distributions);

(ii) stochastic changes and uncertainties in the loads (see below); and

(iii) the combined effect of climatic and load factors.

Relation (3.1) gives the aggregated lake eutrophication model. The model takes into account on a macroscopic level the effect of biological and biochemical processes, interbasin mass exchanges, the sediment furthermore the influence of stochastic factors and uncertainties.

3.2 Nutrient Load Model (NLMP2)

We consider three P sources as indicated in Figure 2:

(i) direct sewage P load, L_S;

(ii) indirect sewage P load when the recipient is a tributary of the lake, L_{SN} (both L_S and L_{SN} can be considered biologically available and deterministic); and

(iii) tributary load to which contribute various point sources (sewage discharges) and non-point sources of the watershed.

The biologically available portion of the tributary load is

$$\underset{\sim}{L}_R = \underset{\sim}{L}_D + \delta(\underset{\sim}{L}_T - \underset{\sim}{L}_D) \,, \tag{3.3}$$

where $\underset{\sim}{L}_D$ is the dissolved reactive P load, $\underset{\sim}{L}_T$ total P load and δ availability ratio of the particulate (not dissolved) P load (the difference of $\underset{\sim}{L}_T$ and $\underset{\sim}{L}_D$), δ is about 0.2.

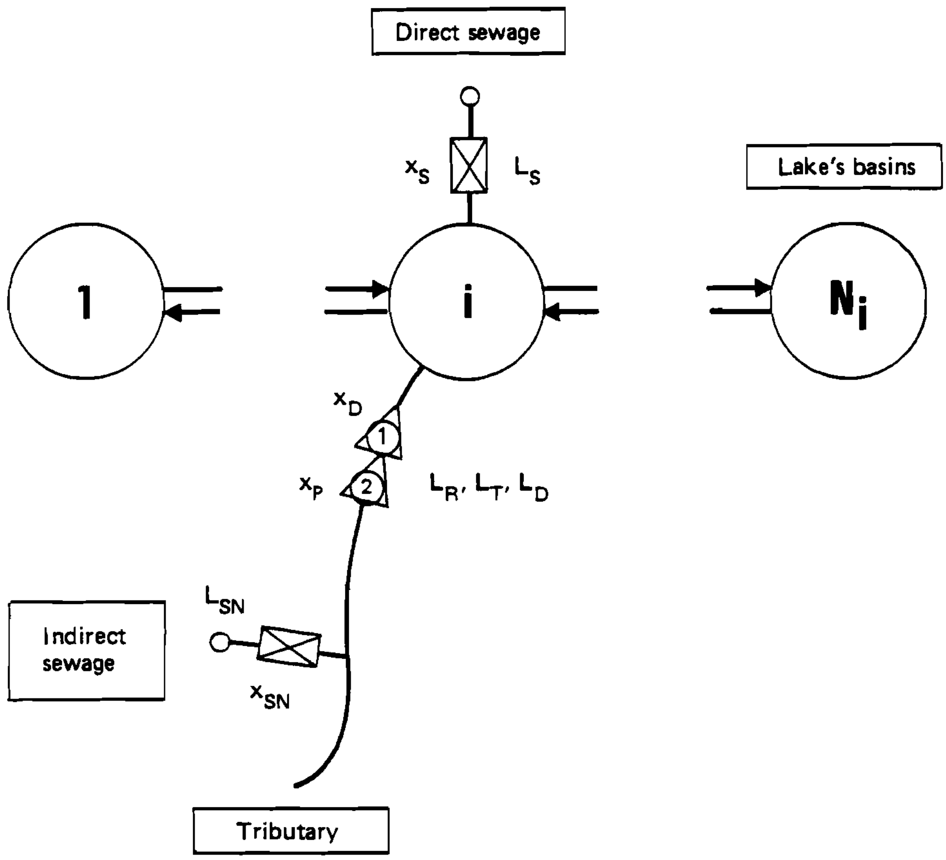


FIGURE 2. Development of models LEMP and NLMP2.

Now, let us consider the basic control options from the above P loads:

(i) and (ii)P precipitation by sewage treatment plants, and

(iii) pre-rersevoir systems established on rivers before they enter the lake. These consist of two parts (Figure 2): the removal of particulate P through sedimentation in the first segment, and the removal of dissolved P (benthic eutrophication in reed-lakes, sorption, etc.) in the second part. The corresponding control variables are x_S, x_{SN}, x_P, and x_D as indicated in Figure 2. All of them can be thought as removal coefficients, with

$$0 \leq r^- \leq x \leq r^+ \leq 1 , \tag{3.4}$$

where r^- and r^+ are lower and upper limits respectively, and if $x = 0$, no action is taken.

Now consider the simple situation given in Figure 2 for the ith basin of the lake. The original, uncontrolled load, $\underset{\sim}{L}_i$ can be expressed as follows:

$$\underset{\sim}{L}^a_{0,i} = \underset{\sim}{L}_D + \delta(\underset{\sim}{L}_T - \underset{\sim}{L}_D) + L_S + L_{SN} + \underset{\sim}{L}_{NC} , \tag{3.5}$$

where $\underset{\sim}{L}_{NC}$ is the portion of the load that is beyond the controls considered (e.g., atmospheric pollution); for the sake of simplicity, we drop the index i from the right hand side of Equation (3.5). The controlled load of the i-th basin is

$$\begin{aligned} \underset{\sim}{L}^a_i = (1 - x_D) \left[\underset{\sim}{L}_D - (1 - r_t) x_{SN} L_{SN}\right] + \delta(1 - x_P)(\underset{\sim}{L}_T - \underset{\sim}{L}_D) \\ + (1 - x_S) L_S + \underset{\sim}{L}_{NC} , \end{aligned} \tag{3.6}$$

where r_t is the retention coefficient. The expression $(1 - r_t)$ replaces a river P transport model and it defines that portion of P that reaches the lake from an indirect sewage discharge at a given point on the tributary ($r_t = 0$ means no retention). It is apparent from Equation (3.6) that the tributary load can be controlled by P precipitation and/or pre-reservoirs. The latter influence linearly both expectation and variance of $\underset{\sim}{L}_i$, while P precipitation only influences the expectation, thus there is an obvious trade-off between these alternatives. With equations (3.5) and (3.6) we can now obtain $\Delta\underset{\sim}{L}_i$ (recall that $\Delta\underset{\sim}{L}_i = \Delta\underset{\sim}{L}^a_i / V_i$):

$$\Delta \underset{\sim}{L}_i^a = x_D \overset{①}{\left[E\{\underset{\sim}{L}_D\} - (1-r_t)\, x_{SN}\, L_{SN}\right]} + (x_D - 1) \overset{②}{\left[\underset{\sim}{L}_D - E\{\underset{\sim}{L}_D\}\right]} \tag{3.7}$$

$$+ \delta \overset{③}{\left\{\left[(x_P - 1)\underset{\sim}{L}_T + E\{\underset{\sim}{L}_T\}\right] - \left[(x_P - 1)\underset{\sim}{L}_D + E\{\underset{\sim}{L}_D\}\right]\right\}}$$

$$+ \overset{④}{(1 - r_t)\, x_{SN}\, L_{SN}} + \overset{⑤}{x_S\, L_S} \, .$$

Terms ① and ④ express reduction in the expectation of the rivers dissolved P load; term ② represents the effects of the fluctuations of the $\underset{\sim}{L}_D$ load; term ③ gives the modification in the particulate P load of the river, while term ⑤ exhibits influence of direct sewage control. If we set all the control variables x to zero in Equation (3.7), we obtain the fluctuations in the original, noncontrolled load, the expectation of which is zero.

Equation (3.7) is nonlinear in the control variables because of the product term $x_D \cdot x_{SN}$, which may cause difficulties when this relation is used in an optimization scheme. Several possibilities are available to overcome this nonlinearity. For example, a new variable

$$x^a = x_D \cdot x_{SN} \tag{3.8}$$

can be introduced, a linear function of x_D and x_{SN}, which is then included in the constraint equation, see e.g.,Loucks et al., 1981. In such a case the optimization requires a parametric analysis involving this variable.

Another possibility is offered by the surface dependent character of the P removal in the pre-reservoir (second element). Generally, for a reed lake one cannot estimate more than the P removal per unit of surface-area, independent of the inflow concentration. Under this approximation, however, x_D can be defined in terms of the original uncontrolled load of the inflow (in contrast to the the conventional definition: actual inflow minus outflow divided by the actual inflow) as a variable which is not influenced by indirect sewage control. The price for such an elimination of nonlinearity is twofold:

(i) An upper limit should be specified for x_D which states that no more nutrients can be removed than those that reach the lake via the particular tributary. In terms of expectations the constraint equation can be written as

$$x_D \, E\{\underset{\sim}{L}_D\} + r_t \, x_{SN} L_{SN} \leq E\{\underset{\sim}{L}_D\}; \tag{3.9}$$

(we note that in a more precise sense the above condition should actually be fulfilled for all the realizations of $\underset{\sim}{L}_D$).

(ii) A new variable x_D^u should be introduced with $x_D \leq x_D^u$ to take into account the fact that the impact of the reservoir on the fluctuation (Equation (3.7), term 2 is not restricted by the (physical) constraints introduced by (3.9).

For a more general situation than illustrated in Figure 2, when the ith lake basin is fed by N_1 direct sewage discharges ($1 < n \leq N_1$) and N_2 tributaries ($1 \leq m \leq N_2$), each with M_m indirect sewage discharges ($1 \leq l \leq M_m$), Equation (3.7) becomes:

$$\underset{\sim}{\Delta L}_i^a = \sum_{m=1}^{N_2} \Big\{ x_D^m \Big[E\{\underset{\sim}{L}_D^m\} - \sum_{l=1}^{M_m} (1 - r_l^{ml})\, x_{SN}^{ml}\, L_{SN}^{ml} \Big] \tag{3.10}$$

$$+ (x_D^m - 1)\, \Big[\underset{\sim}{L}_D^m - E\{\underset{\sim}{L}_D^m\} \Big]$$

$$+ \delta \Big\{ \Big[(x_P^m - 1)\, \underset{\sim}{L}_T^m + E\{\underset{\sim}{L}_T^m\} \Big] - \Big[(x_P^m - 1)\, \underset{\sim}{L}_D^m + E\{\underset{\sim}{L}_D^m\} \Big] \Big\}$$

$$+ \sum_{l=1}^{M_m} (1 - r_l^{ml})\, x_{SN}^{ml} L_{SN}^{ml} \Big\} + \sum_{n=1}^{N_1} x_S\, L_S\,.$$

Equation (3.10) is actually the final NLMP2 model except that we have not yet discussed the derivation of the stochastic load components $\underset{\sim}{L}_T$ and $\underset{\sim}{L}_D$, which is in itself a difficult problem because insufficient (infrequent) observations, short historical data, and our lack of understanding (see e.g. Haith, 1982; Beck, 1982). Since the annual means are used for $\underset{\sim}{L}_T$ and $\underset{\sim}{L}_D$ and the dynamics are less important, the recommendation is to derive the loads from a regression type analysis as functions of the major hydrologic and watershed parameters (error terms should also be included).

Observations and careful analysis of the composition of the load (point vs non-point source contributions) and watershed are required for such a procedure. In general $\underset{\sim}{L}_T$ and $\underset{\sim}{L}_D$ have (positive) lower bounds and can be characterized by strongly skewed distributions. Very often they can be expressed as simple functions of the annual mean streamflow rates, $\underset{\sim}{Q}$, the statistics of which are generally known from much longer records than those available for loads. This way the basic stochastic influence of hydrologic regime can be involved. In many cases, annual means of $\underset{\sim}{L}$ and $\underset{\sim}{Q}$ are estimated from scarce observations. The uncertainty associated can be investigated from a Monte Carlo type analysis or basic statistical considerations (Cochran, 1962; Somlyódy and van Straten, 1985). Taking into account all these factors, the loads can have rather complex distributions composed of various normal, log-normal, gamma,... distributions.

The model LEMP will be used together with NLMP2 given by Equation (3.10) as major components of the eutrophication management-optimization model, EMOM. Let us note that after introducing (3.8) or (3.9) into NLMP2, the water quality indicator will be expressed by the coupled model NLMP2-LEMP as a linear function of the decision variables, x_{il}.

To conclude this section, we return to assumptions (i)-(vii) and check their "rigour" in the light of the knowledge gained in this section:

(i) shallowness is not restrictive from the point of view of using NLMP2;

(ii) non-sequential connection of the basins would simply involve a change in the structure of D;

(iii) P is practically the only element to be controlled even in cases when e.g., nitrogen is limiting algal growth (OECD, 1982; Herodek, 1983); and

(vi) short-term management basically determines the long-term behavior of the lake water quality, thus the results here have strong implications for long-term management;

(vii) controls disregarded here (e.g., sewage diversion) have similar features from a methodological viewpoint than those handled in the present paper.

4. CONTROL VARIABLES, COST FUNCTONS AND CONSTRAINTS

We now consider the management options that are available to control the eutrophication process and the restrictions that limit their choice.

4.1. Control Variables

We already touched upon some of the decision variables as removal rates of sewage treatment plants and reservoirs. Here we give a more detailed discussion of possible control variables, including their character and associated bounds.

Sewage Treatment

(i) If we consider artificial eutrophication as a problem of the lake-watershed system and we are not interested in the details of the engineering design of each treatment plant, x_S and x_{SN} should be handled as real-valued variables; otherwise all the elements of the technological process of all the plants should have been taken into account with the aid of $\{0,1\}$ variables. Lower bounds for x_S and x_{SN} often exist, since in many countries the effluent P concentration is fixed by standards (between 0.5-2 g/m^3) specified by environmental agencies.

(ii) Because of historical reasons — at least in developed countries — P precipitation (chemical treatment) is going to be realized in existing plants designed originally for biological treatment. If the efficiency of he biological treatment is unsatisfactory, which is often the case, it should be upgraded prior to introducing chemical treatment. If decision variable associated to upgrading is x_u, this type of treatment plant management can be taken into account by the condition

$$\operatorname{sgn} x_u \geq \operatorname{sgn} x_S, \tag{4.1}$$

or by

$$x_S \geq x_u \tag{4.2}$$

if x_u is a real-valued variable.

(iii) If the problem incorporates not only decision about P precipitation and the associated upgrading, but also the design of the sewer network, this would again lead to an engineering type planning task (e.g., Kovács et al, 1983) that differs from our present considerations. If, however, we wanted to involve sewer network design in the eutrophication management in a simple way, the network of the subregions could be ordered so as to correspond to treatment plants and handled similarly as explained in item (ii), after introducing decision variables x_{SE}.

(iv) Finally, a rough picture on the spatial distribution of treatment plants can be obtained defining corresponding $\{0,1\}$ variables (0 means again no action).

Pre-reservoirs

(i) As indicated in Section 3, basic P removal processes are surface dependent, thus control variables x_D and x_P specify the size of reservoirs. Lower bounds are generally given by the smallest reasonable retention time that corresponds to a given size.

(ii) Variables x_D and x_P must often be handled as $\{0,1\}$ variables.

4.2. Cost Functions

Costs and cost functions are discussed in the same order as decision variables above.

Sewage Treatment

(i) The costs of *P* precipitation (including investment and operation costs being site specific) grow exponentially with increasing removal rates and decreasing effluent concentrations (see e.g., OECD, 1982; Monteith et al., 1980; Schüssler, 1981). The most straightforward method to approach this type of cost functions is piecewise linearization (see e.g., Loucks et al., 1981) which requires the introduction of dummy variables.

(ii)-(iv) They are fixed costs.

Pre-reservoirs

Cost functions of reservoir systems are strongly depending on which process determines primary *P* removal, construction and operation conditions. Very often not enough knowledge is available to define a cost function other than linear and usually running costs can be neglected or can be assumed to be compensated by the benefits of the reservoirs (e.g., utilization of harvested reeds). As a summary, we can conclude that costs can be expressed as piecewise linear functions of the decision variables, an important feature from the viewpoint of building up the management model EMOM.

4.3. Constraints

As we already discussed most of the physical, technological and logical constraints — cf. (3.4), (3.8), (3.9), (4.1) and (4.2) and related explanations — we consider here only the budgetary constraint which is perhaps the most important one.

In order to select among management alternatives of different investment costs (IC), and operational, maintenance and repair costs (OC), the total annual cost (TAC) term is used

$$\mathrm{TAC} = \sum_{l} \mathrm{OC}_l + \sum_{l} \alpha_l \, IC_l \,, \tag{4.3}$$

where α_l is the capital recovery factor that depends on the discount rate and the lifetime of the project (see e.g., Loucks et al., 1981). This factor can be different for "small" and "large" projects (e.g., introduction of P precipitation in treatment plants and creation of reservoirs of considerable size, respectively) as for "large" investments governments often guarantee finance at low ("pure") interest rates. For this reason, as pointed out by Thomann (1972), α_l should be considered as a model parameter of a certain range and its influence on model performance should be tested.

In most cases the TAC is limited by budgetary considerations

$$\mathrm{TAC} \le \beta \tag{4.4}$$

or reexpressing this in terms of the control variables

$$\textstyle\sum_{i,l} c_{il}(x_{il}) \le \beta \,. \tag{4.5}$$

This constraint, which involves the piecewise linear functions c_{il}, can be replaced by a linear constraint by substituting x_{il}^1, x_{il}^2... for the variable

x_{il}, each x_{il}^j corresponding to a piece of linearity of c_{il}.

5. FORMULATION OF THE EUTROPHICATION MANGEMENT OPTIMIZATION MODEL: EMOM

As already indicated in Section 1, there are a number of variants available in the building of the management optimization model that allow us to capture the stochastic features of the water quality management problem. The NLMP2-LEMP model yields the following description of the eutrophication process:

$$\underset{\sim}{y} = E\,\{\underset{\sim}{y}_o\} + \underset{\sim}{w} - (D + d\underset{\sim}{w})\,\Delta\underset{\sim}{L}\,, \tag{3.1}$$

$$\Delta\underset{\sim}{L} = E\,\{\underset{\sim}{L}_o\} - \underset{\sim}{L}\,, \tag{3.2}$$

where for the sake of convenience, the water quality indicator is denoted by y, and for $i = 1,\ldots,N$,

$$\underset{\sim}{L}^a_{o,i} = \underset{\sim}{L}_{D,i} + \alpha(\underset{\sim}{L}_{T,i} - \underset{\sim}{L}_{D,i}) + L_{S,i} + L_{SN,i} + \underset{\sim}{L}_{NC,i} \tag{3.5}$$

$$\begin{aligned}\Delta\underset{\sim}{L}_i^a = \textstyle\sum_{m=1}^{N_2} \Big\{ & x_D^m \Big[E\{\underset{\sim}{L}_D^m\} - \textstyle\sum_{l=1}^{M_m}(1 - r_i^{ml})\, x_{SN}^{ml}\, L_{SN}^{ml}\Big] \\ & + (x_D^m - 1)\Big[\underset{\sim}{L}_D^m - E\{\underset{\sim}{L}_D^m\}\Big] \\ & + \delta\Big\{\Big[(x_P^m - 1)\,\underset{\sim}{L}_T^m + E\{\underset{\sim}{L}_T^m\}\Big] - \Big[(x_P^m - 1)\,\underset{\sim}{L}_D^m + E\{\underset{\sim}{L}_D^m\}\Big]\Big\} \\ & + \textstyle\sum_{l=1}^{M_m}(1 - r_i^{ml})\, x_{SN}^{ml}\Big\} + \textstyle\sum_{n=1}^{N_1} x_S L_S\,,\end{aligned} \tag{3.10}$$

the last relation being valid under some additional constraints, see (3.9) and the related comments. Substituting, regrouping terms and renumbering (reindexing) the control variables, we obtain an affine relation for the water quality indicators $\underset{\sim}{y}_i (1 \leq i \leq N)$ of the type

$$\underset{\sim}{y} = \underset{\sim}{T} x - \underset{\sim}{h} \tag{5.1}$$

where $\underset{\sim}{h}_i$ incorporates all the noncontrollable factors that affect the water quality $\underset{\sim}{y}_i$ in basin i, and the random coefficients associated to the x-variables in (3.10) determine the entries of the random matrix $\underset{\sim}{T}$ through the transformation: $(D + d\underset{\sim}{w})\Delta\underset{\sim}{L}$. To simplify the presentation of what follows, we string the decision variables in an n-vector $(x_1, \ldots, x_m)$, each x_j corresponding to a specific control measure affecting the load in some basin i. $\underset{\sim}{T}$ is thus a $N \times n$-matrix and $\underset{\sim}{h}$ is an N-vector. We also write

$$\underset{\sim}{y}(x,w) = T(w)\, x - h(w) \tag{5.2}$$

for the preceding equation, the notation $\underset{\sim}{y}(x,w)$ is used to stress the dependence of the water quality indicators $\underset{\sim}{y}_i (1 \leq i \leq N)$ on the decision variables x and on the existing environmental conditions, denoted by w (controllable and noncontrollable) determining the entries of T and h.

The distribution function $G_y(x, \cdot)$ of the random vector $y(x, \cdot)$ depends on the choice of the control measures $x_1,\ldots,x_n$. We could view our objective as finding x^* that satisfies the constraints and such that for every other feasible x

$$G_y(x^*, \cdot) \geq G_y(x, \cdot)\,, \tag{5.3}$$

i.e., such that for all $z \in R^N$

$$\text{prob. } [y(x^*, \cdot) < z] \geq \text{prob. } [y(x, \cdot) < z]\,.$$

If such an x^* existed it would, of course, be the "absolute" optimal solution, since it guarantees the best water quality whatever be the actual realization of the random environment. There always exists such a solution if there are no budgetary limitations: simply build all possible projects to their physical upper bounds ! If this were the case, there would be no need for this analysis and it is precisely because there are budgetary limitations that we are led to choose a restricted number of treatment plants and/or pre-reservoirs, and unless the problem is very unusual, there will be no choice of investment program that will dominate all other feasible programs in terms of the preference ordering suggested by (5.3).

We are thus forced to examine somewhat more carefully the objectives we want to achieve. We could, somewhat unreasonably, see the goal as bringing the water quality indicator to a near zero level (depending on the internal load) in all basins. This would ignore the individual characteristics of each basin, as well as the user-oriented criteria, as for example recreational versus agricultural. A much more sensible approach, as discussed in Section 1, is to choose the control measures so as to achieve certain desired trophic states. Let

$$\gamma_i, \ i = 1,\dots, N$$

be water quality goals expressed in terms of the selected indicator, $(\mathrm{Chl-a})_{max}$, each γ_i corresponding to the particular use of basin i. The sensitivity of the solution to these fixed levels γ_i would have to be a part of the overall analysis of the system. We are thus interested in the quantities:

$$[y_i(x, w) - \gamma_i]_+ \qquad \text{for } i = 1,\dots, N,$$

that measure the deviations between the realized water quality and the fixed goals γ_i, where $[\beta]_+$ denotes the nonnegative part of β:

$$[\beta]_+ \begin{array}{ll} = 0 & \text{if} \quad \beta < 0, \\ = \beta & \text{if} \quad \beta \ge 0. \end{array}$$

The vector

$$[y_i(x, \cdot) - \gamma_i]_+, \qquad i = 1,\ldots,N,$$

is random with distribution function $G(x, \cdot)$ defined on R^N, and the problem is again to choose among all feasible control measures $x_1, \ldots, x_n$, i.e., that satisfy all technological and budgetary constraints, a program x^* that generates the "best" distribution $G(x^*, \cdot)$ by which once could again mean

$$G(x^*, z) \geq G(x, z)$$

for all $z \in R^N$. As already mentioned earlier, such an x^* exists only in very unusual circumstances, and thus we must find a way to compare the distribution functions that takes into account their particular characteristics but leads to a measure that can be expressed in terms of a scalar functional.

5.1. The first possibility would be to introduce a pure ***reliability criterion***, i.e., to fix, in consultation with the decision maker, certain reliability coefficients to guide in the choice of an investment program. More specifically we would fix $0 < \alpha \leq 1$, so that among all feasible x we should restrict ourselves to those satisfying

$$\text{prob. } [y(x, \cdot) < \gamma] \geq \alpha \,, \tag{5.4}$$

where $\gamma = (\gamma_1,\ldots,\gamma_{m_2})$. Or preferably, if we take into account the fact that each basin should be dealt with separably, we would fix the reliability coefficients α_i, $i = 1,\ldots,m_2$, and impose the constraints

$$\text{prob. } [y_i(x, \cdot) < \gamma_i] \geq \alpha_i \,, \; i = 1,\ldots,m_2. \tag{5.5}$$

The scalars α or $(\alpha_i, \; i = 1,\ldots,m_2)$ being chosen sufficiently large so that we would observe the unacceptable concentration level only on rare occasions. In terms of the distribution function G, these constraints become

$$G(x,0) \geq \alpha . \tag{5.6}$$

for (5.4), and

$$G_i(x,0) \geq \alpha_i \quad \text{for } i = 1,\dots,N, \tag{5.7}$$

for (5.5) where the $G_i(x, \cdot)$ are the marginal distributions of the random variables $[y_i(x, \cdot) - \gamma_i]_+$. In the parlance of stochastic optimization models these are ***probabilistic (or chance) constraints***; one refers to (5.4) as a *joint* probabilistic constraint. To find a measure for comparing the distributions $\{G(x, \cdot), x \text{ feasible}\}$ we have a simple accept/reject criterion, namely if *at* 0 $G(x, \cdot)$ is either larger than or equal to α, or for each $i = 1,\dots,N$, $G_i(x, \cdot)$ is larger than or equal to α_i, the investment program x is acceptable and otherwise it is rejected. This means that we "compare" the possible distributions $\{G(x, \cdot), x \text{ feasible}\}$ at 1 point, but we do not consider any systematic ranking. Assuming we opt for the more natural separable version of the probabilistic constraints (5.5), we would rely on the following model for the policy analysis:

$$\begin{array}{lll}
\text{find } x \in R^n \text{ such that} & & (5.8)\\
r_j^- \leq x_j \leq r_j^+, & j=1,\dots,n, & \\
\sum_{j=1}^n a_{ij} x_j \leq b_i, & i=1,\dots,m, & \\
\text{prob } \left| \sum_{j=1}^n t_{ij}(w) x_j - h_i(w) < \gamma_i \right| \geq \alpha_i, & i=1,\dots,N, & \\
\text{and } z = \sum_{j=1}^n c_j(x_j) \text{ is minimized} & &
\end{array}$$

where as before the vector r^- and r^+ are upper and lower bounds on x, the inequalities $\sum_{j=1}^n a_{ij}\, x_j \leq b_i$ describes the technological constraints, including (3.9), and for every j

$$c_j : R \to R_+$$

is the cost function associated to project j, see (4.5). The overall objective would thus be to find the smallest possible budget that would guarantee meeting the present goals γ_i at least a portion α_i of the time.

We did not pursue this approach because it did not allow us to distinguish between situations where we almost met the preset goals γ_i and those that generate "catastrophic" situations, i.e., when some of the values of the $(y_i(x,w),\ i = 1,\ldots,N)$ would exceed by far $(\gamma_i,\ i = 1,\ldots,N)$ and for purposes of analysis of this eutrophication model this is a serious shortcoming. Let us also point out that there are also major technical difficulties that would have to be overcome. Probabilistic constraints involving affine functions with random coefficients are difficult to manage. We have only very limited knowledge about such constraints, and then only if the random coefficients $((t_{ij}(\cdot))_{j=1}^{n}, h_i(\cdot))$ are jointly normally distributed, cf. Section 1 of Wets, 1983a for a survey of the available results and the relevant references. Since in environmental problems the coefficients are generally not normally distributed random variables we could not even use the few results that are available, except possibly by replacing the probabilistic constraints by approximates ones using Chebyshev's inequality, as suggested by Sinha, cf. Proposition 1.26 in Wets, 1983a.

5.2. A second possibility is to recognize the fact that one should distinguish between situations that barely violate the desired water quality or levels $(\gamma_i,\ i = 1,\ldots,N)$ and those that deviate substantially from these norms. This suggests a formulation of our objective in terms of a penalization that would take into account the observed values of $[y_i(x,w) - \gamma_i]_+$ for $i = , \cdots ,N$. We expect such a function

$$\Psi : R^N \to R$$

to have the following properties:

(i) Ψ is nonnegative,

(ii) $\Psi(z) = 0$ if $z_i \leq 0 \quad i = 1,\ldots,N$,

(iii) Ψ is separable, i.e., $\Psi(z) = \sum_{i=1}^{N} \Psi_i(z_i)$.

This last property comes from the fact that the objectives for each basin are or may be different and there are essentially no "joint rewards" to be accrued from having given concentration levels in neighboring basins, the interconnections between the basins being already modeled through the Equation (3.1). A more sophisticated model, would still work with separate penalty functions $\left\{\Psi_1(z_1), \cdots , \Psi_N(z_N)\right\}$ but instead of simply summing these penalties, would treat them as multiple objectives. A solution to such a problem would eventually assign specific weights to each basin, making it equivalent to an optimization problem with single objective function. We shall assume that these weighting factors have been made available to or have been discovered by the model builder, and have been incorporated in the functions Ψ_i themselves; note however that the methodology developed here would apply equally well to a multiple objective version of the model. In addition, to (i)-(iii) we would expect the following properties: for $i = 1, \cdots ,N$,

(iv) Ψ_i is differentiable, with derivative Ψ_i',

(v) Ψ_i' is monotone increasing, i.e., Ψ_i is convex,

(vi) $\Psi_i'(z_i) > 0$ whenever $z_i > 0$,

- relatively small if z_i is "close" to 0,
- leveling off when z_i is much "larger" than 0.

A couple of possibilities, both with $\Psi_i(z_i) = 0$ if $z_i \le 0$, are

$$\Psi_i(z_i) = \beta_i \, z_i^2 \quad \text{if} \quad z_i \ge 0 \, ,$$

with $\beta_i > 0$,

$$\Psi_i(z_i) = \beta_i (e^{z_i} - z_i - 1) \quad \text{if} \quad z_i \ge 0 \, ,$$

also with $\beta_i > 0$.

As we see, there is a wide variety of functions that have the desired properties, what is at stake here is the creation of a (negative) utility function that measures the socio-economic consequences of the deterioration of the environment. We found that the following class of functions provided a flexible tool for the analysis of these factors. Let $\theta : R \to R_+$ be defined by

$$\theta(\tau) := \begin{cases} 0 & \text{if } \tau \le 0 \\ ½ \, \tau^2 & \text{if } 0 \le \tau \le 1 \\ \tau - ½ & \text{if } \tau \ge 1 \end{cases} \tag{5.9}$$

This is a piecewise linear-quadratic-linear function. The functions $(\Psi_i, \, i = 1,\dots,N)$ are defined through:

$$\Psi_i(z_i) = q_i \, e_i \, \theta(e_i^{-1} z_i) \quad \text{for} \quad = 1,\dots,N, \tag{5.10}$$

where q_i and e_i are positive quantities that allow us to scale each function Ψ_i in terms of slopes and the range of its quadratic component. By varying the parameters e_i and q_i we are able to model a wide range of preference relationships and study the stability of the solution under perturbation of these scaling parameters.

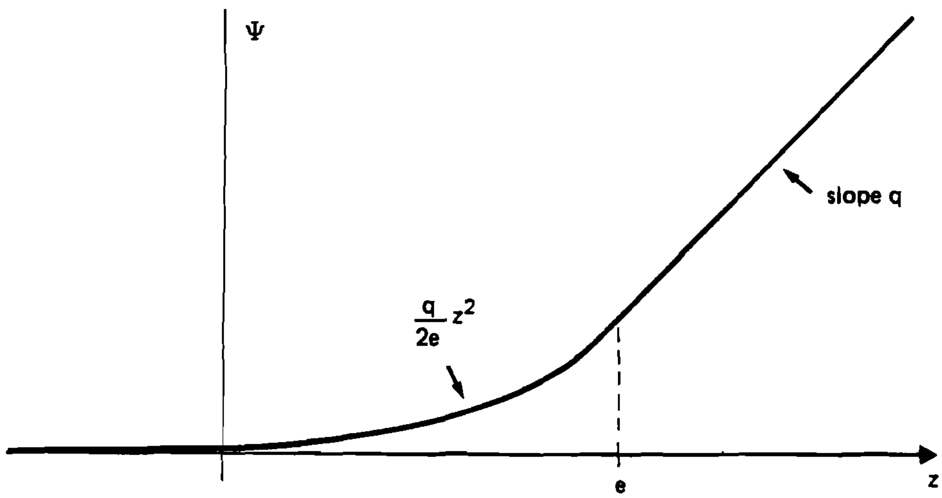


FIGURE 3. Criteria functions.

The objective is thus to find a program that in the average minimizes the penalties associated with exceeding the desired concentration levels. This leads us to the following formulation of the water quality management problem:

$$
\begin{aligned}
&\text{find } x \in R^n \text{ such that} && && (5.11)\\
&r_j^- \le x_j \le r_j^+ , && j=1,\cdots,n ,\\
&\textstyle\sum_{j=1}^n a_{ij}\, x_j \le b_i , && i=1,\cdots,m ,\\
&\textstyle\sum_{j=1}^n c_j(x_j) \le \beta ,\\
&\textstyle\sum_{j=1}^n t_{ij}(w)x_j - y_i(w) = h_i(w), && i=1,\cdots,N,\\
&\text{and } z = E\left\{\textstyle\sum_{i=1}^N q_i\, e_i\, \theta(e_i^{-1}[y_i(w)-\gamma_i])\right\} \text{is minimized}
\end{aligned}
$$

where β is the available budget. This type of stochastic optimization problem goes under the name of ***stochastic program with recourse***: a decision x (the investment program) must be chosen before we can observe the outcome of the random events (the environment modeled here by the random quantities $t_{ij}(w), h_i(w)$) at which time a recourse decision is selected so as to make up whatever discrepancies there may be; the variables y_i are just measuring the difference between $\sum_j t_{ij}x_j$ and h_i. One refers to (5.11) as a program with ***simple recourse*** in that the recourse decision is uniquely determined by the first-stage decision x and the values taken on by the random variables.

It is very important to note that no attempt has been made at combining budgetary considerations and the penalty functions that measure the deviations from the desired concentration levels in a single objective function, although there are financial considerations that may affect the choice of the coefficients q_i and e_i of the penalty terms. In our approach we handle these two criteria separately, we rely on a (discrete) parametric analysis of the solution of (5.11) as a function of β, the available budget. An essentially equivalent approach would have been to formulate (5.11) as a multiobjective program, one objective corresponding to the penalizations terms. the other to the cost function.

In terms of the distribution functions $\{G(x, \cdot), x \text{ feasible}\}$ the entire "tail" of the distributions enters into the comparison not just the value of $G(x, \cdot)$ at 0, as was the case in model (5.8) with probabilistic constraints. Indeed, the objective function can now be expressed as:

$$z = \sum_{i=1}^{N} q_i\, e_i \int_0^{\infty} \theta(e_i^{-1} s)\, dG_i(x,s)\, .$$

5.3. A third possibility is to essentially ignore the stochastic aspects of the eutrophication model and replace the random variables that appear in the formulation of the water quality management problem by fixed quantities. This would lead us to the following ***deterministic optimization problem***:

$$\begin{array}{lll}
\text{find } x \in R^n \text{ such that} & & (5.12)\\
r_j^- \le x_j \le r_j^+ & j=1,\cdots,n, & \\
\sum_{j=1}^{n} a_{ij} x_j \le b_i & i=1,\cdots,m, & \\
\sum_{j=1}^{n} c_j(x_j) \le \beta & & \\
\sum_{j=1}^{n} \hat{t}_{ij}\, x_j - y_i = \hat{h}_i & i=1,\cdots N, & \\
\text{and } z = \sum_{i=1}^{N} q_i\, e_i\, \theta\left[e_i^{-1}[y_i - \gamma_i]\right] \text{ is minimized.} & &
\end{array}$$

The choice of the paramaters $\hat{t}_{ij}$ and $\hat{h}_i$ is left to the model builder. One possibility is to choose

$$\hat{t}_{ij} = \bar{t}_{ij} = E\left[t_{ij}(w)\right],$$
$$\hat{h}_i = \bar{h}_i = E\left[h_i(w)\right],$$

i.e., replace the random quantities by their expectations. Without accepting the solution of (5.12), we could always use it as part of an initialization scheme for solving the stochastic optimization problem (5.11), and this is actually how the algorithm proceeds, see Section 7.

5.4. A fourth model — in which reliability considerations again occupy a central role but in which the shapes of the distribution functions $\{G_i(x, \cdot), i = 1, \ldots, m_2\}$ play a much more important role than just their values at one point -- allows for variables concentration levels. Again let $(\alpha_i, i = 1, \cdots, N)$ be scalars that correspond to desired reliability level. The objective is to find an investment program x such that

$$\text{prob.} \left[y_i(x, \cdot) < v_i \right] \geq \alpha_i \ , \quad i = 1, \ldots, N, \tag{5.13}$$

but now the $(v_i, i = 1, \ldots, N)$ are also decision variables that we like to choose as low as possible. There is a variety of ways measuring "as low as possible", for example by minimizing

$$\sum_{i=1}^{N} q_i [v_i - \gamma_i]_+ \tag{5.14}$$

where the q_i are nonnegative scalars that assign different importance to meeting the desired water quality goals $(\gamma_i, i = 1, \ldots, N)$ in the various basins, or by minimizing

$$\max_i \ v_i \tag{5.15}$$

i.e., by bringing the overall concentration level as far down as possible (at least a certain portion of the time determined by the $\alpha_i' s$), or by minimizing as in model (5.11) the function

$$\sum_{i=1}^{N} q_i \ e_i \ \theta \left[e_i^{-1}(v_i - \gamma_i) \right] \tag{5.16}$$

which penalizes the deviations from γ_i in a nonlinear manner, cf. Figure 3, or still to handle the minimization of the $(v_i, i = 1, \ldots, N)$ as a multiple objective optimization problem, each coordinate of v corresponding to an objective that we seek to minimize.

We shall here formulate our optimization problem in terms of the objective (5.14) but of course any of the other variants could or should also be considered. The optimization problem again involves probabilistic constraints but its structure now resembles much more the stochastic program with recourse (5.11) than the first model (5.8) involving probabilistic constraints. We obtain

$$
\begin{array}{lr}
\text{find } x \in R^n \text{ such that} & (5.17) \\
r_j^- \le x_j \le r_j^+, \qquad\qquad j=1,\dots,n & \\
\sum_{j=1}^n a_{ij}\, x_j \le b_i \qquad\qquad i=1,\dots,m & \\
\sum_{j=1}^n c_j(x_j) \le \beta & \\
\text{prob.}\left[\sum_{j=1}^n t_{ij}(w)x_j - h_i(w) - v_i < 0\right] \ge \alpha_i, \qquad i=1,\dots,N & \\
\text{and } z = \sum_{i=1}^N q_i\,[v_i - \gamma_i]_+ \text{ is minimized .} &
\end{array}
$$

At this point it may be worthwhile to observe that (5.14) is just a limit case of (5.16). Recall that the range over which $q_i\; e_i\; \theta(e_i^{-1}(\cdot - \gamma_i))$ is quadratic is $[0,e_i]$, cf. Figure 3. If we shrink this interval to 1 point, we are let with the piecewise linear function $q_i[\cdot - \gamma_i]_+$.

As for our earlier models, we should study the solution as a function of β, the available budget. However, solving (5.17) presents all the technical challenges mentioned in connection with the first model (5.8) involving probabilistic constraints, the presence of the $(v_i,\ i = 1,\dots,m_2)$ has in no way simplified the problem. We do not know of any direct method for solving (5.17). One possibility is to find an approximation of (5.17) that could be handled by available linear or nonlinear programming techniques. We return to this in the next section.

6. EMOM: AN EXPECTATION-VARIANCE MODEL

In the formulation of the eutrophication management optimization model (EMOM) the objective has always been chosen so as to measure in the most realistic fashion possible the deviations of the observed concentrations indicators from the water quality goals. This led us to the stochastic program with recourse model (5.11) with the associated solution procedure to be discussed in Section 7. Here we introduce a linear programming model (Somlyódy, 1983b) that is based on expectation-variance considerations (for the water quality indicators). The justification of the model relies on the validity of certain approximations and thus in some situations one should accept the so-generated optimal solution with some circumspection. However, as shall be argued, its solution always points in the right direction and is usually far superior to that obtained by solving a "deterministic" problem such as (5.12). In the Lake Balaton case study the results for both this expectation-variance model and the stochastic programming model (5.11) lead to remarkably similar investment decisions as shown by the analysis of the results in Section 8.

As a starting point for the construction of this model, let us consider the following objective function:

$$\sum_{i=1}^{N} q_i \, E \left\{ y_i(x,\cdot) - \gamma_i)_+^2 \right\} \tag{6.1}$$

where, as in Section 5 $y_i(x,w)$ is the water quality level characterized by the selected indicator in basin i given the investment program x and the environmental conditions w, γ_i the goal set for basin i and q_i a weighting factor associated to basin i. The objective being quadratic in the area of interest, and the distribution functions $G_i(x,\cdot)$ of the $y_i(x,\cdot)$ not being too far from normal, one should be able to recapture the essence of its effect on the decision process by considering just expectations and variances. This observation, and the "soft" character of the management problem which in any case means that there is a large degree of flexibility in the choice of the objective, suggest that we could substitute

$$\sum_{i=1}^{N} q_i \left[E\{y_i(x,\cdot) - \bar{y}_{oi}\} + \theta\ \sigma(y_i(x,\cdot) - \bar{y}_{oi})\right] \qquad (6.2)$$

for (6.1), where θ is a positive scalar (usually between 1 and 2.5), $\bar{y}_{oi} = E\{y_{oi}\}$ is the expected nominal state of basin i, and σ denotes standard deviation,

$$\sigma\left[y_i(x,\cdot) - \bar{y}_{oi}\right] = E\left\{\left[y_i(x,\cdot) - E\ \{y_i(x,\cdot)\}\right]^2\right\}^{1/2}. \qquad (6.3)$$

Since for each $i = 1, \cdots, N$, the y_i are affine (linear plus a constant term) with respect to x, the expression for

$$E\left\{y_i(x,\cdot) - \bar{y}_{oi}\right\} = \sum_{j=1}^{n} \mu_{ij} x_j + \mu_{io}$$

as a function of x is easy to obtain from equations (3.1) and (3.10). The μ_{ij} are the expectations of the coefficients of the x_j and the μ_{io} the expectation of the constant term. Unfortunately the same does not hold for the standard deviation $\sigma(y_i(x,\cdot) - \bar{y}_{oi})$. Equations (3.1) and (3.10) suggest that

$$\sigma(y_i(x,\cdot) - \bar{y}_{oi}) \sim (\sum_l \sigma_{il}^2\ x_l^2)^{1/2} \qquad (6.4)$$

where σ_{il} is the part of the standard deviation that can be influenced by the decision variable x_l; for example, the standard deviation of the tributary load $\underset{\sim}{L}_D$. Cross terms are for all practical purposes irrelevant in this situations since the total load in basin i is essentially the result of a sum of the loads generated by various sources that are independently controlled. This justifies using

$$\sum_{i=1}^{N} q_i \left| \left[\sum_{j=1}^{n} \mu_{ij} x_j \right] + \theta \left[\sum_{j=1}^{n} \sigma_{ij}^2 x_j^2 \right]^{1/2} \right| \tag{6.5}$$

instead of (6.2) as an objective for the optimization problem. This function is convex and differentiable on R_+^n except at $x = 0$, and conceivably one could use a nonlinear programming package to solve the optimization problem:

$$\begin{aligned}
&\text{find } x \in R^n \text{ such that} && (6.6)\\
&r_j^- \le x_j \le r_j^+ \qquad j=1,\dots,n\\
&\textstyle\sum_{j=1}^{n} a_{ij} x_j \le b_i \qquad i=1,\cdots,m\\
&\textstyle\sum_{j=1}^{n} c_j(x_j) \le \beta\\
&\text{and } z = \textstyle\sum_{i=1}^{N} q_i \left| \sum_{j=1}^{n} \mu_{ij}\, x_j + \theta \left[\sum_{j=1}^{n} \sigma_{ij}^2\, x_j^2 \right]^{1/2} \right| \text{ is minimized.}
\end{aligned}$$

Assuming that the cost functions have been linearized, i.e. with each c_j piecewise linear, the MINOS package would be an excellent choice since the solution is clearly bounded away from 0.

We can go one step further in simplifying the problem to be solved, namely by replacing the term.

$$\left[\sum_{j=1}^{n} \sigma_{ij}^2\, x_j^2 \right]^{1/2}$$

in the objective, by the linear (inner) approximation

$$\sum_{j=1}^{n} \sigma_{ij}\, x_j .$$

On each axis of R_+^n, no error is introduced by relying on this linear approximation, otherwise we are over-estimating the effect a certain combination of the x_j's will have on the variance of the concentration levels. Thus, at a given budget level we shall have a tendency to start projects that affect more strongly the variance if we use the linear approximation, and this is actually what we observed in practice (see Section 8). Assuming the cost functions c_j are piecewise linear, we have to solve the *linear* program:

$$
\begin{aligned}
&\text{find } x \in R^n \text{ such that} && && (6.7)\\
&r_j^- \le x_j \le r_j^+, && j=1,\cdots,n\\
&\textstyle\sum_{j=1}^n a_{ij} x_j \le b_i, && i=1,\cdots,m\\
&\textstyle\sum_{j=1}^n c_j(x_j) \le \beta\\
&\text{and } t = \textstyle\sum_{i=1}^N q_i \sum_{j=1}^n (\mu_{ij} + \theta\,\sigma_{ij}) x_j \text{ is minimized.}
\end{aligned}
$$

We refer to this problem as the *(linearized) expectation-variance model* (see also Somlyódy, 1983b; and Somlyódy and van Straten, 1985).

For the sake of illustration, let us consider the i-th basin of Figure 2 and suppose that there is no mass exchange with neighboring basins. Then from equations (3.1) and (3.7), recalling that $\underset{\sim}{L}_i = \underset{\sim}{L}_i^a / V_i$, we obtain

$$
E\left\{y_i(x,\cdot) - \bar{y}_{oi}\right\} = x_P\left\{d_{ii}\,\delta\left[E\{\underset{\sim}{L}_D\} - E\{\underset{\sim}{L}_T\}\right]\right\}
$$

$$
+x_D\left\{d_{ii}\left[(1-r_i)x_{SN}L_{SN} - E\{\underset{\sim}{L}_D\}\right]\right\} - x_{SN}(1-r_i)L_{SN} - x_S L_S.
$$

To obtain a linear form in the x_j, we proceed as indicated in Section 3, see equation (3.9). To derive the remaining term in the objective of (6.7) we only need to consider the controllable portion of the variance of $y_i(x,\cdot) - \bar{y}_{oi}$, we rewrite (3.1) as follows

$$
\underset{\sim}{y}_i(x,\cdot) - \bar{y}_{oi} = \underset{\sim}{w}_i - (d_{ii} + d_i \underset{\sim}{w}_i)\,\Delta\underset{\sim}{L}_i
$$

Let us write

$$
\Delta\underset{\sim}{y}_i = (d_{ii} + d_i \underset{\sim}{w}_i)\,\Delta\underset{\sim}{L}_i
$$

from which we obtain

$$\sigma^2(\Delta \underset{\sim}{y}_i) = d_{ii}^2 \; \sigma^2(\Delta \underset{\sim}{L}_i) + d_i^2 \; \sigma_{\underset{\sim}{w}_i} \left[\sigma^2(\Delta \underset{\sim}{L}_i) + E^2\{\Delta \underset{\sim}{L}_i\} \right] \tag{6.8}$$

Now from equations (3.5), (3.6) and (3.7) we have that

$$\sigma^2(\Delta \underset{\sim}{L}_i) = \frac{1}{V_i} \left[\delta^2 \; x_P \; \sigma^2(\underset{\sim}{L}_T) + (x_D - \delta x_P)^2 \; \sigma(\underset{\sim}{L}_D) \right]. \tag{6.9}$$

This would lead to an expression for $\sigma(\Delta \underset{\sim}{y}_i)$ that would be nonlinear in the x variables. To avoid the nonlinearities we specify $\sigma_a(\Delta \underset{\sim}{y}_i)$ and $\sigma_a(\Delta \underset{\sim}{L}_i)$ as the linear combination of the additive terms in (6.8) and (6.9)

$$\sigma_a(\Delta \underset{\sim}{y}_i) := d_{ii} \, \sigma_a(\Delta \underset{\sim}{L}_i) + d_i \, \sigma_{\underset{\sim}{w}_i} \left[\sigma_a(\Delta \underset{\sim}{L}_i) + E\{\Delta \underset{\sim}{L}_i\} \right] \tag{6.10}$$

and

$$\sigma_a(\Delta \underset{\sim}{L}_i) := \frac{1}{V_i} \left[\delta x_P \; \sigma(\underset{\sim}{L}_T) + (x_D - \delta x_P) \; \sigma(\underset{\sim}{L}_D) \right]. \tag{6.11}$$

Note that in equations (6.8) and (6.9) all the coefficients are positive and the behavior of the "new" σ_a is similar to the standard deviations as defined through (6.8) and (6.9). Substituting (6.11) in (6.10) yields.

$$\begin{aligned} \sigma_{ai} = V_i^{-1} \Big\{ & x_P \left[(d_{ii} + d_i \; \sigma_{\underset{\sim}{w}_i}) \, \delta \left[\sigma(\underset{\sim}{L}_T) - \sigma(\underset{\sim}{L}_D) \right] + d_i \, \sigma_{w_i} \delta \left[E\{\underset{\sim}{L}_T\} - E\{\underset{\sim}{L}_D\} \right] \right] \\ & + x_D \left[\left[d_{ii} + d_i \, \sigma_{\underset{\sim}{w}_i} \sigma(\underset{\sim}{L}_D) \right] + d_i \, \sigma_{\underset{\sim}{w}_i} \left[E\{\underset{\sim}{L}_D\} - (1 - r_i) x_{SN} L_{SN} \right] \right] \\ & + x_{SN} \; d_i \; \sigma_{\underset{\sim}{w}_i} L_{SN} + x_s \; d_i \; \sigma_{\underset{\sim}{w}_i} L_S \Big\}. \end{aligned}$$

Collecting terms we obtain the coefficients σ_{ij} that appear in the objective of the linear program (6.7). (A more detailed, but similar, derivation also yields the expression for the standard derivation when there is mass exchange between neighboring basins).

The arguments that we have used to justify the expectation-variance model are mostly of a heuristic nature, in that they rely on a good understanding of the problem at hand and "engineering" intuition. In the formulation of the EMOM models, in Section 5, the objective has usually been formulated in terms of finding control measures such that the observed concentration levels (water quality indicators) are not too far from pre-set goals (given trophic states). If by "not too far" we mean that

$$\left\{E\left[\underset{\sim}{y}_i(x,\cdot)\right]+\sigma\left[\underset{\sim}{y}_i(x,\cdot)\right]\right\}-\gamma_i \tag{6.12}$$

should be as small as possible, we could also reformulate the problem, in terms of the nominal concentration levels. Indeed, with

$$\Delta\underset{\sim}{y}_i=\underset{\sim}{y}_i-\underset{\sim}{y}_{oi}$$

where as in Section 3, $\underset{\sim}{y}_{oi}$ denotes the nominal state in basin i. Then instead of minimizing (6.12), we could maximize

$$E\left[\Delta\underset{\sim}{y}_i(x,\cdot)\right]-\sigma\left[\Delta\underset{\sim}{y}_i(x,\cdot)\right] \tag{6.13}$$

and this should give about the same results. This is actually the motivation behind the formulation of (6.7), see Figure 4.

There is however another approach that does not rely so extensively on heuristic considerations, which leads us to the model (6.6), i.e., the nonlinear version of the expectation-variance model. The fourth model, described in Section 5, which integrates both reliability considerations and penalties for fixing the reliability levels, led us to the nonlinear program find $x \in R^n$ such that (5.17)

$$\begin{aligned}
& r_j^- \le x_j \le r_j^+, && j=1,\cdots,n, \\
& \textstyle\sum_{j=1}^n a_{ij}\, x_j \le b_i && i=1,\cdots,m \\
& \textstyle\sum_{j=1}^n c_j(x_j) \le \beta && \\
& \text{prob.} \left[\textstyle\sum_{j=1}^n t_{ij}(w)x_j - h_i(w) - y_i < 0\right] \ge \alpha_i, && i=1,\cdots,N \\
& \text{and } z = \textstyle\sum_{i=1}^N q_i\, [y_i - \gamma_i]_+ \text{ is minimized .} &&
\end{aligned}$$

Because these probabilistic constraints are very difficult to handle, we may consider finding an approximate solution by replacing the probabilistic constraints by

$$(1-\alpha_i)^{-2} \left[\sum_{j=0}^n \sum_{k=0}^n \sigma_{ijk}\, x_j\, x_k\right]^{1/2} + \sum_{j=0}^n \mu_{ij}\, x_j \le y_i \tag{6.14}$$

$$x_0 = 1$$

where

$$t_{oi}(\cdot) = -h_i(\cdot),$$

and for $j=0,..,n$ and $k=0,...,n$

$$\mu_{ij} := E\left\{t_{ij}(w)\right\},$$

$$\sigma_{ijk} := \mathrm{cov}\left[t_{ij}(\cdot),\ t_{ik}(\cdot)\right].$$

If the random variables $\left[t_{ij}(\cdot),\ j=0,...,n\right]$ are jointly normal, then the restrictions generated by the deterministic constraints (6.14) are exactly the same as those imposed by the probabilistic constraints, but in general they are more restrictive, cf. Propositions 1.25 and 1.26 in Wets, 1983a. Without going into the details, we can see that (6.14) is obtained by applying Chebyshev's inequality and this, in general, determines an upper bound for the probabilistic event

$$\left\{w \mid \sum_{j=0}^n t_{ij}(w)x_j \ge 0\right\}.$$

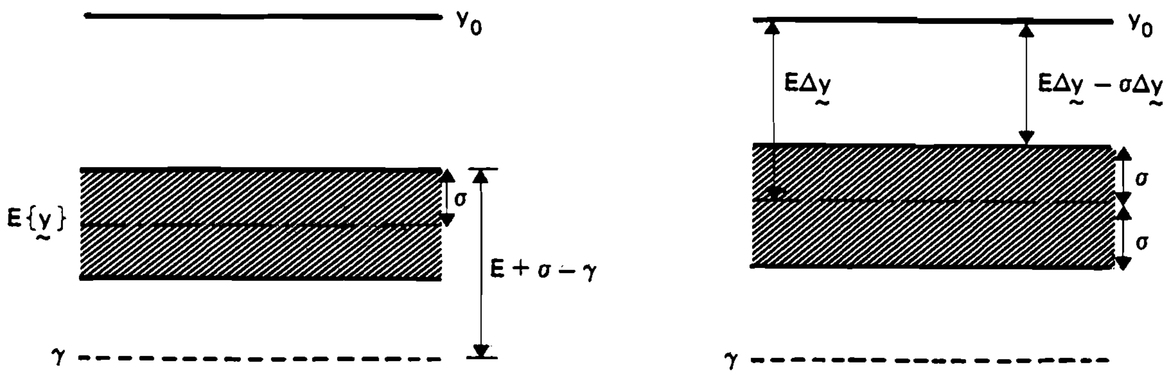


FIGURE 4. Objective of expectation-variance model.

So if we can justify a near normal behavior for the random variable (for fixed x)

$$\sum_{j=1}^{n} t_{ij}(w)x_j - h_i(w) =: y_i(x,w),$$

we can use the constraints (6.14) instead of the probabilistic constraints to obtain an approximate solution of (5.17). Note that in this setting, "near normality" of the $\underset{\sim}{y}_i(x,\cdot)$ is a much more natural, and weaker, assumption than normality of the $\underset{\sim}{t}_{ij}$. Assuming that we proceed in this fashion, we obtain the nonlinear program:

$$\text{find } x \in R^n \text{ such that} \tag{6.15}$$
$$r_j^- \le x_j \le r_j^+, \qquad j=0, \cdots, n$$
$$\sum_{j=1}^{r} a_{ij}\, x_j \le b_i, \qquad i=1, \cdots, m$$
$$\sum_{j=1}^{n} c_j(x_j) \le \beta$$

and

$$z = \sum_{i=1}^{N} q_i \Big| \sum_{j=0}^{n} \mu_{ij} x_j + (1-\alpha_i)^{-2} \sum_{j=0}^{n} \sum_{k=0}^{n} \sigma_{ijk} x_j x_k)^{1/2} - \gamma_i \Big|_+ \text{ is minimized}$$

where $r_0^- = r_0^+ = 1$. We have eliminated the variables $(y_i, i=1, \cdots, N)$ from the formulation of the problem by using the fact that the optimal y_i^* can always be chosen so that (6.14) is satisfied with equality. Moreover, if the desired concentration levels γ_i are low enough, then we know that the optimal solution will always have $y_i^* > \gamma_i$ and thus we can rewrite (6.15) as follows:

$$\text{find } x \in R^n \text{ such that} \tag{6.16}$$
$$r_j^- \le x_j \le r_j^+, \qquad j=0,\dots,n$$
$$\sum_{j=1}^{n} a_{ij}\, x_j \le b_i, \qquad i=1,\dots,m$$
$$\sum_{j=1}^{n} c_j(x_j) \le \beta$$

and

$$z = \sum_{i=1}^{N} q_i \Big| \sum_{j=0}^{n} \mu_{ij}\, x_j + (1-\alpha_i)^{-2} (\sum_{j=0}^{n} \sum_{k=0}^{m} \sigma_{ijk}\, x_j\, x_k)^{1/2} - \gamma_i \Big| \text{ is minimized.}$$

The objective of this optimization problem is sublinear, i.e., convex and positively homogeneous. Assuming that the cost functions c_j are linear, or more realistically have been linearized, see Section 4.3, we are thus confronted with a nearly linear program that we could solve by specially designed subroutines (nondifferentiability at 0), or by a linearization scheme that would allow us to use linear programming packages. Now observe that the nonlinear program (6.16) is exactly of the same type as (6.6) if we make the following adjustments:

(i) in the objective of (6.16) replace the covariance term $\sum_{j=0}^{n} \sum_{k=0}^{n} \sigma_{ijk} x_k x_j$ by the sum of the variances $\sum_{j=1}^{n} \sigma_{ijj} x_j^2$;

and

(ii) if for all $i=1,\ldots,N$, the α_i are the same set $\theta = (1-\alpha_i)^{-1}$, otherwise we replace θ by $\theta_i = (1-\alpha_i)^{-2}$ in (6.6).

To justify (i), we appeal to (6.4).

In the derivation that led us from (5.17) to (6.16), we stressed the fact that the solution of (6.16) and thus equivalently of (6.6), would be feasible for the original program (5.17), and that in fact it would more than meet the probabilistic constraints specified in (5.17). The further linearization of the objective bringing us from (6.6) to (6.7) overstresses (possibly only slightly) the role that the variance will play in meeting the prescribed reliability levels. In terms of model (5.17), we can thus view the solution of (6.7) as a "conservative" solution that overestimates the importance to be given to the stochastic aspects of the problem. In that sense, the solution of (6.7), especially in comparison to that of the deterministic problem (5.12), always indicates how we should adjust the decisions so as to take into account the stochastic features of the problem.

In our analysis (see Section 8), we have used the linear programming version (6.7) of this expectation-variance model; the wide availability of reliable linear programming packages makes it easy to implement, and thus an attractive approach, provided one keeps in mind the reservations expressed earlier.

7. SOLVING EMOM: STOCHASTIC VERSION

We outline here a method for solving the "full" stochastic version of the eutrophication management optimization model (5.11). We recognize (5.11) as a stochastic program with simple recourse, with technology matrix T and right hand sides h stochastic. When only h is stochastic and the objective function (recourse cost function) is piecewise linear, efficient procedures are available for stochastic programs with simple recourse, cf. Wets, 1983b (implemented by Kahlberg and Kusy, 1976); for a report on numerical experimentation with this code, see Kusy and Ziemba, 1983; and Nazareth, 1984; for a different algorithmic approach and a progress report on its implementation. But here we have to deal as with nonlinear (convex) recourse costs. In order to deal with this class of problems, a new procedure was developed which exploits the properties of a dual associated to problem (5.11).

Replacing the budget constraint by the equivalent linear system, cf (4.5) and the related discussion, substituting $\hat{x}_j + r_j^-$ for x_j, defining

$$\hat{h}(\xi) = h(\xi) - T(\xi)\, r_j^-$$
$$\hat{r}_j = r_j^+ - r_j^-$$

and dropping the $\hat{}$, we obtain a version of (5.11) that fits into the following class of problems:

$$\begin{aligned}
&\text{find } x \in R^n \text{ such that} \\
&0 \le x_j \le r_j, && j = 1, \cdots, n \\
&\textstyle\sum_{j=1}^n a_{ij}\, x_j \le b_i, && i = 1, \ldots, m_1 \qquad (7.1)\\
&\textstyle\sum_{j=1}^n t_{ij}(w) x_j - v_i(w) = h_i(w) && i = 1, \cdots, m_2 \\
&\text{and } z = \textstyle\sum_{j=1}^n \left[c_j x_j + \frac{d_j}{2r_j} x_j^2\right] + E\left\{\sum_{i=1}^{m_2} q_i\, e_i\, \theta\, \left[e_i^{-1} v_i(w)\right]\right\} \text{ is minimized}
\end{aligned}$$

to which one refers as *quadratic stochastic programs with simple recourse*. Here $m_2 = N$, $c_j = d_j = 0$, and the system $(\sum_{j=1}^{n} a_{ij} \le b_i,\ i=1,\dots,m_1)$ has been expanded $(m_1 > m)$ to include the budgetary constraints. The θ function is defined by (5.9). For problems of this type, in fact with this application in mind, an algorithm is developed in Rockafellar and Wets, 1983, which relies on the properties of a dual problem which can be associated to (7.1). In particular it is shown that the following problem:

$$
\begin{aligned}
&\text{find } y \in R_+^{m} \text{ and } z(\cdot)\colon\ \Omega \to R^{m_2} \text{ measurable such that} && (7.2)\\
&0 \le z_i(w) \le q_i, && i=1,\cdots,m_2\\
&u_j = c_j - \textstyle\sum_{i=1}^{m_1} a_{ij}\, y_i - E\left\{\sum_{i=1}^{m_2} z_i(w)\, t_{ij}(w)\right\}, && j=1,\cdots,n\\
&\text{and } \textstyle\sum_{i=1}^{m_1} y_i\, b_i - \sum_{i=1}^{m_2} E\left\{h_i(w) z_i(w) + \dfrac{e_i}{2q_i}\, z_i^2(w)\right\}\\
&-\textstyle\sum_{j=1}^{n} r_j\, d_j\, \theta(d_j^{-1} u_j) \text{ is maximized},
\end{aligned}
$$

is dual to the original problem, provided that for $i=1,\dots,m_2$, the e_i and q_i are positive (and that is the case here) and for $j = 1,\dots,n$, the $d_j > 0$ which is not the case. This can be taken care of in two different manners.

First, one should observe that the Duality Theorem proved in Rockafellar and Wets, 1983, remains valid for $d_j = 0$ provided that the function

$$u_j \;\to\; r_j\, d_j\, \theta(d_j^{-1} u_j)$$

be replaced by

$$u_j \to r_j\, \theta^{*}(u_j)$$

when $d_j = 0$, where

$$\theta^{*}(\tau) = \begin{cases} 0 & \text{if } \tau \le 0 \\ \tau & \text{if } \tau \ge 0 \end{cases}$$

(the function $r_j\, \theta^{\infty}(\bullet)$ is the limit of the collection $r_j\, d_j\, \theta(d_j^{-1} \bullet)$ as d_j tends to zero). The same proof applies. An algorithm for solving this problem — i.e., (6.1) but with $r_j\, d_j\, \theta(d_j^{-1} w_j)$, replaced by $r_j\, \theta^{\infty}(w_j)$ in the objective — could be constructed along the same lines as those of the method to be described here below. The basic difference being that θ^{∞} is not differentiable and that needs to be handled appropriately.

Second, to introduce in the objective of (5.11) an artificial perturbation of the objective that would result in a problem of type (7.1) with the prescribed properties. In the algorithm to be described such quadratic terms are added in a most natural fashion. The overall strategy is that of the proximal point algorithm, studied by Rockafellar, 1976. We generate a sequence of feasible solutions $\{x^{\nu},\ \nu = 1, \cdots\}$ of (6.1) with each x^{ν} the (unique) optimal solution of (7.1) where c_j and d_j are replaced by

$$c_j^{\nu} = c_j + \rho_{\nu}\, x_j^{\nu-1}$$

and

$$d_j^{\nu} = d_j + \rho_{\nu}\, r_j\ ;$$

the scalars $\rho_{\nu} \downarrow 0$ with $\rho_{\nu} > 0$. Since for each ν, $d_j^{\nu} > 0$ (the $r_j > 0$ and $\rho_{\nu} > 0$), we can apply the Duality Theorem of Rockafellar and Wets, 1983 (as stated), the dual being exactly (7.2) with c_j and d_j replaced by c_j^{ν} and d_j^{ν} for $j = 1,\ldots,n$. This dual problem is then solved and the (optimal) multipliers associated to the equations

$$u_j = c_j^{\nu} - \textstyle\sum_{i=1}^{m_1} a_{ij}\, y_i - \sum_{i=1}^{m_2} E\left\{z_i(w)\, t_{ij}(w)\right\} \tag{7.3}$$

give us x^{ν} the optimal solution of problem (7.1) with the adjusted coefficients c_j^{ν} and d_j^{ν}. We then generate $c_j^{\nu+1}$ and $d_j^{\nu+1}$ and again the resulting dual is solved. We repeat this until an error bound (involving the evaluation of the objective function at x^ν) is below a preset level. A detailed description of the method (in a somewhat more general setting), its convergence and various error bounds can be found in Rockafellar and Wets, 1985.

Every iteration of the algorithm, from x^ν to $x^{\nu+1}$, involves the solution of a problem of type (7.2), which can almost be viewed as a nonstochastic problem except for the simple constraints

$$0 \leq z_i(\cdot) \leq q_i, \qquad i = 1,\dots,m_2 .$$

The method that we use relies on a finite element representation of the $z_i(\cdot)$ functions, that is chosen so that the above constraints are automatically satisfied. Let $\zeta_{il}(\cdot)$, $l = 1,\dots,L$ be a finite collection of functions such that for all l,

$$0 \leq \zeta_{il}(\cdot) \leq q_i, \qquad i = 1,\dots,m . \tag{7.4}$$

If this collection of functions is rich enough, or if chosen so that the optimal $z_i^*(\cdot)$ that is part of an optimal solution of (7.2), lies close to their linear span, we write

$$z_i(\cdot) = \sum_{l=1}^{L} \lambda_{il} \; \zeta_{il}(\cdot) \tag{7.5}$$

where the λ_{il} are weights associated to the functions $\zeta_{il}(\cdot)$ such that

$$\sum_{l=1}^{L} \lambda_{il} = 1, \;\; \lambda_{il} \geq \text{ for all } l .$$

The generated functions $z_i(\cdot)$ are always between 0 and q_i and the constraints (7.4) can be ignored. With the substitution (7.5), the dual problem (7.2) becomes a *finite dimensional* quadratic program find $y \in R_+^{m_1}$ and for $i=1,\ldots,m_2$, $l=1,\ldots,L$, $\lambda_{il} \geq$ such that (7.6)

$$u_j = c_j - \sum_{i=1}^{m} a_{ij} y_i - \sum_{l=1}^{L} \lambda_{il} \bar{t}_{ijl}, \qquad j=1,\ldots,n \text{ and}$$

and

$$\sum_{i=1}^{m_1} b_i y_i + \sum_{i=1}^{m_2} \sum_{l=1}^{L} \lambda_{il} \bar{h}_{il} + \tfrac{1}{2} \sum_{i=1}^{m_2} \sum_{l=1}^{L} \sum_{l'=1}^{L} \lambda_{il} \lambda_{il'} \bar{e}_{ill'}$$

$$+ \sum_{j=1}^{n} r_j\, d_j\, \theta(d_j^{-1} u_j) \text{ is minimized}$$

where

$$\bar{t}_{ijl} = E\left\{\zeta_{il}(w)\, t_{ij}(w)\right\},$$

$$\bar{h}_{il} = E\left\{\zeta_{il}(w)\, h_i(w)\right\},$$

$$\bar{e}_{ill'} = E\left\{e_i\, q_i^{-1}\, \zeta_{il}(w)\, \zeta_{il'}(w)\right\}.$$

The only question is to know if the collection $\{\zeta_{il}(\cdot),\ l=1,\ldots,L\}$ is rich enough to yield a solution of (7.1) by solving (7.6). This is resolved in the following fashion. Let $(\hat{y}, \hat{z}(\cdot))$ with $\hat{z}(\cdot) = \sum_{l=1}^{L} \hat{\lambda}_{il}\, \zeta_{il}(\cdot)$ be the solution generated by solving (7.6). This is a feasible solution of (7.2), but not necessarily an optimal solution. We then solve the following problems: for $i=1,\ldots,m_2$

$$\text{find } \zeta_i^{L+1} \in \operatorname{argmin}\left[\Phi(\hat{y}, \zeta) \mid 0 \leq \zeta \leq q_i\right] \tag{7.7}$$

where $\Phi(\hat{y}, \cdot)$ is the objective function of (7.2) obtained by setting $y = \hat{y}$. This is a very simple optimization problem, in fact its solution is given by the formula: for $i = 1,\ldots,m_2$

$$\zeta_{i,L+1}(w) = q_i\, \theta'\left[e_i^{-1}\left[\sum_{j=1}^{n} t_{ij}(w)\hat{x}_j - h_i(w)\right]\right] \tag{7.8}$$

where $\hat{x}$ are the multipliers associated to the equation (7.3), for the optimal solution $(\hat{y}, \hat{\lambda})$ of (7.6) and θ' is the (particularly simple) derivative of θ. If we add these functions to those used to represent $z(\cdot)$, by solving the augmented problem (7.6) with $L + 1$, instead of L, we are guaranteed to obtain an improved solution of (7.2) unless $(\hat{y}, \hat{z}(\cdot))$ is already optimal for (7.2). For further details, including an analysis of the convergence rate, consult Rockafellar and Wets, 1985.

An experimental version of this algorithm was implemented at IIASA by A. King (and is available through IIASA as part of a collection of codes for solving stochastic programs), see King, 1985. We start the procedure by solving the deterministic problem (5.12) with $\hat{h}_i = \bar{h}_i$ and $\hat{t}_{ij} = \bar{t}_{ij}$. This gives us a vector x^{ν} with $\nu = 1$ that is used to define the vectors c^{ν} and d^{ν}, and for $i = 1, \cdots, m_2$,

$$\zeta_{i1}(w) = q_i\, \theta' \left\{ e_i^{-1} \left[\sum_{j=1}^{n} t_{ij}(w) x_j^{\nu} - h_i(w) \right] \right\}.$$

In addition to this function ζ^1, we also include in our starting collection of finite elements $\zeta^{-1} \equiv 0$ and $\zeta^0 \equiv q$. The optimal solution of the quadratic program (7.6) and the associated multipliers are then used to generate through (7.8) a new function ζ^l (leading up to a larger version of (7.6)) or the solution is recognized as optimal (or more exactly ε-optimal). In this latter case, we use the (optimal) value of the multipliers $x^{\nu+1}$ associated to the equations that define u_j, $j = 1, \cdots, n$ as the starting point of a new major cycle, redefining $c^{\nu+1}$ and $d^{\nu+1}$ and proceeding from there as indicated above. The algorithm is stopped when the gap, between the objective of the original problem (7.1) evalued at $x = x^{\nu+1}$ and the objective of its dual evaluated at $y^{\nu+1}$ and $z^{\nu+1}(\cdot)$ obtained from the (last) solution of (7.6), is sufficiently small.

The distribution of the random elements $t_{ij}(\cdot)$ and $h_i(\cdot)$ is replaced by a discrete distribution obtained from the output of NLMP2 and LEMP by sampling. For a number of reasons including numerical stability considerations and the speed-up of the computation, it is recommended to start with a relatively small sample increasing its size only for verification purposes. We have observed that a relatively small sample, about 50, will give surprisingly good results.

To solve the quadratic program (6.6) we used MINOS (Murtagh and Saunders, 1983). Much could be done to improve the performance of that part of the algorithm, a specially designed quadratic programming subroutine should be able to exploit not only the structure of this quadratic program but also the fact that when introducing an additional finite element we can use the previous solution to simplify the search for a new basis, etc.

8. APPLICATION TO LAKE BALATON

8.1. Background for Lake Balaton

8.1.1. Description of the lake, its watershed and possible control measures

Lake Balaton (Figure 5), one of the largest shallow lakes of the world, which is also the center of the most important recreational area in Hungary, has recently exhibited the unfavorable signs of artificial eutrophication. An impression of the major features of the lake-region system, of the main processes and activities, about the underlying research, data availability and control alternatives can be gained from Figure 5 and Table 1 (for details, see Somlyódy et al, 1983; and Somlyódy and van Straten, 1985). Four basins of different water quality can be distinguished in the lake (Figure 5) determined by the increasing volumetric nutrient load from east to west* (the biologically available load, BAP, is about ten times higher in Basin I than in Basin IV, Table 1, line 7). The latter is associated to the asymetric geometry of the system, namely the smallest western basin drains half of the total watershed, while only 5% of the catchment area belongs to the larger basin (Table 1, lines 2 and 4).

Based on observations for the period 1971-1982 the average deterioration of water quality of the entire lake is about 20% (in terms of Chl-a). According to the OECD classification, the western part of the lake is in a hypertrophic, while the eastern portion of it is in an eutrophic stage (Table 1, line 9).

* The absolute loads are roughly equal for the four basins

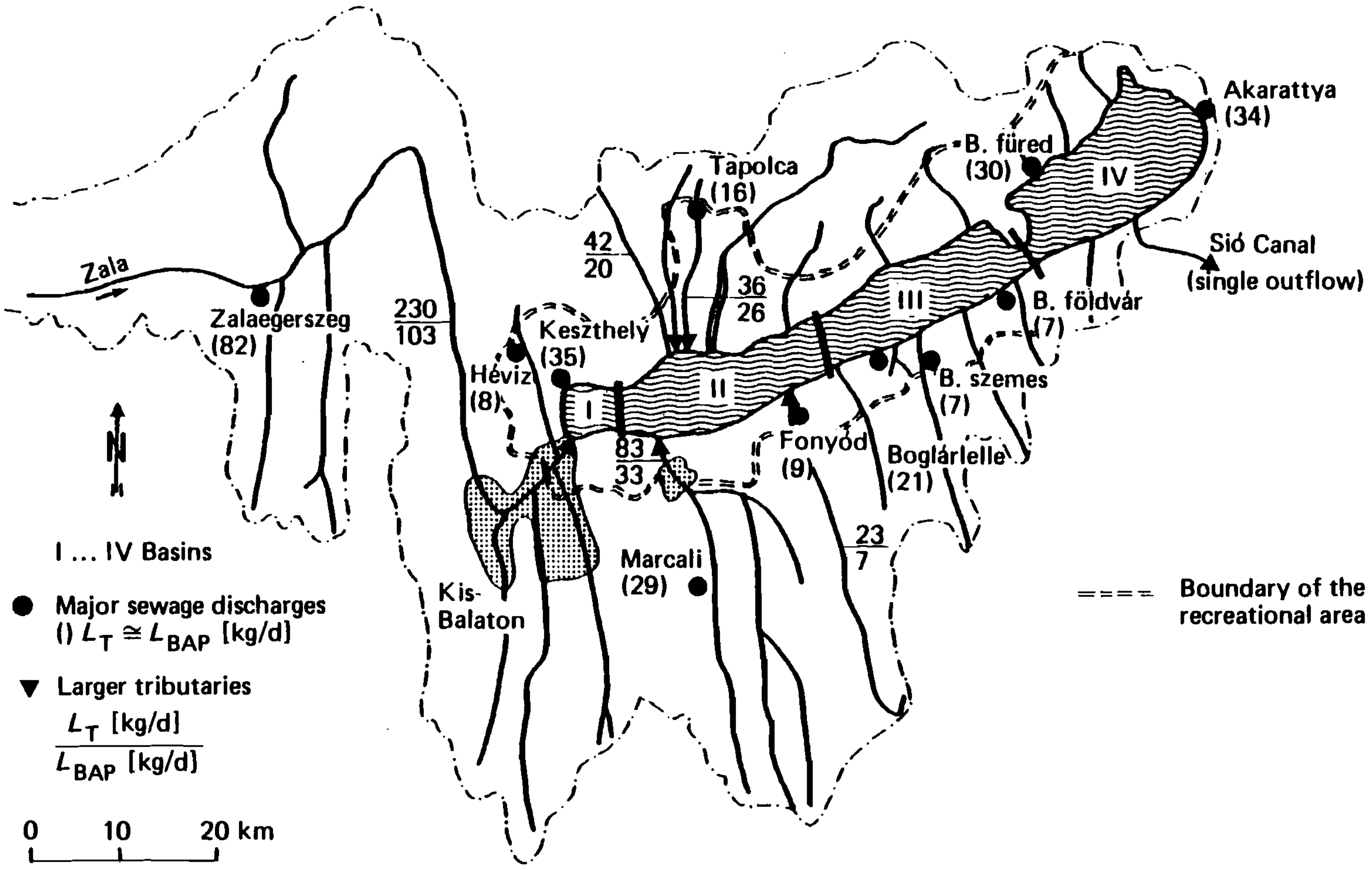


FIGURE 5. Major nutrient sources and control options.

TABLE 1. Major features of the lake and its watershed.

1.	Basin	I	II	III	IV	Whole lake
2.	Watershed area [km^2]	2750	1647	534	249	5180
3.	Use of the watershed	agriculture and intensive tourism (main season: July and August)				
4.	Lake surface area [km^2	38	144	186	228	536
5.	Volume [10^6m^3]	82	413	600	802	1907
6.	Depth [m]	2.3	2.9	3.2	3.7	3.2
7.	BAP load [mg/m^{3}d]	1.70	0.30	0.15	0.14	0.25
8.	Climatic influences	no stratification; large fluctuation in temperature (up to 25-28°C); 2-month ice cover; strong wind action				
9.	Water quality (Chl–a)$_{max}$ [mg/m^3]	(a) 75	38	28	20	late seventies
		(b) 150	90	60	35	1982
		hyper-eutrophic state; *P* limitation till the end of the seventies; large year-to-year fluctuation in Chl-a depending on meteorology and hydrology; 20% per year increase in Chl-a during 1971-1982; marked longitudinal gradient				
10.	Sediment	internal load nowadays is roughly equal to the external BAP load				
11.	Data	long hydrological and wheater records; regular water quality and load survey since 1971 and 1975, respectively				
12.	Research	increasing activity in Hungary in various institutes during the past 30 years; joint study of IIASA, the Hungarian Academy of Sciences and the Hungarian National Water Authority, 1978-1982, see Somlyódy et al. (1983).				
13.	Models developed	various alternative models for the components indicated in Figure 1, see Somlyódy et al. (1983)				
14.	Methodologies	0D and PDE models, regression analysis, Kalman filtering, time series analysis, Monte Carlo simulations, uncertainty analyses, optimization techniques				
15.	Measures of short-term control	P precipitation on existing treatment plants; pre-reservoirs				
16.	Policy making	Government decision in 1983: P control is being under realization				

The lakes' total P, L_T is in an average $315t/yr$ (the BAP load is $170t/yr$), but depending on the hydrologic regime it can reach $550t/yr$. 53% of L_T is carried by tributaries (30% of which is of sewage origin -- indirect load, see e.g., the largest city of the region, Zalaegerszeg in Figure 5), 17% is associated to direct sewage discharges (the recipient is the lake). Atmospheric pollution if responsible for 8% of the lake's TP load and the rest is formed by direct runoff (urban and agricultural). Tributary load increases from east to west, while the change in the direct sewage load goes in the opposite direction. The sewage contribution (direct and indirect loads) to L_T is 30%, while about 52% to L_{BAP} (the load of agricultural origin can be estimated as 47 and 33%, respectively) suggesting the importance of sewage load from the viewpoint of the short term eutrophication control. Figure 5 indicates also the loads of sewage discharges and tributaries which were involved in the management optimization model. These cover about 85% of the nutrient load* which we consider controllable on the short term (e.g. atmospheric pollution and direct runoff are excluded).

Control alternatives are sewage treatment (upgrading and P precipitation) and the establishment of pre-reservoirs as indicated in Figure 5 (see e.g. the Kis-Balaton reservoir system planned for a surface area of about 75 km^2). Besides Hungarian research activities, the problem of Lake Balaton was studied in the framework of a four-year cooperative research project on Lake Balaton involving the International Institute for Applied Systems Analysis, IIASA (Laxenburg, Austria), the Hungarian Academy of Sciences, and the Hungarian National Water Authority (Somlyódy, 1982 and

* The rest represented by several small creeks and sewage outlets were neglected for the sake of simplicity.

1983a; Somlyódy et al, 1983). The development of the management model to be discussed here formed a part of the Case Study. The results achieved were then utilized in 1982* in the policy making procedure associated with the Lake Balaton water quality problem which was completed by a governmental decision in 1983 (Láng, 1985).

8.1.2. Specification of elements of EMOM for Lake Balaton

(a) Lake eutrophication model, LEMP

Starting from a four-compartment, four-box dynamic phosphorus cycle model (van Straten, 1981), LEM, the aggregated model LEMP of a structure defined by Equation (3.1) was derived by a systematic analysis as described in Section 2. As seen from Figure 6, illustrating the deterministic version of LEMP, the $Y_i = (\mathrm{Chl-a})_{\max,i}$ indicators ($1 \leq i \leq 4$) are really linearly dependent on the sum of the external and internal loads (see also line 10 in Table 1). The matrix D takes a specific bloc-diagonal structure in this case, since neighboring basins are only related in a unidirectional way from west to east, which indicates that any management actions performed on the eastern subwatersheds have no effect on water quality of the western basins (d_{ii} range between 20 and 35, while elements related to interbasin exchange are smaller by an order of magnitude).

* At that time only the results of the expectation-variance model were available, the development of which was the fastest.

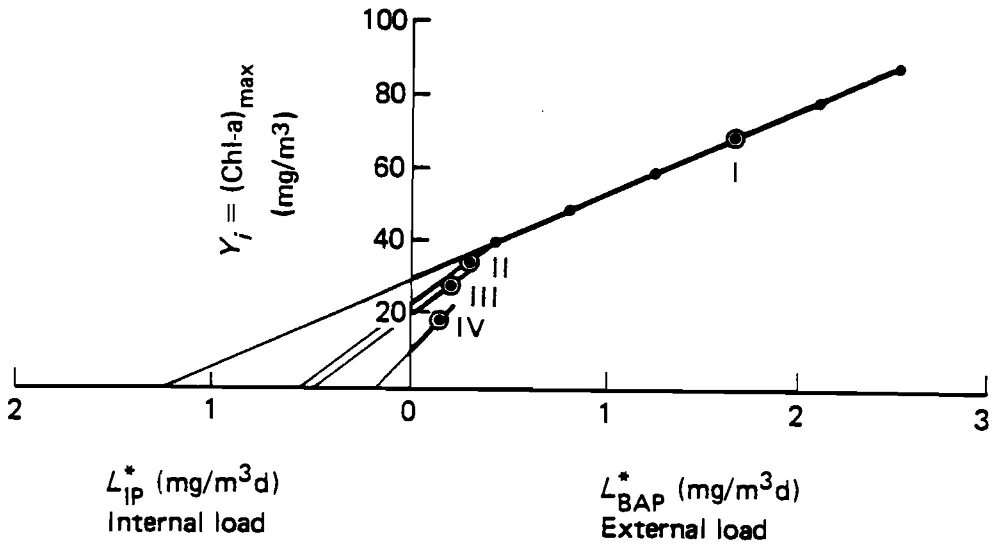


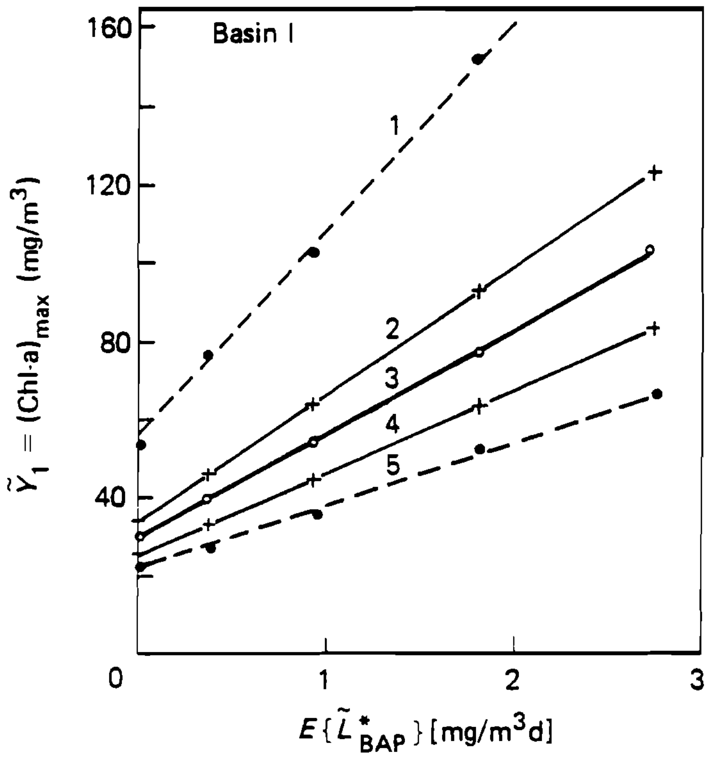
FIGURE 6. Aggregated lake eutrophication model: deterministic version of LEMP.

The stochastic version of LEMP was derived from a Monte-Carlo type usage of LEM under synthetic time series generated by models METG and NLMP1 (Figure 1). The analysis has shown that in the non-controlled case $(\text{Chl-a})_{max}$ for Basin I can range between 30 and 90 mg/m^3 (±40% around the mean) depending solely on meteorological factors; a strikingly wide domain. Such fluctuations can mask the effect of considerable load reductions. The sensitivity of the other basins in the lake is smaller: the coefficients of variation of $(\text{Chl-a})_{max}$ for Basin II ... IV are 10, 6 and 5% (it is 13% for Basin I).

If stochastic variations in loads are also taken into account, the coefficient of variation for Basin I goes to 20% and the extreme value of $(\text{Chl-a})_{max}$ can reach 150 mg/m^3*. Upper extremes for subsequent basins are 60, 35 and 25 mg/m^3, respectively. As apparent from Figure 7,

* A value observed in 1982 (Table 1, line 9) when the prognosis was already available.

referring to Basin I under pre-reservoir control, linearity is preserved as before, and not only for the mean, but also for statistical properties of the typically skewed distributions (standard deviations and extreme values).



3: mean; 2 and 4: ± standard deviation; 1 and 5: extremes; E is the operator of expectation.

FIGURE 7. Aggregated lake eutrophication model: stochastic version.

The analysis outlined here led to the specification of Equation (3.1) for Lake Balaton; for details the reader is referred to Somlyódy (1983b) and Somlyódy and van Straten (1985).

(b) Nutrient load model, NLMP2

The nutrient load model for Lake Balaton can be derived on the basis of Figure 5 from relation (3.10). The tributary loads, $\underset{\sim}{L}_T$ and $\underset{\sim}{L}_D$ are computed from regression models (Somlyódy and van Straten, 1985)

$$\underset{\sim}{L} = (L_0 + a_1 \underset{\sim}{Q} + \underset{\sim}{L}_\rho)(\xi^- + \underset{\sim}{\xi}) \tag{8.1}$$

where $\underset{\sim}{Q}$ is the stream flow rate, $\underset{\sim}{L}_\rho$ is the residual, and the variable $\underset{\sim}{\xi}$ accounts for the influence of infrequent sampling (ξ^- is the lower bound). The most detailed data set including 25 years long continuous records for Q and 5 years long daily observations for the loads was available for the Zala River* (Figure 5) draining half of the watershed and representing practically the total load of Basin I. For the Zala River $\underset{\sim}{L}_\rho$ was found to have a normal distribution, while $\underset{\sim}{Q}$ was approached by a lognormal distribution. The time step of the original model employed in NLMP 1 (Figure 1) was a month and that of the aggregated version used in NLMP2 a year. The loads of other tributaries were established on the basis of much more scarce observations. For modeling the uncertainty component of $\underset{\sim}{\xi}$, first a Monte-Carlo analysis was performed on the Zala River data by assuming various sampling strategies. Subsequently, the conclusions were extended to the other rivers and the parameters of the (assumed) gamma distributions of $\underset{\sim}{\xi}$ were estimated.

It should be noted that the nonlinearity related to the product terms ($x_D \cdot x_{SN}$) in NLMP2 was handled in both ways, see (3.8) and (3.9), indicated in Section 3 in the expectation-variance model, while through the constraint equation (3.9) in other management models discussed in Section 5.

---

* Its annual load estimated from daily data can be considered accurate.

(c) Control variables and cost functions

All the optimization models implemented use real-valued control variables. Integer $\{0,1\}$ variables for the two reservoir systems (see Figure 5) were also used by simply fixing the variable values at 0 or 1 as part of the input. The elaboration of cost functions was based on analyzing a variety of technological process combinations (leading to different removal efficiencies) for treatment plants included in the analysis (Figure 5). As an example, the cost function for the largest treatment plant, Zalaegerszeg (see Figure 5) the capacity of which is $Q_c = 15000 m^3/d$, is given in Figure 8*. Three groups of expenses are illustrated in the figure:

(i) Investment cost required for upgrading biological treatment;

(ii) Investment cost of P precipitation which increases rapidly with increasing requirements. The use of piecewise linear cost functions required the introduction of three dummy variables for each treatment plants.

(iii) Running cost

* Roughly US $1 is equivalent to 50 Forints (Ft).

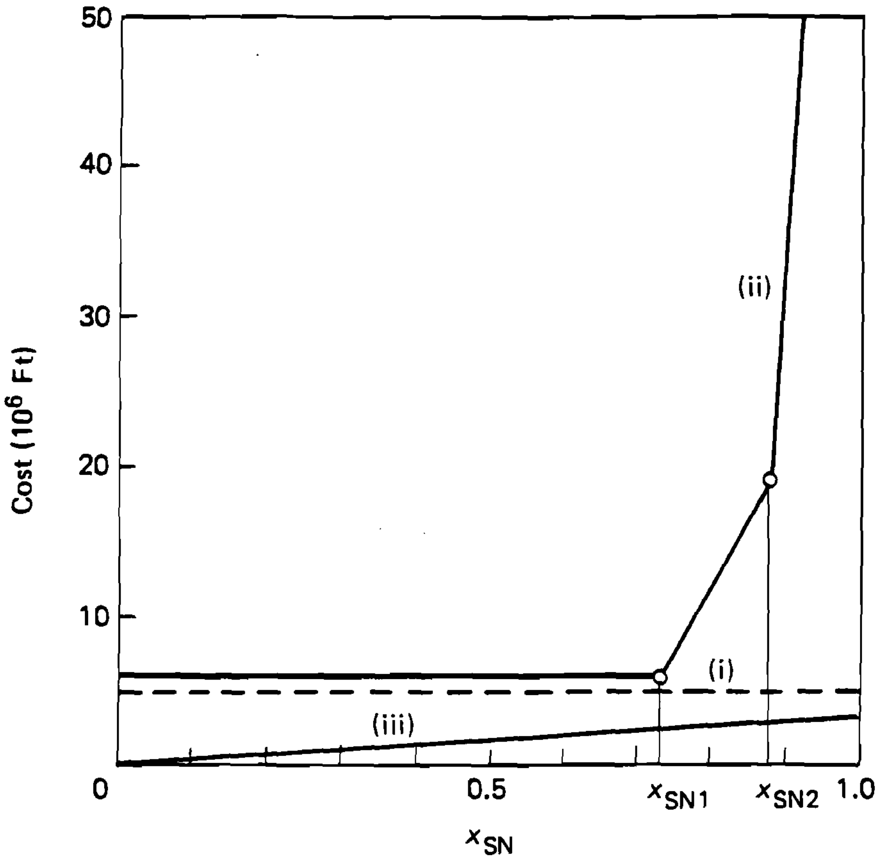


FIGURE 8. Costs of sewage treatment (Zalaegerszeg).

For pre-reservoirs linear cost functions of the surface area (and control variable) were used, see Section 4.2.

8.2. Results of the expectation-variance model

In order to gain an impression of the character of the problem and the behavior of the solution, first we specify a "basic situation" (which is close to the real case) having the following features and with the following assumptions (Somlyódy, 1983b):

(i) control variables are continuous;

(ii) no effluent standard prescription is given;

(iii) no P retention takes place in rivers ($r_l^{ml} = 0$, see Equation (3.10));

(iv) the capital recovery factor is equal for all the projects, $a_l = a = 0.1$ and

(v) equal weighting is adopted (see q_l and θ in Section 6).

With these assumptions optimization was performed under different budgetary conditions (TAC $\leq \beta = 0.5\text{--}25.10^7 Ft/yr$). Statistical parameters (expectation, standard deviation and extremes) of the water quality indicators gained from Monte Carlo procedure* are illustrated in Figure 9 for the Keszthely basin as a function of the total annual cost, TAC **.

In Figure 10, we record the changes in the two major control variables (x_{SN1} and x_{D1}) associated to the treatment plant of Zalaegerszeg and the reed lake segment of the Kis-Balaton system (see Figure 5). There is a significant trade-off between these two variables. For decision making purposes, it is important to observe that there are four ranges of possible values of β (the budget), in which the solution has different characteristics.

* 1000 simulations were performed in each case.
** Running cost is about ten times larger than TAC.

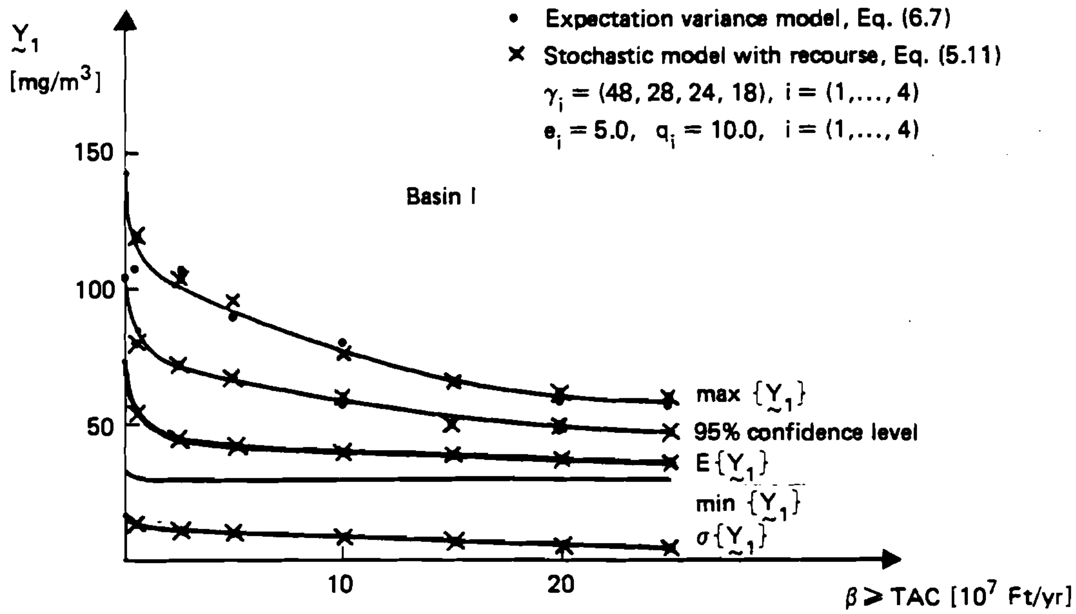


FIGURE 9. Water quality indicator $(\mathrm{Chl-a})_{max}$ as a function of the total annual cost.

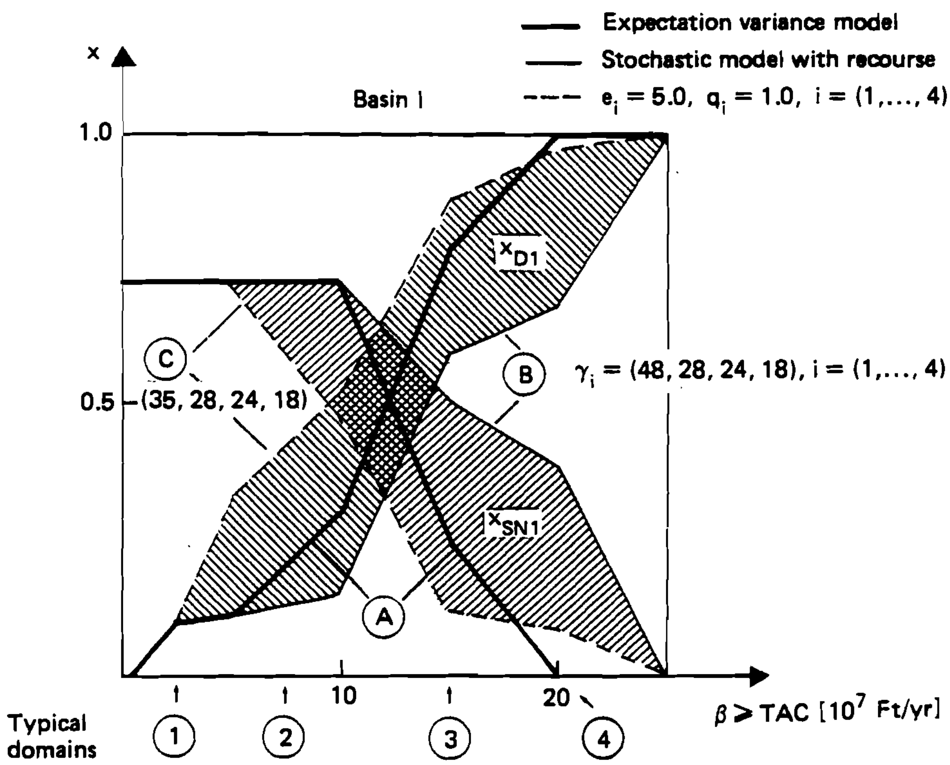


FIGURE 10. Change of major decision variables.

(i) In the range of $\beta = 0.5-5.10^7 Ft/yr$, it appears that sewage treatment can be intensified and tertiary treatment introduced. Expectation of the concentration levels will decrease considerably, but not the fluctuations. Under very small costs (~0.3.$10^7 Ft$ investment costs, IC) it turns out that only the sewage of Zalaegerszeg (Figure 5) should be treated. Under increasing budget, potential treatment plants are built, going from west to east.

(ii) If β is between 5.10^7 and $10.10^7 Ft/yr$, the effectiveness of sewage treatment cannot be increased further but reservoir system are still too expensive.

(iii) At about $\beta = 15.10^7 Ft/yr$ the solution is a combination of tertiary treatment and reservoirs. Fluctuations in water quality are reduced by the latter control alternatives.

(iv) Finally around $\beta = 20.10^7 Ft/yr$, tertiary treatment is dropped in regions where reservoirs can be built. After constructing all the reservoirs no further water quality improvement can be achieved.

Concerning the model sensitivity on major parameters, the following conclusions can be drawn (for details see Somlyódy, 1983b; and Somlyódy and van Straten ,1985):

(i) Fixed water quality standard not reflecting the properties of the system (spatial non-uniformities) can result in a strategy far from the optimal one, since the distribution of a portion of the budget is *a priori* determined by the pre-set standard.

(ii) Under increasing P retention in rivers the improvement in water quality is less remarkable in the budget range $0-10.10^7 Ft/yr$

than in the basic case. The worst — nevertheless nearly unrealistic — situation is if all the phosphorus was removed along the river and still treatment has to be performed: the budget should be partially allotted for investments having no influence on the lake's load.

(iii) If only deterministic effects are considered ($\theta = 0$), reservoir projects enter the solution under much larger budget values, see Section 8.4.

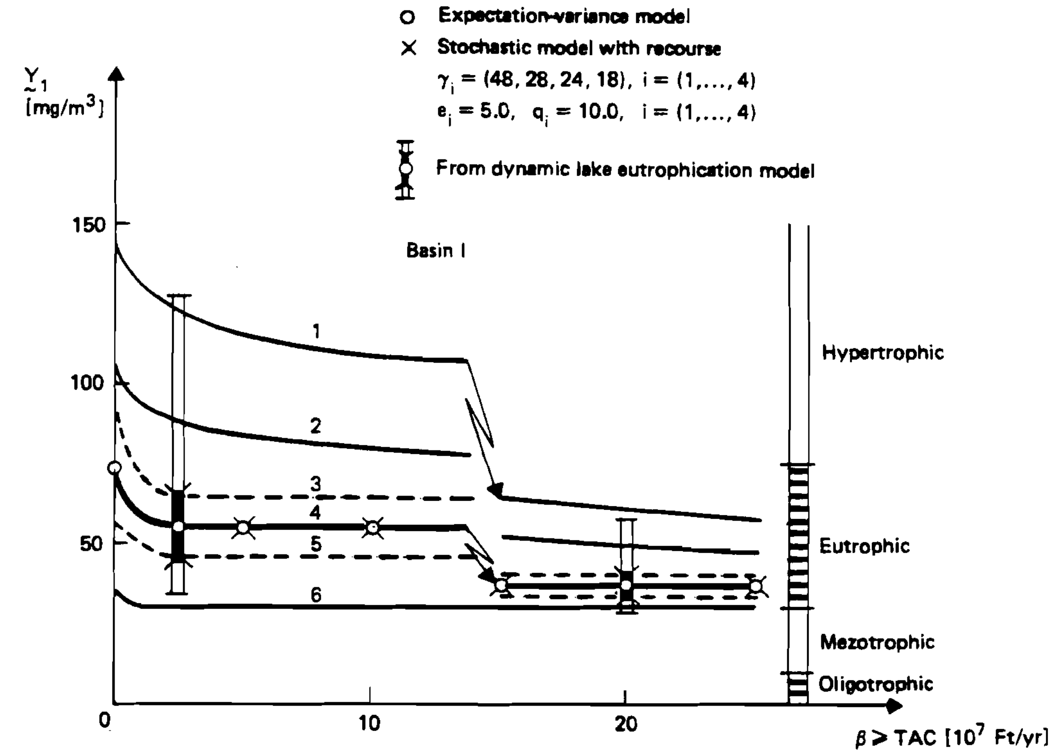
(iv) If the capital recovery factor is smaller for reservoir projects than for sewage treatment plants, see Section 4.3, reservoir projects start to be feasible at smaller budgets. Errors in the efficiency or in costs of reservoirs cause similar shifts in the solution.

(v) When selecting properly the model parameters, the combination of the absolute load reductions for the four basins is maximized by the model (as it is suggested most frequently in the literature, see the Introduction). Since, however, the absolute loads alone do not reflect the spatial changes in water quality, the policy drastically differs from the optimal one.

Subsequently we give the "realistic" solution for the Lake Balaton management problem by using

- actual retention coefficients (ranging between 0.3 - 0.5)
- upper limits 0.9 for the P removal rate of reservoirs; and
- fixed variables $\{0, 0.9\}$ for the Kis-Balaton reservoir system.

Figure 11*, which refers again to the Keszthely bay, shows remarkable differences as compared to Figure 9. First of all the drastic effect of reservoirs upon expectation but even stronger upon fluctuation of water quality is stressed. Reservoirs enter the solution between 15.10^7 and $17{,}5.10^7 Ft/yr$ total annual cost resulting in a reduction in the mean $(Chl-a)_{max}$ concentration from about 55 to 35 mg/m ^{3} and in the extreme values from more than 100 to about 60 mg/m ^{3}.



4: expectation; 3 and 5: $\pm\sigma$; 2: 95% confidence level; 1 and 6: extremes.

FIGURE 11. Solutions of EMOM for lake Balaton, Basin I.

* In the Figure $\pm$ standard deviation and the upper 95% confidence level are also illustrated (the distributions are bound towards small $\underset{\sim}{Y}_1$ concentrations and the lower 95% confidence level values are close to the minimum).

While Figure 9 offers several solutions for a decision maker depending on the budget available, on the basis of Figure 11, only two feasible alternatives come to mind:

(i) If total annual cost of about $2.5.10^7 Ft/yr$ is available, all the sewage projects can and should be realized (going from west to east). Through this alternative the expectation of $\underset{\sim}{Y}_1 = (\mathrm{Chl-a})_{max}$ is reduced to about 55 mg/m ^{3} (tertiary treatment affects the water quality at a slightly smaller extent than in the basic case due to P retention of tributaries) but still extremes larger than 110 mg/m ^{3} can occur (hypertrophic domain according to the classification of OECD, 1982). Further increase in the budget (up to $10.10^7 Ft/yr$) has no impact on water quality (under the alternatives included in the analysis).

(ii) If budget around $20.10^7 Ft/yr$ is given (not only the Kis-Balaton, but) all the reservoirs can be established and tertiary treatment can be realized for direct sewage sources. The mean $(\mathrm{Chl-a})_{max}$ concentration is about 35 mg/m ^{3} while the maximum about 60 mg/m ^{3} (eutrophic stage).

In Figure 11 the results of a detailed simulation model for two optimal solutions ($TAC = 2.5.10^7 Ft$ and $20.10^7 Ft$) are also given. The agreement between the calculated concentration indicators suggests that the aggregated lake eutrophication model is quite appropriate for our present purpose (Figure 1, Phase 4).

Figure 12 compares the typically skewed probability density functions of two considerably different solutions ($\beta = 2.5.10^7 Ft$ and $20.10^7 Ft$

respectively) for four basins, derived from Monte Carlo simulations (the non-controlled state is also given in this figure). Also from this figure we can conclude that tertiary treatment is more effective than reservoirs (when both alternatives are available) for controlling the mean concentration, but fluctuation can be controlled by reservoirs only. In the case ① ($\beta = 2.5 . 10^7 Ft / yr$) Basin I remains hypertrophic, Basins II and III eutrophic, whilst Basin IV mesotrophic. In the second situation ($\beta = 20.10^7 Ft / yr$) the spatial differences and stochastic changes are much smaller: Basin I...III are eutrophic and Basin IV mesotrophic (the long-term improvement of water quality is certainly larger than the short term one discussed here).

From all what we learned through the management model, it follows that in order to realize the optimal short term strategy of eutrophication management

- tertiary treatment of direct sewage discharges should be introduced (from west to east);
- depending on the budget available tertiary treatment of indirect sewage discharges of pre-reservoirs (again from west to east) should be realized.

For further details of the management strategy worked out for Lake Balaton and other management models not discussed in this paper, the reader is referred to Somlyódy (1983b) and Somlyódy and van Straten (1985).

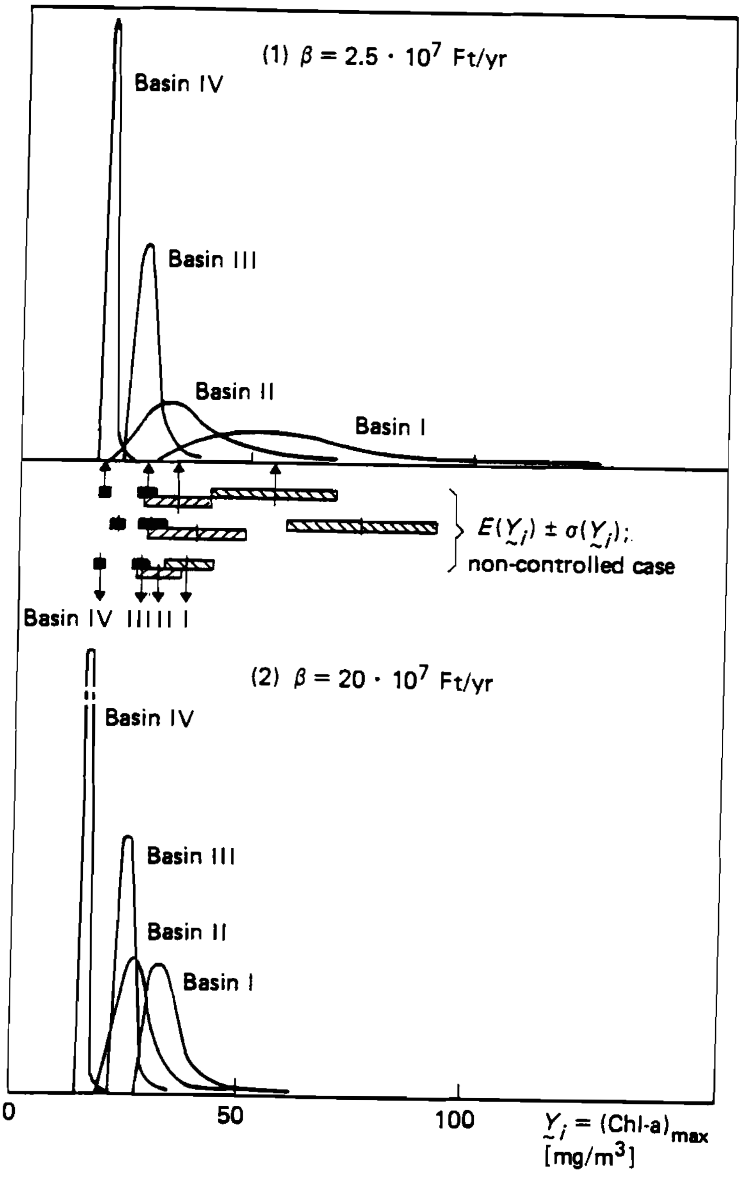


FIGURE 12. Probability density functions for two different situations (from 1000 Monte Carlo simulations).

8.3 Results of the Stochastic Recourse Model

As seen from Table 1 (line 9), the nominal state of water quality is given by the indicator vector Y_{0i} = (75,38,28,20),(i =(1,...,4)). Goals were specified by γ_i = (48,28,24,18) expressing the desire that Basin I should be shifted to the eutrophic, and other segments to the mesotrophic state (see Figure 11), but without forcing a completely homogeneous water quality in the entire lake on the short term, which would be unrealistic.

The definition of these goals, however, means that the improvement intended to be achieved is quite uniform for the four basins in a relative sense: as compared to the maximal possible reduction in the water quality indicator on the short term (see Figure 6), we plan 50-60% improvement for Basins I, II, and III. Basin IV (with 20%) is the only exception as its water quality is presently quite good, but this segment plays a secondary role from the viewpoint of the management problem.

Other parameters of the objective function (see (5.11) and Figure 3), e_i and q_i, were selected uniformly for the four basins: $e_i = 5$ and $q_i = 10$, $i = 1, \cdots, 4$. This corresponds, in the region $z_i \leq 5$, to a "variance formulation" of the objective function (being similar to (6.1)) as $q/2e = 1$. With these parameter values, the quadratic portion of the utility function is predominant in Basins II, III, and IV, while for Basin I the upper linear portion of the utility functional is also of importance.

Results of the stochastic optimization model with recourse are also illustrated in Figures 9-11, in comparison with that of the expectation-variance model. As seen from Figures 9-11, the two models produce practically the same results in terms of the water quality indicator (including also its distribution). With respect to details there are minor deviations. According to Figure 10, the expectation-variance model gives more emphasis to fluctuations in water quality (see also further on) and consequently to reservoir projects, than the stochastic recourse model (with the parameters specified above), and this is in accordance with the remarks made in Section 6, i.e., the role of the variance is overstressed in the expectation-variance model.

From this quick comparison of the performances of the two models, we may conclude that the more precise stochastic model validates the use of the expectation-variance model in the case of Lake Balaton.

For a more systematic comparison of the two models, the difference in the objective functions should be kept in mind: the stochastic model has more parameters than the expectation-variance model. In particular, the exclusion of the water quality goal from the expectation-variance model version plays an important role; the goal can be included only indirectly by modifying the weighting factors in (6.7). Figure 10 illustrates clearly that the prescription of the goal close to the lowest realizable value for Basin I (see Figure 6, 90% of the possible improvement in the short term) leads to a stronger emphasis on reservoirs as compared, not only to the original case, but also to the results of the expectation-variance model. The faster increase in x_{D1} (as a function of β, see Figure 10) is associated with a decrease in x_{SW1} — as expected — in addition to smaller budget allocations for the other basins. Depending on the value of γ_1, in the range (35, 48), the solutions lay in the shadowed areas indicated in Figure 10 and the solution of the expectation-variance model is located in the "center" of these domains.

As mentioned before, the expectation-variance model gives more weight to variance than the stochastic recourse model and for computational justification we compared curves (A) and (B) in Figure 10. The rationality for this comparison is that in lack of water quality goals and with equal weighting, the expectation-variance model follows to some extent the principle of "equal relative" water quality improvement in all basins (other factors — e.g. the distribution of costs for basins — play also important roles), and in this sense its solution can best be compared to solution (B) of the stochastic method.

Further discussion on the role of parameters γ_i, e_i and q_i, and on the comparison of the deterministic model (5.12) and the stochastic model is given in the subsequent section.

8.4 Comparison of the Deterministic and Stochastic Solutions -- Sensitivity Analysis

In order to gain experience with systems different from Lake Balaton, we changed some of the parameters of the Balaton problem and continue our presentation on the application of EMOM with this modified example. We call the hypothetical water body: Lake Alanton.

Lake Alanton differs from Balaton in the following respects:

(i) The volume of Basin IV is only $60 \cdot 10^6 m^3$ (see Table 1, line 5);

(ii) Two direct sewage loads are increased in the region of Basin IV (see Figure 5) to 50 and 70 kg/d, resulting in the same absolute BAP load than that of Basin I (the volumetric load is however larger due to (i); 2.3 $mg/m^3/d$, see Table 1, line 7);

(iii) The nominal water quality indicator and the slope of the response line was modified for Basin IV in such a manner, that response lines of Basins I and IV coincide (see Figure 6);

(iv) Cost functions of the two treatment plants were modified (they became similar to the one illustrated in Figure 9).

As a consequence of these changes, the water quality of Lake Alanton is approximately equally "bad" at its two ends. Still, however, an important difference exists between Basins I and IV: the load of Basin I is governed by "stochastic" tributary load, while that of Basin IV, is given by "deterministic" sewage discharges. This way the low water quality of Basin I is associated with large fluctuations (as seen in previous Sections), but randomness is of secondary importance for Basin IV.

The longitudinal distribution of the water quality is now quite different from that of Lake Balaton, and a manager may have the intention to establish a uniform quality by control decisions. Accordingly we fixed the goals to $\gamma_i = 30$ mg/m^3 $i=1, \cdots, 4$, and maintained the same parameter values as used in Section 8.3. Results for the "basic situation" (Section 8.2) are given in Figures 13 and 14.

From Figure 13 the same conclusions can be drawn for Basin I than from Figure 9. The only difference is that at a fixed budget the water quality improvement is slightly smaller than for Lake Balaton as a part of the budget is utilized for Basin IV. As seen from the figure, P precipitation is an effective tool for improving the water quality of Basin IV: the concentration $(\mathrm{Chl-a})_{\max}$ is reduced from 80 to about 40 mg/m^3 already at a budget of $2.5.10^7 Ft / yr$. The increase of β results in nearly no further change and

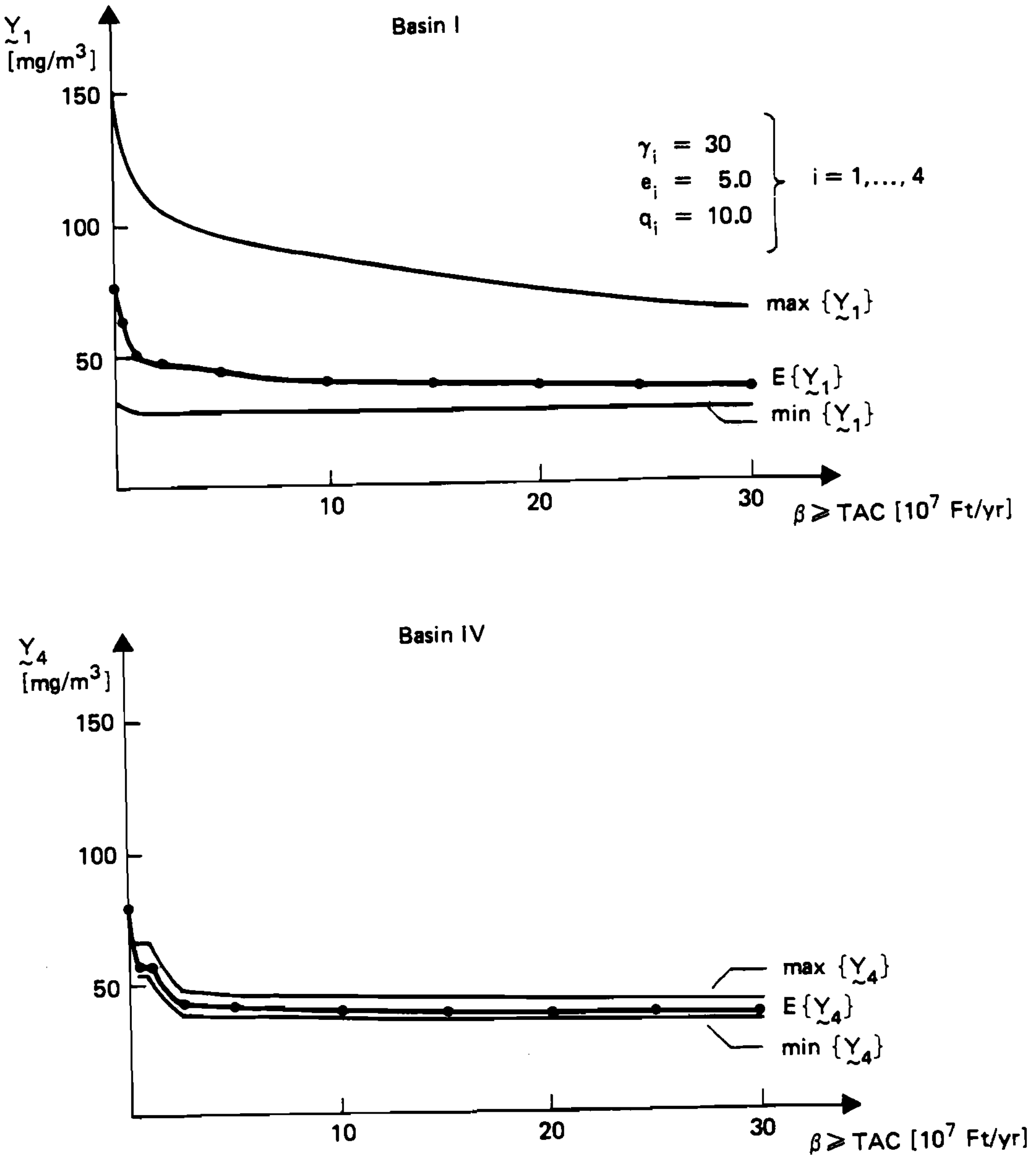


FIGURE 13. Lake Alanton: water quality indicator as a function of the total annual cost.

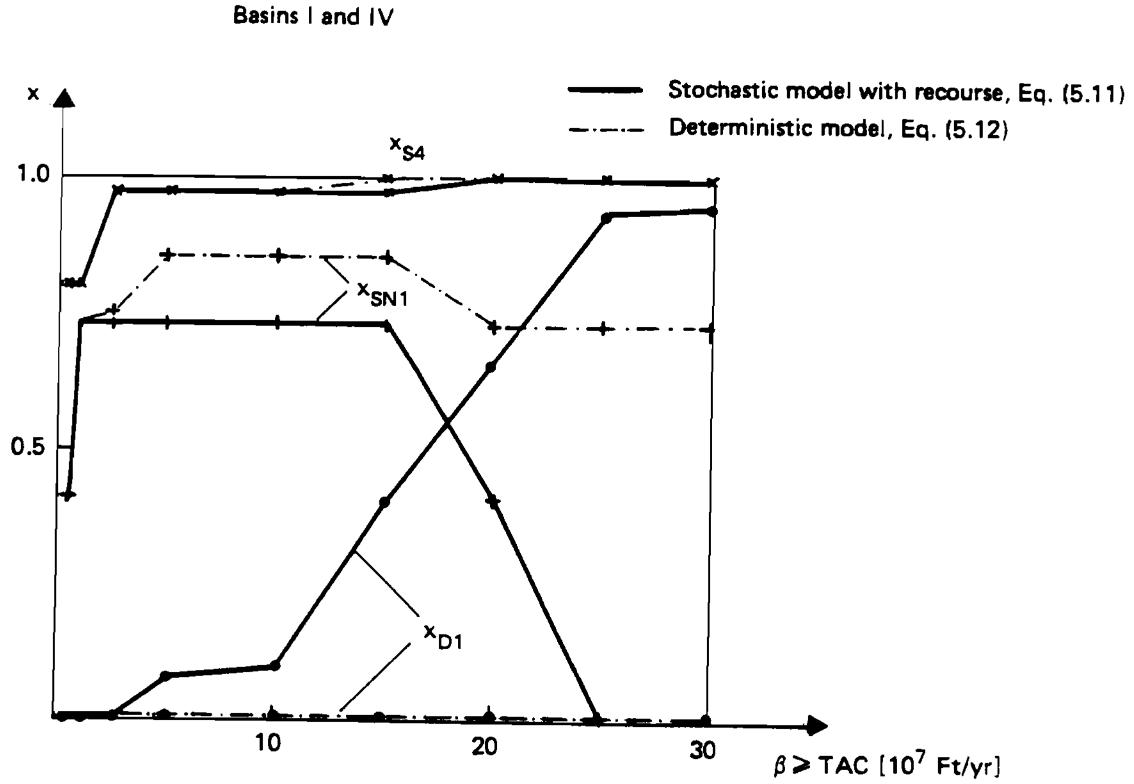


FIGURE 14. Lake Alanton: major decision variables.

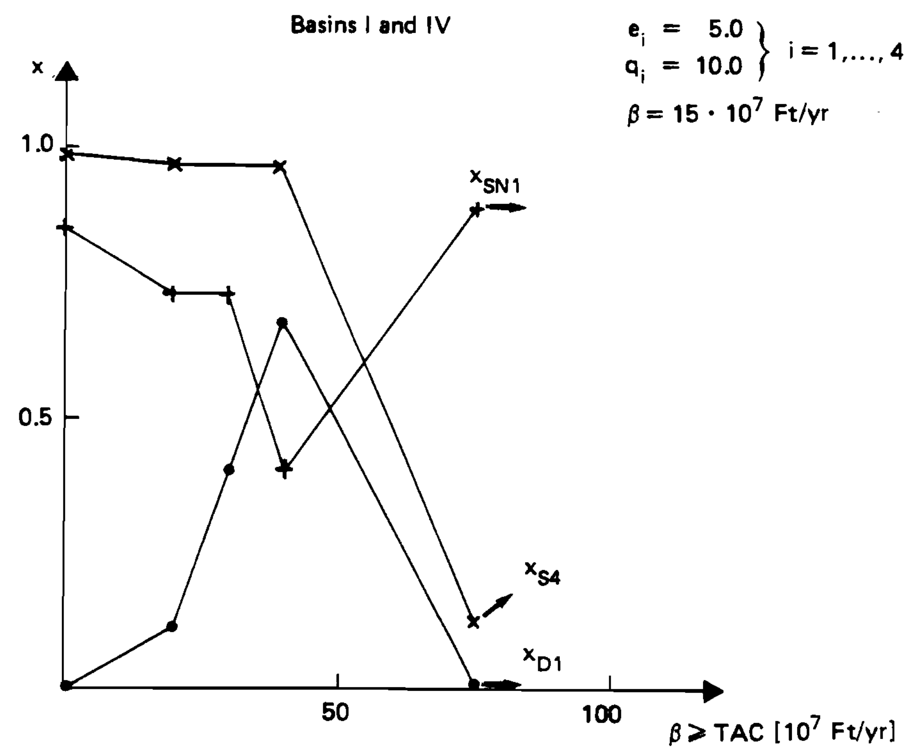


FIGURE 15. Sensitivity with respect to water quality goals.

the fluctuations also remain approximately constant, nevertheless small in comparison to Basin I.

In Figure 14, x_{S4} belongs to the largest treatment plant in the region of Basin IV. The character of $x_{D1}(\beta)$ and $x_{SN1}(\beta)$ gained from the stochastic model is the same as for Lake Balaton (Figure 10), while x_{S4} is practically constant above $\beta = 2.5.10^7 Ft/yr$. The most important conclusion of this figure can be drawn from the comparison of the stochastic and deterministic solutions: the deterministic quadratic model (5.12) excludes (incorrectly) the reed-lake part of the largest reservoir project even under the larger budget values, which is a consequence of neglecting the random elements of the problem. The $(\mathrm{Chl{-}a})_{max}$ concentration is reduced on an average to about 40 mg/m^3; however, extreme values still can exceed 100 mg/m^3; as contrasted to the solution of the stochastic model when the maximum is about 60 mg/m^3 This example shows clearly if the analyst is not able to recognize the stochastic features of the problem, the control strategy worked out may lead to serious failures.

The sensitivity of the solution with respect to the water quality goals is illustrated in Figure 15 referring to a typical budget $15.10^7 Ft/yr$. From the analysis performed and the figure, the following conclusions can be drawn:

(i) If the goal is set unrealistically far from the nominal value and from the values which can be realized by the available control measures (e.g. 0 or 200 mg/m^3 in this case), the penalty function has nearly no influence and thus the solution is equivalent to the deterministic one ($x_{D1} = 0$);

(ii) The solution is quite sensitive on the choice of the γ_i. The variable x_{D1} has a maximum at about 40 mg/m^3. If the goals uniformly or individually for Basins I and IV are close to 75-80 mg/m^3 the corresponding major decision variables are close to zero (no action is taken).

In summary, we can state that the water quality goal has a major influence on the solution, it forces indeed the solution towards the desired levels of water quality in the different lake basins. This feature is the primary advantage of the stochastic model with recourse as contrasted to the expectation-variance model for the Lake Balaton case and its variant Lake Alanton.

As to the role of the other parameters of the penalty function of the stochastic model is concerned, the systematic analysis performed did not lead to unambiguous conclusions. For both examples, Lake Balaton and Lake Alanton, the influence of e_i and q_i is of secondary importance as compared to the effect of γ_i. In general, it can be said that the influence is minor for small and large budgets. In the middle budget range it is difficult to separate the impact of e_i and q_i, especially because it strongly depends also on the preset goals γ_i, and the stochastic features of the water quality for the different basins.

The experience gained in the frame of the present study suggests the choice of a piecewise linear-quadratic utility function with $q / 2e = 1$ as a first step, see (5.9) and Figure 3, and to perform a thorough sensitivity analysis on the parameters γ_i, e_i and q_i in the subsequent steps.

We complete this section with the following conclusions:

(i) The stochastic optimization model with recourse justified the applicability of the much simpler expectation-variance model for Lake Balaton;

(ii) Deterministic version of the stochastic objective function and the solution of the corresponding deterministic quadratic optimization problem leads to strikingly different and incorrect management strategy as compared to the stochastic model;

(iii) The major parameter of the stochastic optimization model with recourse is the water quality goal prescribed for different basins. The inclusion of the goal in the objective function is the primary advantage in comparison with the expectation-variance model. The advantage of the latter model is, of course, simplicity and fast implementation;

(iv) Further experimentation is needed in the selection of parameters e_i and q_i (in the objective function of the stochastic model).

9. SUMMARY

In this paper we dealt with the development and application of stochastic optimization models for lake eutrophication management. We considered primarily shallow lakes which are strongly influenced by hydrologic and meteorologic factors and thus stochasticity should be a key component of water quality control.

Major elements of the study performed are as follows:

(i) Identification of important steps of eutrophication management in practice,

(ii) Based on the principle of decomposition and aggregation, an approach is presented how to develop a eutrophication management optimization model, EMOM, which preserves the scientific details of diverse in-lake and watershed processes needed at the decision-making level. The procedure combines simulation and optimization in the framework of EMOM.

(iii) We describe the proposed stochastic, planning type lake eutrophication and nutrient load models, LEMP and NLMP2, respectively, which are major components of EMOM.

(iv) We discuss control variables, cost functions and various constraints to be used in EMOM.

(v) Alternative management optimization models are formulated which use the same LEMP, NLMP2, control variables etc., and differ primarily in the objective function and solution technique to be adopted. Three of them were selected for implementation: a full stochastic model, an expectation-variance model, and the deterministic (quadratic) version of the full stochastic method. For the first one a new stochastic programming procedure had to be developed, while for the other two standard packages could be employed. The objective function of the full stochastic model has one more parameter as compared to the expectation-variance model: including the possibility of selecting the water quality goal to be achieved by the management, which makes this model especially attractive.

(vi) Application to Lake Balaton. This part of the study had a direct impact on the policy making procedure performed in Hungary in 1982 which ended up with a government decision in 1983.

(vii) Comparison of the two stochastic models, furthermore deterministic and stochastic approaches on the example of Lake Balaton and on a modified, hypothetical system (the comparison was associated with a detailed sensitivity analysis).

ACKNOWLEDGEMENT

In addition to the development of the computer code for the method described in Section 7, Alan King (University of Washington, Seattle) has contributed much to this project by his pertinent comments concerning algorithmic implementation and data analysis. We hope he finds a small sign of our appreciation in seeing the variant of Lake Balaton named in his honor.

REFERENCES

Beck, M.B. 1982. Time-Series Analysis of Zala River Nutrient Loadings. WP-82-116. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Biswas, A.K. (ed.) 1981. Models for Water Quality Management, pp.329. New York: McGraw-Hill International Book Company.

Bogárdi, J., L. Dávid, and L. Duckstein. 1983. Trade-off Between Cost and Effectiveness of Control of Nutrient Loading into a Water Body. RR-83-19. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Canale, R. 1976. Modeling Biochemical Processes in Aquatic Ecosystems. *Ann Arbor Science*. Ann Arbor, Michigan.

Cochran, W.G. 1962. *Sampling Techniques*. New York: John Wiley.

Dubois, D.M. (ed.) 1981. Progress in Ecological Engineering and Management by Mathematical Modelling. Liege, Belgium: Cebedoc.

Haith, D.A. 1982. Models for Analyzing Agricultural Nonpoint-Source Pollution. RR-82-017. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Herodek, S. 1983. Biochemical Processes in Lake Balaton, in *Eutrophication of Shallow Lakes: Modeling and Management. The Lake Balaton Case Study*, pp. 101-146. L. Somlyódy, S. Herodek, and J. Fischer (eds.). CP-83-53. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kallberg, J., and M. Kusy. 1976. Code Instruction for S.L.P.R., A Stochastic Linear Program with Simple Recourse. Vancouver: University of British Columbia.

King, A. 1985. An Implementation of the Lagrangian Finite-Generation Method, in *Numerical Methods for Stochastic Programming*, Y. Ermoliev and R. Wets (eds.). IIASA Collaborative Volume.

Kovács, L.B., E. Boros, and F. Inotay. 1983. A Two Stage Approach for Large Scale Sewer Systems Design with Application to the Lake Balaton Resort Area, in *Eutropication of Shallow Lakes: Modeling and Management. The Lake Balaton Case Study*, pp.315-333. L. Somlyódy, S. Herodek, and J. Fischer (eds.). CP-83-53. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kusy, M., and W. Ziemba. 1981. A Bank Asset and Liability Management Model. Technical Report, University of British Columbia, Vancouver, *Operations Research* (forthcoming).

Lam, D.C.L., and L. Somlyódy. 1983. A Comparison Between Lake Erie and Lake Balaton Water Quality Studies, in *Eutrophication of Shallow Lakes: Modeling and Management. The Lake Balaton Case Study*, pp.287-314. L. Somlyódy, S. Herodek, and J. Fischer (eds.). CP-83-53. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Láng, I. 1985. Impact on Policy Making: Background to a Governmental Decision, in *Modeling and Managing Shallow Lake Eutrophication with Application to Lake Balaton*. L. Somlyódy and G. van Straten (eds.) Springer Verlag (in press).

Lijklema, L., P. Gelencsér, and F. Srilágyi. 1983. Sediments and Sediment-Water Interaction, in *Eutrophication of Shallow Lakes: Modeling and Management. The Lake Balaton Case Study*, pp.81-100. L. Somlyódy, S. Herodek, and J. Fischer (eds.). CP-83-53. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Loucks, D.P., J.R. Stedinger, and D.A. Haith. 1981. *Water Resource Systems Planning and Analysis*, pp.559. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Monteith, T.J., R.C. Sullivan, and W.C. Sonzogni. 1980. Phosphorus Control Strategies at Municipal Waste-Water Treatment Plants in the U.S. Great Lakes Basin. GLEPS, Contribution No. 14, pp.43.

Murtagh, B., and M. Saunders, 1983. MINOS 5.0 User's Guide, Technical Report 83-20. Systems Optimization Laboratory, Stanford University.

Nazareth, L. 1984. Design and Implementation of Stochastic Optimizer with Recourse and Tenders, in *Numerical Methods for Stochastic Programming*, Y. Ermoliev and R. Wets (eds.). IIASA Collaborative Volume.

OECD. 1982. Eutrophication of Waters. Monitoring, Assessment and Control, pp.154. Paris.

Orlob, G.T. (ed.) 1983. Mathematical Modeling of Water Quality: Streams, Lakes, and Reservoirs. *12 International Series on Applied System Analysis*. J. Wiley, New York.

Rockafellar, R.T. 1976. Monotone Operators and the Proximal Point Algorithm. *SIAM J.Control and Optimization* 14:877-898.

Rockafellar, R.T., and R. Wets. 1983. A Dual Solution Procedure for Quadratic Stochastic Programs with Simple Recourse, in *Numerical Methods*, pp.252-265. V. Pereyra and A. Reinoza (eds.). Lecture Notes in Mathematics 1005. Springer Verlag, Berlin.

Rockafellar, R.T., and R. Wets, 1985. A Lagrangian finite Generation Technique for Solving Linear-Quadratic Problem in Stochastic Programming, in *Stochastic Programming: 1984*, A. Prekopa and R. Wets (eds.). *Mathematical Programming Study*.

Scavia, D., and A. Robertson. 1979. Perspectives on Lake Ecosystem Modeling. *Ann Arbor Science*. Ann Arbor, Michigan.

Schüssler, H. 1981. Kosten der Phosphatelimination in kommunalen Kläranlagen. *GWF-Wasser/Abwasser* 122. H.6, pp.251-257.

Somlyódy, L. 1982. Modeling a Complex Environmental System: The Lake Balaton Case Study. *Mathematical Modeling*, Vol. 3.

Somlyódy, L. 1983a. A Systems Approach to Eutrophication Management with Application to Lake Balaton. *Water Quality Bulletin* Vol. 9 No.1,25-37.

Somlyódy, L. 1983b. Lake Eutrophication Mangement Models, in *Eutrophication of Shallow Lakes: Modeling and Management. The Lake Balaton Case Study*, pp.207-250. L. Somlyódy, S. Herodek, and J. Fischer (eds.). CP-83-53. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Somlyódy, L., S. Herodek, and J. Fischer (eds.). 1983. Eutrophication of Shallow Lakes: Modeling and Management, pp.367. *The Lake Balaton Case Study*. CP-83-53. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Somlyódy, L. and G. van Straten (eds.) 1985. *Modeling and Managing Shallow Lake Eutrophication with Application to Lake Balaton*. Springer Verlag (in press).

Thomann, R.V. 1972. *Systems Analysis and Water Quality Management*, pp.293. New York: McGraw-Hill Book Company.

van Straten, G. 1981. Analysis of Model and Parameter Uncertainly in Simple Phytoplankton Models for Lake Balaton, in *Progress in Ecological Engineering and Management by Mathematical Modeling*, pp.107-134. D.M. Dubois (ed.). Liege, Belgium: Cebedoc.

Wets. R. 1983a. Stochastic Programming: Solution Techniques and

Approximation Schemes, in *Mathematical Programming: The State-of-the-Art*, pp.566-603. A. Bachem, M. Grötschel and B. Korte (eds.) Springer Verlag.

Wets, R. 1983b. Solving Stochastic Programs with Simple Recourse, *Stochastics* 10:219-242.