Working Paper

STOCHASTIC PROGRAMMING IN WATER RESOURCES SYSTEM PLANNING: A CASE STUDY AND A COMPARISON OF SOLUTION TECHNIQUES

J. Dupačová
A. Gaivoronski
Z. Kos
T. Szántai

August 1986
WP-86-40

International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

STOCHASTIC PROGRAMMING IN WATER RESOURCES SYSTEM PLANNING: A CASE STUDY AND A COMPARISON OF SOLUTION TECHNIQUES

J. Dupačová
A. Gaivoronski
Z. Kos
T. Szántai

August 1986
WP-86-40

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

FOREWORD

Stochastic programming methodology is applied in this paper to a capital investment problem in water resources. The authors introduce possible model formulations and then find the solutions by using a number of specific solution techniques. These are partly based on the SDS/ADO tape for stochastic programming problems and also on standard linear and nonlinear programming packages. Such an approach allows a thorough analysis of the solution as well as a comparison of the algorithmic procedures used to obtain these solutions.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

AUTHORS

Dr. Jitka Dupačová, research scholar at SDS in 1985–1986 is professor at the Department of Statistics, Faculty of Mathematics and Physics, Charles University, Prague.

Dr. Alexei Gaivoronski, research scholar at SDS is senior researcher of the V. Glushnov Institute of Cybernetics, Kiev, USSR.

Ing. Dr. Zdeněk Kos, research scholar at IIASA in 1980, is senior researcher at the Faculty of Engineering of the Technical University, Prague.

Dr. Tomás Szántai is professor at the Department of Operations Research of the Eötvös Loránd University, Budapest.

ABSTRACT

To analyze the influence of increasing needs upon a given water resources system in Eastern Slovakia and to get a decision on the system development and extension, several stochastic programming models can be used. The two selected models are based on individual probabilistic constraints for the minimum storage and for the freeboard volume supplemented by one joint probabilistic constraint on releases or by a nonseparable penalty term in the objective function. Suitable numerical techniques for their solution are applied to alternative design parameter values. As a result, the paper gives an answer to the case study which is based on multi modeling within the framework of stochastic programming and, at the same time, it gives a comparison of various solution techniques partly included in the SDS/ADO collection of stochastic programming codes.

STOCHASTIC PROGRAMMING IN WATER RESOURCES SYSTEM PLANNING: A CASE STUDY AND A COMPARISON OF SOLUTION TECHNIQUES

J. Dupačová, A. Gaivoronski,
Z. Kos and T. Szántai

1. INTRODUCTION

Operation of reservoirs in water resources system is a multistage stochastic control problem. Water resources planning has to consider multiple users and objectives, reservoir operation policies need to be analyzed to obtain an effective use of water resources. In the planning procedure, a great variety of water resources systems designs or operation plans need to be confronted and their economic, environmental and social impacts evaluated. Two basic tasks are often solved: (a) a new system is developed or an existing one is enlarged by some proposed investments in reservoirs, pipes, canals, pumping stations, hydroelectric water plants, etc., or (b) the management and control of an existing system has to be altered to accommodate to the new conditions. Analysis in both these cases rests on mathematical modeling, the objectives and constraint of the problem have to be expressed mathematically. The mathematical models involve the selection of many engineering, design and operating variables. The optimization means the determination of the best values of these variables regarding the constraints. A lot of the variables and parameters occurring in the objective function and in the constraints has to be taken as stochastic as they describe a stochastic real life problem.

Out of many possible goals, the water supply for industry and irrigation, flood control and recreation purposes are considered in this study. The most important decision variable is the storage capacity of reservoirs.

There were drawn up a lot of stochastic programming models for inventory control and water storage problems by Prékopa (1973). One of these models was applied for designing serially linked reservoir systems (Prékopa et al 1978). Another application of these models concerning the flood control reservoir design was described by Prékopa and Szántai (1978). In Czechoslovakia where the analyzed system is located, there is an old tradition to use chance constrained models for the considered type of problems. Applications of these models were described by Dupačová and Kos (1979), Kos (1979) in conjunction with the linear decision rule or with the direct control. Here, we shall follow this tradition and in addition, we shall consider the possible use of alternative stochastic programming models.

It is the evident stochastic nature of inflows which causes the most of the statistical and modeling problems and which causes the necessity of a detailed analysis of the data and of the variables occurring in any of model formulations. In water resources modeling, four types of variables occur:

(1) constant coefficients and parameters that are used for design values (e.g., active storage of the reservoir, reliability, cost coefficients) or variables with small variation (e.g., the withdrawal of water for the thermal power station);

(2) uncontrolled random variables with a known or estimated distribution (monthly flows, meteorologic variables, e.g., precipitation, potential evaporation);

(3) partly controlable random variables with known distribution (e.g. irrigation water requirements with controlable acreage);

(4) random variables with an incomplete knowledge of distribution, such as the future demands, future prices and costs.

In our paper the type (3) of stochastic variables were incorporated into the model taking into account their *joint* probability distribution. In this way the interrelations between the succeeding months values of irrigation water requirements could be considered. It is the theory of logconcave measures developed by Prékopa (1971) which gives the theoretical background for the handling of joint probability constraints in stochastic programming problems.

We pursued in this study two objectives. One is to develop optimization stochastic models to determine the required reservoir capacity under increasing needs. The goal of the model is to show how the irrigation, flood control and recreation needs influence the reservoir operation and the reliability in meeting the multiple goals of the water resources system. This was done in section 2, 3 and 6 which contain the description of the water resources system, as well as several suitable stochastic programming models and an application of one of them to the basin of the Bodrog River in the Eastern part of Slovakia (Czechoslovakia), see

Figures 1 and 2.

Another of our objective was to use this comparatively simple but meaningful and real life problem to test various approaches for solution of stochastic programming problems and co..pare various solution techniques which are implemented in IIASA and constitute SDS/ADO collection of stochastic programming codes. The results of this comparison can be found in sections 4 and 5.

2. THE WATER RESOURCES SYSTEM

The water resources system consists of three reservoirs, V., D., K., two of which are in operation (D., V.) and the third one is to be built or not (K.) – see Figure 3. The main purposes of the water resources system are the irrigation water supply, industrial uses – mainly water withdrawal for the thermal power station, flood control (better flood alleviation), environmental conservation and recreation. The subsystem of the Laborec river (see Figure 4) was originally designed for industrial water supply and flood control. During the operation of the reservoir V., an accelerated development of the recreation occurred and the demands for maintenance of the minimum recreation pool during the summer period were supported by authorities. The area of irrigation grew and according to the plan will be increased substantially. The main questions for the decision-makers are:

- Can the presented water resources system still meet all the requirements? And if so, with which reliability?
- Is the construction of the reservoir K. necessary and when it will be necessary?

The analysis of this problem was divided into two steps. The first one comprises the screening modeling and it is discussed in this paper. Optimization models, however, cannot reflect all the details of the water resources system operation. Therefore the results of the stochastic optimization model are supposed to be verified using the stochastic simulation model with the input generated by the methods of stochastic hydrology. For the stochastic programming screening model, which is the subject of the present paper, an aggregated model was used and the monthly flows and the irrigation water requirements were aggregated into four periods:

FIGURE 1 Location of the water resources system of the Bodrog river.

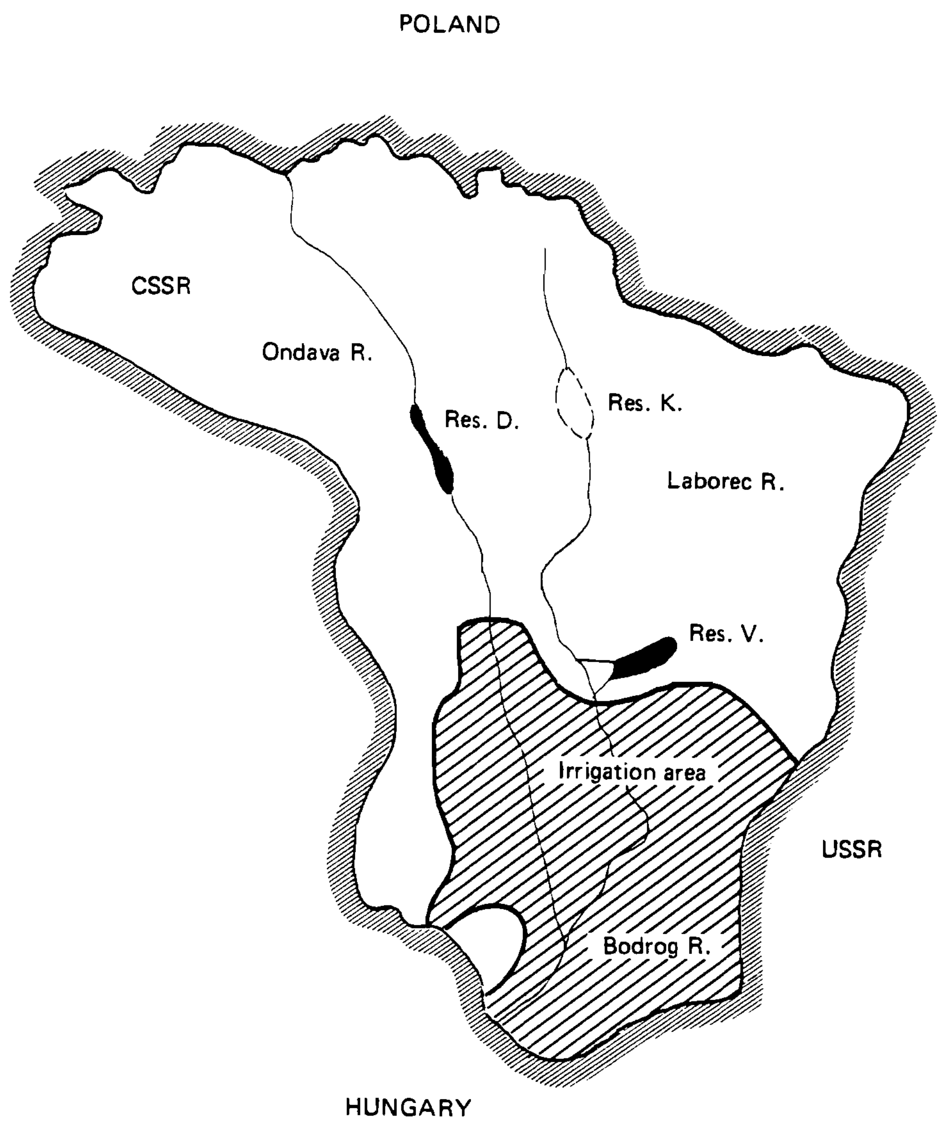

FIGURE 2 The water resource system allocation.

(1) November till April of the following year,

(2) May and June,

(3) July and August,

(4) September and October.

The first period starts at the beginning of the hydrological year and comprises the winter and the spring periods filling the reservoirs. As the main interest of the irrigation is concentrated on the vegetation period and recreation season, the aggregation of the six months is acceptable. The second and fourth periods include irrigation and industrial demands, the third period includes in addition the recreation demands. The requirements for the minimum pool due to environmental

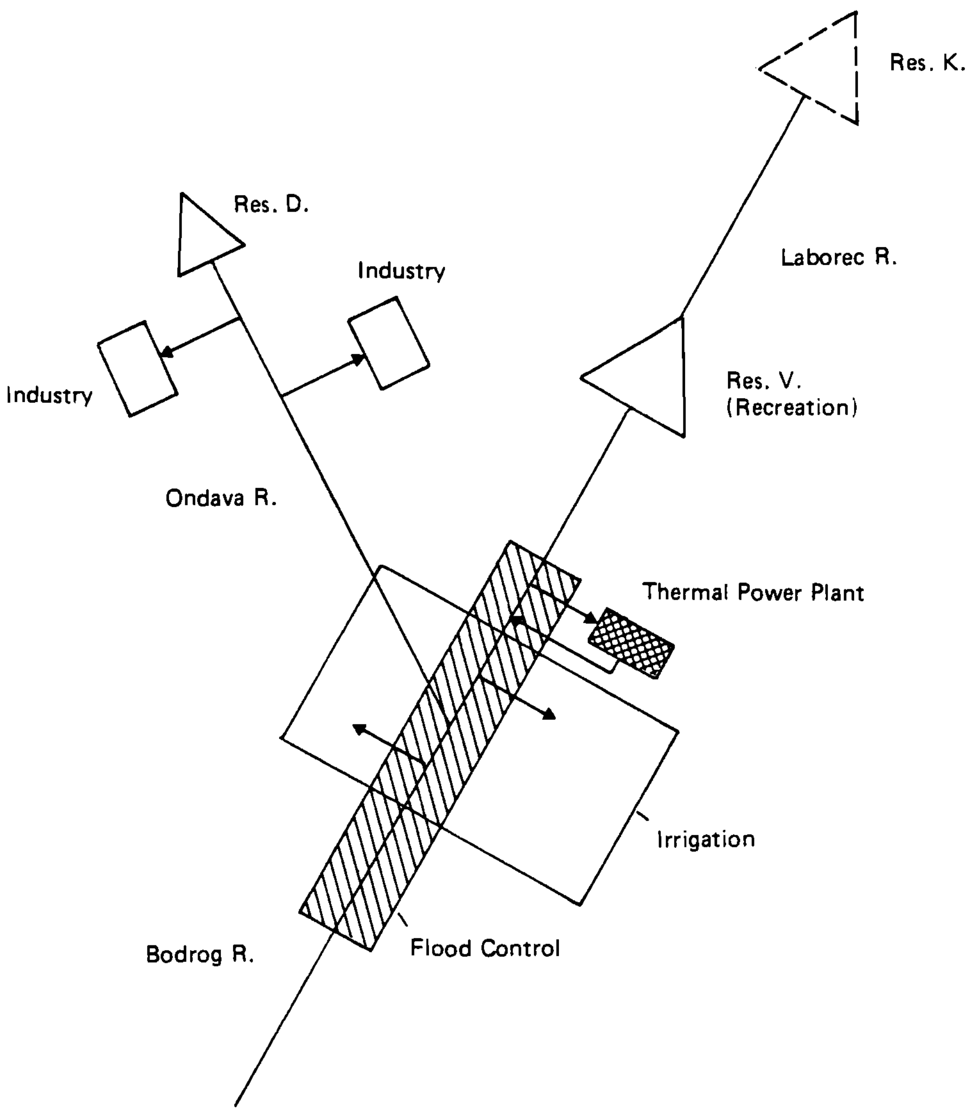

FIGURE 3 Schematic representation of the water resources system of the Bodrog river.

control and enhancement and flood control pertain to all the periods.

3. THE MATHEMATICAL MODELS

The models were designed for screening alternatives on the cost basis.

Using ***probabilistic constraints***, the result identifies the capacity x_0 of reservoir V. that should meet the needs with a prescribed reliability. The task is formulated as the cost of reservoir V. minimization. As the cost is an increasing

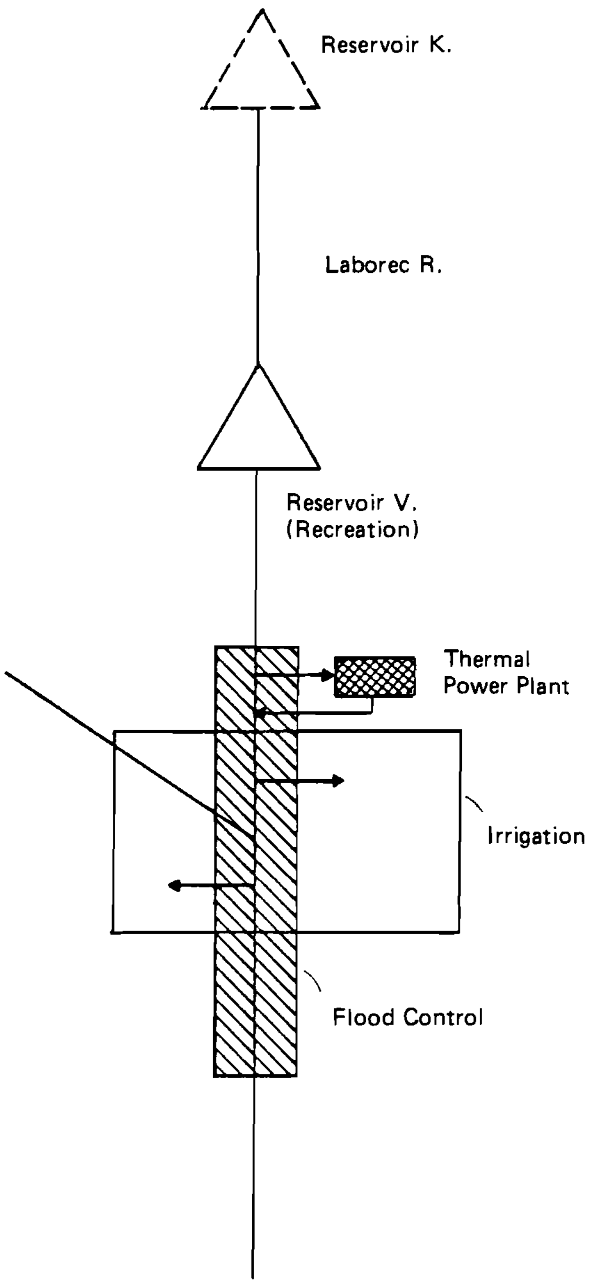

FIGURE 4 Schematic representation of the subsystem of the Laborec river.

function of reservoir capacity x_0, we can evidently minimize the capacity x_0 instead of minimizing the cost.

The constraints involve exceeding of water supply need $\beta_i + d_i$ by the released volume x_i in periods 2, 3 and 4 (vegetation periods). The needs consist of the fixed demand d_i (minimum flow and industrial water needs) and random demand β_i (irrigation water requirements). As to the first period, the fixed demand d_1

(caused mainly by the needs of the thermal power station) should be met with such a high probability that the deterministic constraint $x_1 \geq d_1$ was used. Taking into account the intercorrelations of random demands β_i, $i = 2, 3, 4$, the constraint for the vegetation period, was formulated as follows

$$P\{x_i \geq \beta_i + d_i, \ i = 2, 3, 4\} \geq \alpha \ , \qquad (1)$$

where α is the required joint probability level.

The second type of constraints reflects the environmental control and enhancement, fishing and recreation needs. These requirements are expressed in the form of maintaining a minimum pool or minimum reservoir storage m_i. In periods 1, 2 and 4, the environmental and fishing needs are reflected, in period 3 the recreation needs are added. This constraint is then as follows

$$P\{s_i \geq m_i\} \geq \alpha_i, \ i = 1, \ldots, 4 \ , \qquad (2)$$

where s_i is the reservoir storage, α_i the required individual probability thresholds in period i.

The third type of constraint expresses the flood control target assuming that some flood control storage v_i will be free in the reservoir operation during the whole period i with probability γ_i, i.e.,

$$P\{s_i + v_i \leq x_0\} \geq \gamma_i, \ i = 1, \ldots, 4 \ . \qquad (3)$$

The resulting optimization problem has the form

$$\begin{aligned} & \text{minimize } x_0 \\ & \text{subject } P\{x_i \geq \beta_i + d_i, \ i = 2, 3, 4\} \geq \alpha \ , \\ & \qquad P\{s_i \geq m_i\} \geq \alpha_i, \ i = 1, \ldots, 4 \ , \\ & \qquad P\{s_i + v_i \leq x_0\} \geq \gamma_i, \ i = 1, \ldots, 4 \end{aligned} \qquad (4)$$

and subject to additional constraints

$$\begin{aligned} & d_1 \leq x_1 \leq u_1 \\ & l_0 \leq x_0 \leq u_0 \\ & \quad x_i \leq u_i, \ i = 2, 3, 4 \ , \end{aligned} \qquad (5)$$

which stem mostly from natural hydrological and morphological situation. The upper bounds u_i, $i = 1, \ldots, 4$, are given by the volume of flood with respective duration and probability of exceedance.

Using the direct (zero-order) decision rule and neglecting losses due to evaporation, the reservoir storages s_i can be expressed via the water inflows and releases in the relevant periods. Let r_j denote the water inflow in the j-th period and let ζ_i denote the cumulated water inflow,

$$\zeta_i = \sum_{j=1}^{i} r_j, \; i = 1, \ldots, 4 \; .$$

Denote further by s_0 the initial reservoir storage at the beginning of the hydrological year. As a rule, we can put $s_0 = m_4$, i.e., the reservoir storage is supposed to be at its minimum after the vegetation period. Repeated use of the continuity equation gives

$$\begin{aligned}
s_1 &= m_4 + \zeta_1 - x_1 \\
s_2 &= m_4 + \zeta_2 - x_1 - x_2 \\
s_3 &= m_4 + \zeta_3 - x_1 - x_2 - x_3 \\
s_4 &= m_4 + \zeta_4 - x_1 - x_2 - x_3 - x_4
\end{aligned} \tag{6}$$

Substituting into (2) and (3) yields

$$\begin{aligned}
&P\{\zeta_1 \geq m_1 - m_4 + x_1\} \geq \alpha_1 \\
&P\{\zeta_2 \geq m_2 - m_4 + x_1 + x_2\} \geq \alpha_2 \\
&P\{\zeta_3 \geq m_3 - m_4 + x_1 + x_2 + x_3\} \geq \alpha_3 \\
&P\{\zeta_4 \geq x_1 + x_2 + x_3 + x_4\} \geq \alpha_4
\end{aligned} \tag{7}$$

$$\begin{aligned}
&P\{\zeta_1 \leq x_0 + x_1 - m_4 - v_1\} \geq \gamma_1 \\
&P\{\zeta_2 \leq x_0 + x_1 + x_2 - m_4 - v_2\} \geq \gamma_2 \\
&P\{\zeta_3 \leq x_0 + x_1 + x_2 + x_3 - m_4 - v_3\} \geq \gamma_3 \\
&P\{\zeta_4 \leq x_0 + x_1 + x_2 + x_3 + x_4 - m_4 - v_4\} \geq \gamma_4
\end{aligned} \tag{8}$$

Using the corresponding 100p% quantiles $z_k(p)$ of the distribution of the random variables ζ_k, $k = 1, 2, 3, 4$, we can rewrite the individual probabilistic constraints (7) and (8) in the form

$$\sum_{i=1}^{k} x_i \le z_k(1-\alpha_k) + m_4 - m_k,\ k = 1,\ldots,4 \tag{9}$$

$$\sum_{k=0}^{k} x_i \ge z_k(\gamma_k) + m_4 + v_k,\ k = 1,\ldots,4\ . \tag{10}$$

Unfortunately, this simple device does not apply to the joint probabilistic constraint (1).

The resulting optimization problem

minimize x_0

subject $P\{x_i \ge \beta_i + d_i,\ i = 2, 3, 4\} \ge \alpha$

$$\sum_{i=1}^{k} x_i \le z_k(1-\alpha_k) + m_4 - m_k,\ k = 1,\ldots,4\ ,$$

$$\sum_{i=0}^{k} x_i \ge z_k(\gamma_k) + m_4 + v_k,\ k = 1,\ldots,4\ , \tag{11}$$

$$l_0 \le x_0 \le u_0$$

$$d_1 \le x_1 \le u_1$$

$$x_i \le u_i,\ i = 2, 3, 4$$

can be solved e.g. by special techniques developed by Prékopa et al. (1978), see Section 4.

Alternatively, stochastic programming decision models can be built solely on the evaluation and minimization of the overall expected costs, which contain not only the cost of the reservoir of the capacity x_0 but also the losses connected with the fact that the needs were not fulfilled and/or the requirements on the minimum reservoir storage and on the flood control storage were not met. This type of models is called mostly *stochastic programs with recourse or with penalties*.

Suppose first that the constraints on water supply, minimum water storage and flood control storage do not contain random variables, i.e., we have (besides of the upper and lower bounds (5))

$$x_i \ge \beta_i + d_i,\ i = 2, 3, 4$$

$$\sum_{i=1}^{k} x_i \le \zeta_k + m_4 - m_k, \ k = 1, \ldots, 4 \tag{12}$$

$$\sum_{i=0}^{k} x_i \ge \zeta_k + m_4 + v_k, \ k = 1, \ldots, 4 \ .$$

In case of random β_i, $i = 2, 3, 4$ and ζ_k, $k = 1, \ldots, 4$, the chosen decision x_0, $x_1, \ldots, x_4$ need not fulfill constraints (12) for the actual (observed) values of β_i, ζ_k. If this is the case, costs evaluating losses on crops (not being irrigated on a sufficiently high level), on the decrease of recreation (due to the lower reservoir pool) and the economic losses due to flood can be attached to the discrepancies.

Let $c(x_0)$ denote the cost of reservoir of the capacity x_0, let the penalty functions be of the type

$$\begin{aligned} \varphi(y) &= 0 \quad \text{if } \ y \le 0 \\ &\ge 0 \quad \text{and nondecreasing if } \ y > 0 \ . \end{aligned} \tag{13}$$

Denote by φ_i^1, φ_k^2, φ_k^3 the penalty functions corresponding to the considered three types of constraints in (12), $i = 2, 3, 4$, $k = 1, \ldots, 4$. We try to find such a decision for which the total expected cost will be minimal subject to inequality constraints (5):

$$\text{minimize } c(x_0) + E\left\{ \sum_{i=2}^{4} \varphi_i^1(\beta_i + d_i - x_i) \right. \tag{14}$$

$$\left. + \sum_{k=1}^{4} \varphi_k^2\left[\sum_{i=1}^{k} x_i - \zeta_k - m_4 + m_k\right] + \sum_{k=1}^{4} \varphi_k^3\left[\zeta_k + m_4 + v_k - \sum_{i=0}^{k} x_i\right]\right\}$$

$$\text{subject to } l_0 \le x_0 \le u_0$$

$$d_1 \le x_1 \le u_1$$

$$x_i \le u_i, \ i = 2, 3, 4 \ .$$

The choice of the penalty functions should be based on a deep economic and environmental analysis of the underlying problem. On the other hand, for a screening study, it seems satisfactory to restrict the choice to piece-wise linear or piece-wise quadratic penalty functions.

a) *piece-wise linear penalty* (simple recourse model) All penalty functions φ are of the form

$$\varphi(y) = qy^+, \quad \text{where } q \geq 0 \quad \text{and } y^+ = \max(0, y) \ .$$

The coefficient q has to be given by the decision maker; it corresponds to the unique costs for the different considered discrepancies. As a result, we have to

$$\text{minimize } c(x_0) + E\left\{ \sum_{i=2}^{4} a_i(\beta_i + d_i - x_i)^+ + \sum_{k=1}^{4} b_k \left[\sum_{i=1}^{k} x_i + m_k - m_4 - \zeta_k \right]^+ + \sum_{k=1}^{4} c_k \left[\zeta_k + m_4 + v_k - \sum_{i=0}^{k} x_i \right]^+ \right\} \qquad (15)$$

subject to the constraints (5).

The solution method is mostly based on approximation of the marginal distributions of β_i, ζ_k by discrete ones and, in case of $c(x_0)$ linear, the resulting program can be solved by simplex method with upper bounded variables (see e.g. the SPORT program (by Nazareth) contained in the ADO/SDS tape). For other types of piece-wise linear penalties see e.g. Dupačová (1980).

b) If the cost function $c(x_0)$ is strictly quadratic on the considered interval $<l_0, u_0>$, a special type of *piece-wise quadratic penalty* function was suggested by Rockafellar and Wets (1985), namely,

$$\begin{aligned} \varphi(y) &= 0 \quad \text{for } y \leq 0 \\ &= \frac{1}{2} y^2/p \quad \text{for } 0 \leq y \leq pq \\ &= qy - \frac{1}{2} pq^2 \quad \text{for } y \geq pq \end{aligned}$$

(see Figure 5).

For solving problem (15) with penalty functions of the mentioned type, program LFGM (by King) implementing Rockafellar and Wets' Lagrangian Finite Generation Method is at disposal on the ADO/SDS tape. The parameters p, q of the individual penalty functions have to be given by the decision maker.

The use of the stochastic program (15) with penalties does not require the knowledge of the joint distribution of the random needs for irrigation water. It means, that due to the assumed separability of the penalty functions (i.e., due to the fact that shortage in irrigation water is penalized in each of the three vegetation periods separately and the total penalty is taken as the sum over the three

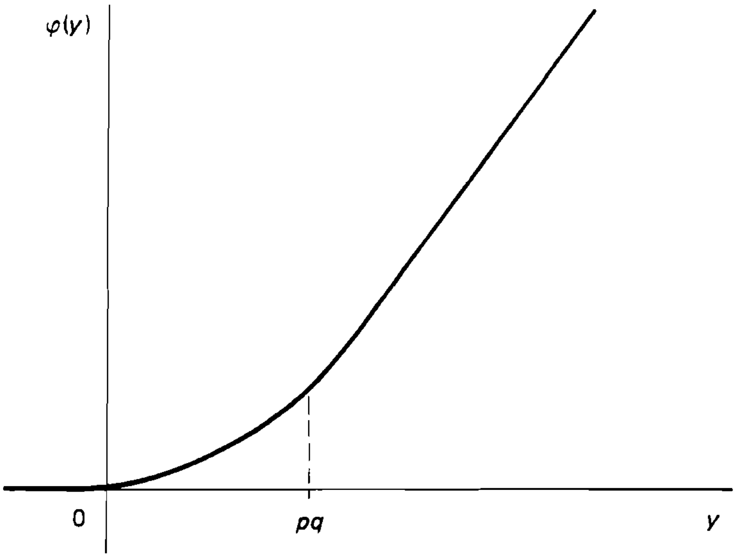

FIGURE 5

periods), no intercorrelations are considered. Alternatively, we can attach a penalty cost to the situation, when the total needs for water irrigation in the vegetation period as a whole were not met. In that case, we can take e.g.

$$\tilde{\varphi}(y) = \varphi(\max_{i} y_i)$$

with φ of the form (12) and minimize

$$c(x_0) + E\left\{\varphi\left[\max_{i=2,3,4}(\beta_i + d_i - x_i)^+\right] + \sum_{k=1}^{4} \varphi_k^2\left[\sum_{i=1}^{k} x_i - \zeta_k - m_4 + m_k\right] \right. \tag{16}$$

$$\left. + \sum_{k=1}^{4} \varphi_k^3\left[\zeta_k + m_4 + v_k - \sum_{k=0} x_i\right]\right\}$$

subject to constraints (5).

Finally, it is possible to combine the probalistic constraints and the penalization: one can define the set of admissible decisions by means of probabilistic constraints (4) and inequalities (5) and, at the same time penalize the occurrence of the discrepancies in (12) by corresponding penalty terms in the objective function.

In this case study, the set of admissible solutions was defined by the individual probabilistic constraints on a minimum pool and on the flood control storage, i.e., by the system of inequalities (9), (10), and by inequalities (5). Instead of the joint probabilistic constraint on the water supply needs, one penalty term of the form

$$\tilde{\varphi}(\beta_i + d_i - x_2,\ i = 2, 3, 4) = c\left\{\max_{i=2,3,4} (\beta_i + d_i - x_i)^+\right\}$$

$$= c \max\left\{0,\ \max_{i=2,3,4} (\beta_i + d_i - x_i)\right\}$$

was used. The resulting problem

$$\text{minimize } x_0 + cE\left\{\max_{i=2,3,4} (\beta_i + d_i - x_i)^+\right\}$$

$$\text{subject to } \sum_{i=1}^{k} x_i \le z_k(1 - \alpha_k) + m_4 - m_k,\ k = 1, \ldots, 4 \ ,$$

$$\sum_{i=0}^{k} x_i \ge z_k(\gamma_k) + m_4 + v_k,\ k = 1, \ldots, 4 \ ,$$

$$l_0 \le x_0 \le u_0 \ ,$$

$$d_i \le x_i \le u_i,\ i = 1, \ldots, 4$$

can be solved by the stochastic quasigradient method Ermoliev (1976) or by techniques designed to solution of the complete recourse problem; see Section 4. For the optimal solution of (17), the values of the joint probability

$$p(x) = P\{x_i \ge \beta_i + d_i,\ i = 2, 3, 4\}$$

were computed.

Observe that in (17), the cost $c(x_0)$ of the reservoir of capacity x_0 is supposed to be linear in the interval $l_0 \le x_0 \le u_0$ and that the coefficient c of the penalty term evaluates unit losses due to water shortage *relative* to the cost per unit capacity of the reservoir.

4. SOLUTION TECHNIQUES USED TO SOLVE THE PROBLEM

In this section we shall describe briefly the solution techniques used for numerical experiments with models (11) and (17) described in the previous section. We could choose among several stochastic optimization programs from the SDS/ADO stochastic optimization library available at IIASA.

For solving problem (11) with the joint probability constraint, nonlinear programming techniques can be used. The choice among them depends on the properties of the set of feasible solutions. For log-concave probability measures, which is the case of multidimensional normal, gamma, un'form, Dirichlet distributions of β, the set described by

$$p(x) = P\{x_i \geq \beta_i + d_i, \; i = 2, 3, 4\} \geq \alpha$$

is convex and among others, the method of feasible directions, supporting hyperplane method and penalty methods supplemented by an efficient routine for computing the values of the function $p(x)$ and its derivatives can be applied. (For a survey see Prékopa (1978), (1986).) For multinormal distribution, the supporting hyperplane method was implemented at IIASA by Szántai as PCSP code.

For solving problem (17), one possibility is to approximate the original distribution of β by a sequence of *discrete* distributions. The penalty $\tilde{\varphi}$ can be given implicitly via so called *second stage program*: for fixed values of x_i, β_i, d_i, $i = 2$, 3, 4, $\tilde{\varphi}(\beta_i + d_i - x_i, \; i = 2, 3, 4)$ equals to the optimal value of the objective function in the following linear program

$$\begin{aligned} &\text{minimize } cy \\ &\text{subject to } x_i + y \geq \beta_i + d_i, \; i = 2, 3, 4 \\ &\qquad\qquad y \geq 0 \end{aligned}$$

so that the problem (17) can be considered as the complete recourse problem and solved accordingly. For discrete distribution the approach leads to linear programming problems of special structure which should be exploited by adequate solution techniques. One such technique is L-shaped algorithm by Van Slyke and Wets (1969) implemented at IIASA by Birge in NDSP code. For further exposition see e.g. Kall (1979), Wets (1983).

Finally, to stochastic programming models of expectation type such as (17), stochastic quasigradient (SQG) method can be applied, see Ermoliev (1976), Ermoliev and Gaivoronski (1984). The stochastic optimization solver STO based on this method solves the following problem

$$\text{minimize } E_\omega f(x, \omega) = F(x) \tag{18}$$

subject to constraints

$$x \in X \tag{19}$$

where $f(x, \omega)$ is a functicn which depends on decision variables x and random parameters ω. The user has to provide an algorithmic description of this function. The set X is the set of constraints and in the current implementation, it would be the set defined by linear constraints, say,

$$X = \{x : Ax \geq b\} \ .$$

The optimal solution of (18)–(19) is reached iteratively starting from an initial point x^0 by applying the following iterative procedure:

$$x^{s+1} = \pi_X(x^s - \rho_s \xi^s) \tag{20}$$

where ρ_s is the stepsize, ξ^s – step direction, π_X – projection operator on the set X:

$$\| \pi_X(x) - x \| = \min_{z \in X} \| z - x \|, \ \pi_X(x) \in X \ .$$

The projection in the STO is performed using QPSOL quadratic programming package, see Gill et al. (1983).

The step direction ξ^s should, roughly speaking, be in average close to the gradient of the objective function $F(x) = E_\omega f(x, \omega)$ at point x^s, although individual ξ^s may be far from actual values of the gradient. This is expressed with the help of conditional expectations:

$$E(\xi^s \mid x^0, \ldots, x^s) = F_x(x^s) + \alpha_s$$

where α_s is some vanishing term. Each particular strategy of choosing sequence of stepsizes ρ_s and step directions ξ^s lead to particular algorithm and many such strategies are implemented in the program STO some of them are fairly sophisticated. It is also possible to change strategies interactively during optimization process. Detailed description of this program is given in Ermoliev and Gaivoronski (1984), for theoretical background and further references see Ermoliev (1976). Here we shall describe only features relevant to the numerical experiments conducted with water resource models.

First of all it was necessary to put the problem in the form (18)–(19). Observe that the model with the joint probabilistic constraint cannot be easily put in the frame-work (18)–(19) while the expectation models (14), (15), (16) and (17) are already fc~mulated in the required fashion. But expectation models do not give the value of probability

$$P\{x_i \geq \beta_i + d_i, \; i = 2, 3, 4\}$$

which is important for decision making. Therefore we conducted numerical experiment in two stages.

1 Solve the problem based on the expectation model:

minimize

$$E_\beta f(x, \beta) = x_0 + cE \max \left\{0, \max_{i=2,3,4} \{\beta_i + d_i - x_i\}\right\} \tag{21}$$

subject to constraints

$$\sum_{i=1}^{k} x_i \leq z_k(1 - \alpha_k) + m_4 - m_k, \; k = 1, \ldots, 4$$

$$\sum_{i=0}^{k} x_i \geq z_k(\gamma_k) = m_4 + \nu_k, \; k = 1, \ldots, 4 \tag{22}$$

and additional constraints (5) (see (17)). Program STO generates certain approximation x^* to the optimal solution.

2 Evaluate the value of probability constraint, that is compute

$$p(x^*) = P\{x_i^* \geq \beta_i + d_i, \; i = 2, 3, 4\} \tag{23}$$

This was done using the generator of multivariate normal distribution from IMSL library which was also used on the first stage. If $p(x^*)$ appeared to be less than admissible level α then we increased the penalty coefficient c from (21) and solved the problem (21)–(22) again. This process was repeated until desirable value of $p(x^*)$ was reached. Clearly, $p(x^*) \to 1$ when $c \to \infty$.

Now we shall discuss strategies which were used for solving (21)–(22). In this case it is possible to compute subgradients $f_x(x^s, \beta^s)$ of function $f(x, \beta)$ for given values of $x = x^s$ and random parameters $\beta = \beta^s$, the components are

$$f_{x_0}(x^s, \beta^s) \equiv 1, \; f_{x_1}(x^s, \beta^s) \equiv 0 \;,$$

$$f_{x_i}(x^s, \beta^s) = \begin{cases} -c & \text{if } \beta_i^s + d_i - x_i^s = \\ & = \max\{0, \max_i \{\beta_i^s + d_i - x_i^s\}\} \\ 0 & \text{otherwise} \end{cases}$$

and take $\xi^s = f_x(x^s, \beta^s)$ in method (20). We decided, however, to use a method which is based only on values of $f(x^s, \beta^s)$ and is applicable therefore for more complex models. Various types of random search techniques and finite differences are implemented in STO. For this model the following analog of random search was used:

- Generate l random vectors $\eta^{s1}, \ldots, \eta^{sl}$: $\eta^{si} = (\eta_0^{si}, \ldots, \eta_4^{si})$ where η_j^{si} are uniformly and independently distributed on $[-\delta_s, \delta_s]$.
- Generate l independent random vectors $\beta^{s1}, \ldots, \beta^{sl}$, $\beta^{si} = (\beta_2^{si}, \beta_3^{si}, \beta_4^{si})$ with multivariate normal distribution.
- Compute ξ^s:

$$\xi^s = \sum_{i=1}^{l} \frac{f(x^s + \eta^{si}, \beta^{si}) - f(x^s, \beta^{si})}{\|\eta^{si}\|} \eta^{si}$$

  During computations we took $1 \leq l \leq 5$.

- Select ρ_s and perform one step in (20).

The value of search step was taken constant $\delta_s = 10$. Stepsize ρ_s was updated interactively, but after several trials the following pattern emerged, which was repeated afterwards:

$$\rho_s = \begin{cases} 10 & 1 \leq s \leq 50 \\ 1 \div 3 & 50 \leq s \leq 150 \\ 0.2 \div 1 & \text{afterwards} \end{cases}$$

In order to get exact solution it is necessary to take $\delta_s \to 0$. But in this particular case it appeared that sufficiently good approximation was obtained even with fixed δ_s. Usually it took 200–300 iterations to get solution. The results obtained by this method for solving water resource problems with different parameters are discussed from the application point of view in the section 6. One of the sets of problem parameters was used to compare performance of various models and solution techniques described at the beginning of this section.

5. COMPARISON OF THE SOLUTION TECHNIQUES

This is the particular problem of the type (17) which was used for comparison:

minimize $F(x) = E_\omega f(x, \omega) = x_0$

$$+ c E_\omega \left\{ \max \left\{ 0, \max_{1=2,3,4} \{ \omega_{i-1} + d_i - x_i \} \right\} \right\}$$

subject to constraints

$$\begin{aligned}
x_1 + x_2 &\le 156.448 \\
x_1 + x_2 + x_3 &\le 201.866 \\
x_1 + x_2 + x_3 + x_4 &\le 225.297 \\
x_0 + x_1 &\ge 512.886 \\
x_0 + x_1 + x_2 &\ge 592.872 \\
x_0 + x_1 + x_2 + x_3 &\ge 654.152 \\
x_0 + x_1 + x_2 + x_3 + x_4 &\ge 720.183
\end{aligned} \tag{24}$$

$100.0 \le x_0 \le 500.0$

$38.1 \le x_1 \le 102.319$

$00.0 \le x_2 \le 252.0$

$00.0 \le x_3 \le 252.0$

$00.0 \le x_4 \le 252.0$

where $d_i = 12.7$, $i = 2, 3, 4$. The random vector $\omega = (\omega_1, \omega_2, \omega_3)$ is distributed normally with

expectations $e = (20.2, 27.37, 10.65)$

standard deviations $q = (8.61, 10.65, 6.00)$

and correlation matrix $\begin{bmatrix} 1. & 0.360 & 0.125 \\ 0.360 & 1. & 0.571 \\ 0.125 & 0.571 & 1. \end{bmatrix}$,

penalty coefficient c was equal 100. This problem is the same as discussed in the case study in Section 6, only the upper bound $u_0 = 500$ mil. m^3 was used instead of $u_0 = 334$ mil. m^3 and the reliabilities $\gamma_k = 0.75$ were taken instead of $\gamma_k = 0.4$. The convenient feature of this problem is that we can easily obtain very good lower bound for solution by minimizing x_0 subject to constraints stated above. This gives $F(x^*) \ge 494.88$ where x^* is the optimal point.

5.1. *Approximations of stochastic problem by large scale linear programming problem*

To use this approach it is necessary to approximate initial continuous distribution by discrete distribution consisting of N points. Various ways of doing this are described in Birge, Wets (1986). Here we used the following two schemes:

1 "Intelligently"

- variance matrix was computed from the given correlation matrix and variances;
- eigenvalues r_i and eigenvectors v_i, $i = 1, 2, 3$ of the variance matrix were computed;
- number k was chosen and for one-dimensional normal distribution points b_i, $i = 1, \ldots, k$ were selected such that

$$P\{z \le b_1\} = P\{z \ge b_k\} = P\{b_i \le z \le b_{i+1}\} = \frac{1}{k+1}$$

for $i = 1, \ldots k - 1$

- approximating discrete distribution consisted of all points

$$g_{ijl} = b_i v_1 \sqrt{r_1 / s} = b_j v_2 \sqrt{r_2 / s} = b_l v_3 \sqrt{r_3 / s} + e$$

for all $i = 1, \ldots, k$, $j = 1, \ldots, k$, $1 = 1, \ldots, k$ and

$$s = \sum_{i=1}^{k} b_i^2$$

These points were taken with equal probability $1/(k^3)$. Approximating distribution chosen in this way is symmetric and has the same expectation and variance matrix as original distribution.

2 Just throwing specific number N of points distributed according to original distribution and assign to each of them probability $1/N$.

After such discretization is performed, the original problem becomes equivalent to the following stochastic optimization problem with complete recourse:

$$\text{minimize } x_0 = \frac{c}{N} \sum_{j=1}^{N} y_j$$

subject to constraints

$$x_i + y_j \geq \omega_{i-1}^j + d_i, \ i = 2, 3, 4, \ j = 1, \ldots, N$$

and constraints (24),

where $\omega^j \in R^3$, $j = 1, \ldots, N$, are the points where the discrete distribution is concentrated.

This is linear programming problem which could be of very high dimension because large N is needed for accurate approximation. We did not, however, make very accurate approximations and therefore were able to use a general purpose linear programming tool, namely linear programming part of MINOS 4.0, see Murtagh and Saunders (1983).

The results are summarized in the Table 1. To each entry in the left column of this table correspond two rows of numbers in the columns 2–4. The upper number corresponds to "intelligent" approximation and the lower to "just throwing". The column marked STO presents results obtained by the variant of SQG described in section 4 after 100 iterations, each iteration required 40 observations of random function. It is amazing that for much bigger problem of 1331 points in case of "intelligent" approximation MINOS found solution much faster than for 343 points. It is also worth noting what a big difference makes "intelligent" approximation as compared with "just throwing": approximation which uses only 125 points is better than one with 1331 points. The estimate of the value of the objective function in Table 1 was made using a sample of 10 000 points generated by subroutine for multivariate normal distribution from the IMSL library.

We complete the discussion by comparison of the empirical expectation and correlation matrix computed using the same 10 000 random numbers with the corresponding preasigned parameter values. It gives an idea how accurate the estimates are:

expectations of random vector ω e = (20.2000, 27.3700, 10.6500)

empirical expectations from 10 000 points (20.2855, 27.3912, 10.6841)

$$\text{correlation matrix of } \omega \begin{bmatrix} 1. & 0.360 & 0.125 \\ 0.360 & 1. & 0.571 \\ 0.125 & 0.571 & 1. \end{bmatrix}$$

$$\text{empirical correlation matrix} \begin{bmatrix} 1. & 0.354429 & 0.107631 \\ 0.354429 & 1. & 0.572717 \\ 0.107631 & 0.572717 & 1. \end{bmatrix}$$

TABLE 1

	$k = 5$	$k = 7$	$k = 11$	STO
Total number of points in approximating distribution	125	343	1331	
x_0	494.88	494.88	494.88	494.886
	494.88	494.88	494.88	
x_1	38.1	38.1	64.409	39.2441
	38.1	38.1	38.1	
x_2	88.019	84.918	54.752	71.4884
	118.35	114.939	92.088	
x_3	61.223	63.057	65.338	74.0301
	43.35901	42.421	61.811	
x_4	37.958	39.225	40.801	40.5345
	25.491	29.84	33.301	
Optimal value of the complete recourse objective function	494.88	494.88	494.88	
	494.88	494.88	494.88	
Estimate of the value of original objective function	505.204	501.401	499.854	495.604
	833.044	820.388	513.261	
Number of MINOS iterations	133	349	8	
	7	13	79	
MINOS user time	55.46	488.92	264.76	
	27.18	71.72	385.96	

5.2. *Stochastic quasigradient method*

In the STO implementation the user can choose between two options: interactive and automatic. In the interactive option the user chooses the stepsize ρ_s from (20) himself, his judgment is based on such notions as "oscillatory behavior" of variables or "visible trend" aided by some additional "measures of processs quality". These are displayed on the terminal along with current point and user can interrupt iterative processs and change the value of stepsize. In order to use this possibility effectively the user has to be quite experienced. However, as practice shows even an inexperienced user can quickly get necessary skills. All this is described in more detail in Gaivoronski (1986).

In experiments with interactive mode the step direction ξ^s from (20) was computed using certain analogue of random search: generate random vector $h = (h_1, h_2, h_3)$ where h_i are independent random variables uniformly distributed on $[-G, G]$ and G is chosen interactively from the interval [1, 20], compute $\|h\|$ compute vector $u = (u_1, u_2, u_3)$ such that

$$u = h(f(x^s + h, \omega_1) - f(x_s, \omega_0))/\|h\|$$

where ω_i are independent observations of ω

repeat the first two steps l times with different independent observations of random variables ω and h, obtain l vectors u and take as step directions ξ^s average of these vectors (in example given below $l = 10$).

Stepsize was chosen interactively. At first sufficiently large stepsize chosen (10.0) which was reduced if irregular behavior of the x variables was observed. The results are in Table 2.

TABLE 2

Step number	Stepsize	x_0	x_1	x_2	x_3	x_4	$F(x)$
The value of step G in random search was set 10.0							
2	10.000	497.966	54.924	56.572	57.126	56.675	522.663
4	10.000	495.456	54.017	56.136	57.751	56.822	517.106
6	10.000	500.000	45.750	110.698	28.866	34.869	1702.02
8	10.000	494.886	96.359	60.089	2.818	66.061	4222.28
10	10.000	500.000	89.514	63.600	1.039	71.145	4405.28
11	2.000	500.000	102.319	51.587	3.529	62.748	4156.32
12	2.000	497.288	95.581	0.	106.282	21.029	3796.11
14	2.000	500.000	91.400	37.497	18.361	42.926	754.445
16	2.000	500.000	38.100	54.772	108.994	19.945	948.331
18	2.000	494.886	39.740	64.965	84.687	35.905	498.726
20	2.000	494.886	38.100	67.452	78.986	40.759	492.378
22	2.000	494.886	38.629	67.744	78.421	40.503	495.428
24	2.000	494.886	38.431	67.806	78.528	40.532	495.420
26	0.200	497.901	39.045	66.946	75.445	43.491	498.291
28	0.200	497.767	39.001	66.934	75.449	43.477	498.133
30	0.200	497.606	39.040	66.956	75.478	43.490	497.490
At this point the step in random search was changed to 5.0							
31	0.200	497.569	39.067	66.981	75.472	43.474	497.879
35	0.200	497.032	39.240	67.186	75.473	43.398	497.353
40	0.200	496.698	39.242	67.206	75.422	43.425	497.042
50	0.200	496.046	39.252	67.200	75.354	43.429	496.358
60	0.200	495.400	39.224	67.255	75.324	43.358	495.802
70	0.200	495.524	39.086	71.344	73.790	40.440	496.192
80	0.200	495.056	39.180	71.461	74.000	40.518	495.664
90	0.200	494.886	39.168	71.501	74.037	40.591	495.595
100	0.200	494.886	39.244	71.488	74.030	40.535	495.604

During actual computations information was displayed more frequently and, of course, without last column which contains the estimates of the values of the objective function $f(x)$ at the current points using the same 10000 observations of the random vector ω which were used for estimations of MINOS results. These estimates were obtained afterwards (and consumed a lot of computer time!). From the table it is evident that very good solution (better than 1331 points approximate scheme) was obtained after 20 iterations, that is after 800 function evaluations and all subsequent points oscillated in close vicinity.

The major disadvantage of the interactive option is that it requires too much from the user. Therefore automatic option was developed in which computer simulates behavior of an experienced user. For experiment with automatic option we choose the simplest version of SQG: one which uses for step direction the values of $f_x(x, \omega)$, which are very easy to compute here:

$$f_x(x, \omega) = \{1., 0., -ct_1, -ct_2, -ct_3\}$$

where $t_j = 1$ if

$$\omega_j + d_{j+1} - x_{j+1} = \max\{0, \max_{i=2,3,4}\{\omega_{i-1} + d_i - x_i\}\}$$

and $t_j = 0$ otherwise, $c = 100$ – penalty coefficient.

On each iteration only one observation of $f_x(x^s, \omega)$ was computed, which was used as step direction.

Stepsize ρ_s was computed automatically according to the simple rule:

- on each iteration one observation $f(x^s, \omega^s)$ was made and these observations were used to compute estimate $F(s)$ of the current value of the objective function:

  $$F(s) = \sum_{i=1}^{s} f(x^i, \omega^i)/s$$

  current path length $L(s)$ was computed also:

  $$L(s) = \sum_{i=1}^{s} \|x^{i+1} - x^i\|$$

- initial stepsize ρ_1 was chosen sufficiently large (in examples below it was 5.0) and each M iterations the condition for reducing stepsize was checked (in examples below $M = 20$, that is conditions were checked on iterations number 20, 40, 60,...). This condition is the following:

  $$\rho_{s+1} = D\rho_s \text{ if } (F(s-K) - F(s))/(L(s) - L(S-K)) \le A$$

  $$\rho_{s+1} = \rho_s \text{ otherwise .}$$

In examples below $D = 0.5$, $K = 20$, $A = 0.01$, starting point was (1000, 100, 100, 100, 100). Each iteration required one observation of random function $f(x, \omega)$ and one observation of its gradient. Below there are two runs with different sequences of random vectors ω. One is "very good" another – "not as good but quite reasonable". When current point approaches optimum the event $\{t_j = 1\}$ becomes less and less likely. Therefore the method spends much of iterations standing at the same point (for instance iterations 260–400 in Tables 3 and 4). The last column again was obtained afterwards using the same 10 000 observations of w as in the previous tests. It shows that the algorithm reaches quite good vicinity of solution after 120 iterations (except a slight jump on iteration 420). In Table 4 we have a big jump on iteration number 220, the method, however, quickly reaches the vicinity of solution again. This jump is due to too big stepsize. After iteration number 240 everything is OK again.

The problem with SQG is the stopping criterion and currently the criterion implemented is based on assumption that if stepsize becomes too small we are in the vicinity of optimum. This of course is not necessarily true but experience shows that nevertheless this criterion is quite reliable if coupled with repeated runs.

The same input data were used for solving the problem (11) with joint probability constraint. New variables $\tilde{x}_0 = x_0 - l_0$ and $\tilde{x}_1 = x_1 - d_1$ were introduced to eliminate the individual lower bounds. For the resulting program

minimize $\tilde{x}_0$

$$\text{subject to } P\begin{bmatrix} x_2 & \le \omega_1 + d_2 \\ x_3 & \ge \omega_2 + d_3 \\ x_4 & \ge \omega_3 + d_4 \end{bmatrix} \ge \alpha$$

$$\begin{aligned}
\tilde{x}_1 &\le z_1(1 - \alpha_1) + m_4 - m_1 - d_1 \\
\tilde{x}_1 + x_2 &\le z_2(1 - \alpha_2) + m_4 - m_2 - d_1 \\
\tilde{x}_1 + x_2 + x_3 &\le z_3(1 - \alpha_3) + m_4 - m_3 - d_1 \\
\tilde{x}_1 + x_2 + x_3 + x_4 &\le z_4(1 - \alpha_4) - d_1 \\
\tilde{x}_0 + \tilde{x}_1 &\ge z_1(\gamma_1) + m_4 - v_1 - l_0 - d_1 \\
\tilde{x}_0 + \tilde{x}_1 + x_2 &\ge z_2(\gamma_2) + m_4 + v_2 - l_0 - d_1 \\
\tilde{x}_0 + \tilde{x}_1 + x_2 + x_3 &\ge z_3(\gamma_3) + m_4 + v_3 - l_0 - d_1 \\
\tilde{x}_0 + \tilde{x}_1 + x_2 + x_3 + x_4 &\ge z_4(\gamma_4) + m_4 + \nu_4 - l_0 - d_1
\end{aligned}$$

$$0 \le \tilde{x}_0 \le u_0 - l_0;\ 0 \le \tilde{x}_1 \le u_1 - d_1;\ 0 \le x_2 \le u_2;\ 0 \le x_3 \le u_3;\ 0 \le x_4 \le u_4 \ ,$$

two solution techniques were implemented:

TABLE 3 "Very good".

Step number	Stepsize	x_0	x_1	x_2	x_3	x_4	$F(x)$
20	5.000	494.886	56.324	56.324	56.324	56.324	524.380
40	5.000	494.886	40.657	57.329	61.280	66.031	504.566
60	2.500	494.886	40.657	57.329	61.280	66.031	504.566
80	1.250	494.886	40.657	57.329	61.280	66.031	504.566
100	0.625	498.438	38.881	55.553	94.306	36.557	502.248
120	0.625	493.189	39.662	57.021	87.596	41.018	497.002
140	0.313	494.886	40.314	57.672	86.945	40.366	495.706
200	0.313	494.886	40.314	59.672	86.945	40.366	792.706
220	0.313	497.969	39.116	55.788	72.861	57.533	499.343
140	0.313	497.344	39.428	56.100	88.173	41.595	198.337
260	0.313	494.886	40.657	57.329	86.912	40.366	495.782
400	0.313	494.886	40.657	57.329	86.945	40.366	495.782
420	0.313	499.844	38.178	54.850	89.423	42.845	501.282
440	0.156	498.281	38.959	55.631	88.642	42.064	499.416
460	0.156	496.719	39.741	56.413	87.861	41.283	497.641
480	0.156	495.156	40.522	57.194	87.080	40.502	496.001
500	0.156	494.886	40.657	57.329	86.945	40.366	495.782
520	0.156	494.886	40.657	57.329	86.945	40.366	492.782
540	0.156	494.886	38.100	68.598	82.589	36.010	498.586
560	0.156	494.886	38.100	63.390	77.380	46.427	495.158
1000	0.156	494.886	38.100	63.390	77.380	46.427	495.158

TABLE 4 "Not as good as previous but quite reasonable".

Step number	Stepsize	x_0	x_1	x_2	x_3	x_4	$F(x)$
20	5.000	494.886	56.324	56.324	56.324	56.324	524.380
40	2.500	494.886	48.993	48.993	61.280	66.031	513.800
60	2.500	496.000	40.100	56.772	61.280	67.145	505.757
80	1.250	497.000	39.600	56.272	61.280	68.145	506.897
100	1.250	494.886	40.657	57.329	61.280	66.031	504.566
200	1.250	494.886	40.657	57.329	61.280	66.031	504.566
220	1.250	499.500	38.350	55.022	108.494	23.431	736.529
240	0.625	494.886	40.657	57.329	74.812	52.499	495.735
420	0.625	494.886	40.657	57.329	74.812	52.499	495.735
440	0.625	496.875	39.662	56.335	75.806	53.494	497.946
460	0.313	494.886	40.657	57.329	74.812	52.499	495.735
940	0.156	494.886	40.657	57.329	74.812	52.499	495.735
1000	0.156	494.886	40.657	57.329	74.812	52.499	495.735

5.3. *The PCSP code by Szántai*

The program solves problems of stochastic programming with joint probability constraints under assumption of multinormal distribution of random right-hand sides ($\omega_{i-1} = \beta_i$, $i = 2, 3, 4$ in our case). It is based on Veinott's supporting hyperplane algorithm (see Veinott (1967)). The individual upper bounds on variables are handled separately and the parameters of the multinormal distribution are used to get a starting feasible interior solution. For constructing the necessary linear and stochastic data files, one can turn to the brief documentation by Edwards (1985). Some computational results of the test problem with data given at the beginning of this Section are given in Table 5.

TABLE 5 Results of the calculations by using the PCSP code.

x_0	x_1	x_2	x_3	x_4	Prob. lev.	CPU time	No. of cutting planes
494.880	49.982	58.302	81.970	35.046	0.973	25.52	3
494.880	42.980	55.822	63.614	62.935	0.983	36.38	5
494.880	41.494	59.400	63.155	61.251	0.981	32.37	4
494.880	41.307	69.226	74.595	40.172	0.997	41.87	5
494.880	39.776	66.921	76.140	42.463	0.999	34.45	5

5.4. *The application of the nonlinear version of the MINOS system*

For this purpose one has to write a separate subroutine named CALCON which calculates the value of the probabilistic constraints and its gradient. The subroutines are contained in the PCSP code. They were coded on the base of an improved simulation technique by Szántai (1985). Computational results given in this section show that the direct application of the MINOS system for the solution of the optimization problem (11) is less comfortable than the use of the PCSP code. We obtained that without giving a good initial setting on the values of the nonlinear variables the MINOS system failed to find a feasible solution.

Finally in Table 7 there is the summary of experiments.

In the first column there is a short description of experiment including name of the program, number of approximating points and type of approximation (for MINOS, *I* and *R* mean the same as before), number of iterations and indication of interactive or automatic mode (for STO), value of probability constraint (for PCSP).

TABLE 6 Results of the calculations by using the nonlinear MINOS system directly. (We applied the x_2 = 63.035, x_3 =77.345, x_4 = 30.0 initial settings. For these initial values the probabilistic constraint is infeasible with value 0.866.)

x_0	x_1	x_2	x_3	x_4	Prob. lev.	CPU time	No. of major iterations
494.886	51.221	63.067	77.345	33.694	0.973	46.46	11
494.886	49.997	63.067	77.395	34.918	0.983	56.22	14
494.886	49.910	63.067	77.345	35.005	0.984	58.06	14
494.886	46.169	63.067	77.395	38.746	0.997	84.48	23
494.886	43.414	63.037	77.345	41.501	0.999	112.58	31

In the column 7 there are values of objective function (24) computed by averaging 10 000 observations generated by IMSL subroutine for multivated normal distribution. In the column 8 are the values of probability constraint computed using the same random numbers and in the last column cpu time of the VAX 780. For interactive mode of STO cpu time is not included because the crucial factor there was user's response.

Experiments suggest that both considered models give comparable results and that the methods STO and PCSP contained in the ADO/SDS library of stochastic programming codes performed better on this particular problem than the direct use of standard LP or NLP packages. To solve the case study, the stochastic quasigradient method was chosen as its implementation is not essentially limited by the assumed type of distribution and through increasing the penalty coefficient value an approximate optimal solution which fulfills the joint probability constraint can be achieved.

6. THE CASE STUDY

In the case study of the water resources system in the Bodrog River basin the model (17) was used with added constraints (see discussion):

$$d_i \le x_i \quad i = 2, 3, 4 \ . \tag{25}$$

The input values of the 3-dimensional multinormal distribution of β_i were as follows (with exception of correlation matrix in mil. m^3)

TABLE 7

Experiment		x_0	x_1	x_2	x_3	x_4	Function value	Prob. value	CPU time
MINOS	125_I	494.88	38.1	88.019	61.223	39.958	505.204	0.9728	55.46
MINOS	125_R	494.88	38.1	118.35	43.359	25.491	833.044	0.4878	27.18
MINOS	343_I	494.88	38.1	84.918	63.057	39.225	501.401	0.9832	488.92
MINOS	343_R	494.88	38.1	114.939	42.421	29.84	820.388	0.5578	71.72
MINOS	1331_I	494.88	64.409	54.752	65.338	40.801	499.854	0.9844	264.76
MINOS	1331_R	494.88	38.1	92.088	61.811	33.301	513.261	0.9391	385.96
STO I	100	494.886	39.2441	71.4884	74.0301	40.5345	495.604	0.9978	
STO I	24	494.886	38.431	67.806	78.528	40.532	495.420	0.9981	
STO A1	1000	494.886	38.100	63.390	77.380	46.427	495.158	0.9994	51.52
STO A2	1000	494.886	40.657	57.329	74.812	52.499	495.735	0.9972	49.38
PCSP	0.75	494.882	38.234	106.391	51.945	28.729	605.040	0.751	
PCSP	0.973	494.880	49.982	58.302	81.970	35.046	500.890	0.9713	25.52
PCSP	0.983	494.880	42.930	55.822	63.614	62.935	501.199	0.9827	36.38
PCSP	0.984	494.880	41.494	59.400	63.155	61.251	501.530	0.9849	32.37
PCSP	0.989	495.375	43.281	58.230	86.255	37.534	497.382	0.9894	
PCSP	0.997	494.880	41.307	69.226	74.595	40.172	495.668	0.9972	41.87
PCSP	0.999	494.880	39.776	66.921	76.140	42.463	495.172	0.9989	34.45

Period	$E(\beta_i)$	$\sigma(\beta_i)$	Correlation matrix		
2	20.2	8.61	1		
3	27.37	10.65	0.360	1	
4	10.65	0.00	0.125	0.571	1

The parameters of the marginal normal distributions of cumulated monthly inflows ζ_i were

Period	$E(\zeta_i)$	$\sigma(\zeta_i)$
1	303.47	122.28
2	375.94	133.43
3	432.61	140.27
4	486.26	158.64

As the values $\alpha_k = 0.g$, $k = 1, 2, 3, 4$ and $\gamma_k = 0.4$, $k = 1, 2, 3, 4$ were used (see discussion) the following values of quantiles z_k were obtained (in mil. m^3

Period k	1	2	3	4
$z_k(1 - 0.9)$	146.8	205.0	252.9	283.0
$z_k(0.4)$	272.5	342.1	397.1	446.1

The values d_i were as follows (mil. m^3)

Period i	1	2	3	4
d_i	38.1	12.7	12.7	12.7

The minimum and maximum reservoir capacity x_0 was: l_0 = 100 mil. m^3 and u_0 = 334 mil. m^3. (In the alternative discussed in the previous section a less realistic value u_0 = 500 mil. m^3 was used.) The upper bounds for variables were 252 mil. m^3 (u_i = 252, i = 1, 2, 3, 4 in mil. m^3) that is the volume of a long-term flood. This constraint was not effective and therefore it was not analysed.

Due to recreation purposes, the acceptable minimum storage in the 3-rd period is m_3 = 194 mil. m^3. However, the comparison of the third inequality of (9), the third inequality of (10) together with $x_0 \leq 334$ gives an upperbound of 189,4 for the sum $m_3 + v_3$, so that the parameter value m_3 = 194 would lead to contradictory constraints. Thats why the minimum storage value m_3 has been put up to 137 mil. m^3 (see alternative C) and the storage values m_k, k = 1, 2, 3, 4 have been kept fixed over all periods (see alternatives A and B). The reliability of maintaining the summer reservoir pool has been evaluated ex post.

6.1. *Choice of reliability values*

The very important parameters of the model are the required probabilities α, γ_i and α_i. The value α is the required joint probability of the water supply. The tests with the model have shown that it is necessary to add the deterministic constraint (25) in order to secure the required values of constant industrial water demands. The penalty term is often so weak that the constraint (25) may be violated. Using the deterministic constraint (25) a relatively low value α, e.g. α = 0.85 may be acceptable. However, the value of joint probability (1) is the output of the model; therefore the condition (1) is tested and if it cannot be fulfilled, the alternative is rejected.

The values α_i refer to the relatively strong environmental and technical requirements for maintaining the minimum reservoir pool. Therefore α_i = 0.9, i = 1, 2, 3, 4 was chosen.

The choice of the values γ_i was rather difficult. They refer to the important constraints imposed on the reservoir V operation that arise from flood control requirements that stipulate that a certain space – flood control storage, be held empty. This requirement cannot be easily expressed in the model due to its aggregated character. The probability that the freeboard storage is empty means also that there is no spill during this period. If the period is short, e.g. one day, the probability γ may correspond to the required reliability. For longer periods e.g. one month, two months, half a year, extremely large storage capacities will be necessary if no spill shall take place during this period with probability equal 0.75. (This value has been used in the previous section and the total reservoir capacity x_0 as high as approx. 500 mil. m^3 was necessary.) Therefore the flood control problems are often treated in a separate model and the required probabilities γ are adapted to the resulting values of this separate model. It is necessary to interpret properly the meaning of these probabilities. For instance the probability $\gamma = 0.4$ of maintaining the freeboard storage does not say that the flood control is on a low level but that the freeboard space will be filled (and possibly some spill may occur) with this probability during the period chosen. With this fact in mind and according to the results of the separate flood control model the value $\gamma_i = 0.4$ $i = 1, 2, 3, 4$ was chosen in this section.

6.2. *Results*

In the first resulting alternative A of the reservoir V design and operation the input parameters were:

$m_k = 57$ mil. m^3, $k = 1, 2, 3, 4$ – minimum reservoir storage,

$v_k = 70$ mil. m^3, $k = 1, 2, 3, 4$ – flood control storage (freeboard storage),

which gave

$x_0 = 291.6$ mil. m^3 – the total reservoir capacity

$x_1 = 107.9$ mil. m^3 – the total release in the 1-st period

$x_2 = 69.6$ mil. m^3 – the total release in the 2-nd period

$x_3 = 69.8$ mil. m^3 – the total release in the 3-rd period

$x_4 = 35.7$ mil. m^3 – the total release in the 4-th period.

Started at $s_0 = m_4 = 57$ mil. m^3, using the computed optimal releases and the cumulated monthly inflows equal to the quantiles z_k (1 – 0.9) and z_k (0.4) we can compute the corresponding storages s_i (in mil. m^3):

Period i	0	1	2	3	4
min storage $\underline{s}_i$	57	95.9	84.5	62.6	$\underline{57.0}$
max storage $\bar{s}_i$	57	$\underline{221.6}$	$\underline{221.6}$	206.8	219.9

The underlined values are critical, i.e., the corresponding inequalities in (9) and (10) are fulfilled as equations. As the existing total reservoir capacity is 304 mil. m^3, we observe from the active constraints

$$x_0 = \bar{s}_1 + v_1 \quad (291.6 = 221.6 + 70)$$

$$x_0 = \bar{s}_2 + v_2$$

that the freeboard storage could be increased from 70 mil. m^3 to 82.4 mil m^3. The computed joint reliability of water supply is $\alpha \geq 0.979$.

In this alternative A the requirements for maintaining the reservoir pool during the third period were not met with probability high enough. The acceptable recreation storage in the third period is s_3 = 194 mil. m^3. Using the relation

$$s_3 = m_4 + \zeta_3 - x_1 - x_2 - x_3 \ ,$$

the corresponding value of ξ_3 can be obtained:

$$\zeta_3 = s_3 - m_4 + x_1 + x_2 + x_3 = 194 - 57 + 247.3 = 384.3 \ .$$

For values $\zeta_3 \geq 384.3$, the condition on recreation storage will be fulfilled. Using the parameters of the marginal normal distribution of ζ_3 we get the corresponding reliability $\hat{\alpha}_3$:

$$\hat{\alpha}_3 = P\{\zeta_3 \geq 384.3\} = P\left\{\frac{\zeta_3 - 432.6}{140.3} \geq 0.34\right\} = \phi(0.34) = 0.633 \ ,$$

where ϕ denotes the distribution function of the $N(0, 1)$ distribution. The resulting reliability of recreation pool is not acceptable.

In the second resulting alternative B of the reservoir V design and operation, the input parameters were changed to m_4 = 131 mil. m^3 – the minimum reservoir storage including the environmental and recreation goals (m_k = 131 for k = 1, 2, 3, 4) v_k = 10, k = 1, 2, 3, 4 the freeboard storage. The resulting solution (in mil. m^3)

$$x_0 = 304.1$$

$$x_1 = 109.4$$

$$x_2 = 69.6$$

$$x_3 = 65.1$$

$$x_4 = 38.9$$

does not differ substantially from that in alternative A. The minimum and maximum reservoir storage volumes (in mil. m^3) were

Period i	0	1	2	3	4
min storage $\underline{s}_i$	131	168.4	157.0	139.8	$\underline{131.0}$
max storage $\bar{s}_i$	131	$\underline{294.1}$	$\underline{294.1}$	282.8	$\underline{294.1}$

where the critical values are underlined again. The computed joint reliability of water supply is $\alpha \geq 0.987$.

For values $\zeta_3 \geq 194 - 131 + 244.1 = 307.1$, the condition on acceptable recreation storage will be fulfilled; the corresponding reliability

$$\hat{\alpha}_3 = P\{\zeta_3 \geq 307.1\} = \phi(0.894) = 0.814$$

of recreation pool is high enough.

In the last alternative C which is a slight modification of the alternative A the input parameters were:

$$m_k = 57 \text{ mil.m}^3,\ k = 1, 2, 4,\ m_3 = 137 \text{ mil.m}^3 \ ,$$

$$v_k = 70 \text{ mil.m}^3,\ k = 1, 2, 4,\ v_3 = 10 \text{ mil.m}^3 \ .$$

The resulting solution (in mil. m^3)

$$x_0 = 334.00$$

$$x_1 = 67.55$$

$$x_2 = 67.55$$

$$x_3 = 37.80$$

$$x_4 = 110.10$$

differs substantially from the previous ones. The minimum and maximum storage volumes (in mil. m^3) were

Period i	0	1	2	3	4
min storage $\underline{s}i$	57	136.2	126.9	$\underline{137.0}$	$\underline{57.0}$
max storage $\bar{s}i$	57	261	$\underline{264.0}$	281.2	220.1

where the critical values are underlined again. Observe that the upper bound of the reservoir capacity has been reached (so that the existing reservoir capacity has been surpassed). At the same time, the computed joint reliability of water supply is only $\alpha \geq 0.412$, which is not acceptable.

For values $\zeta_3 \geq 194 - 57 + 172.9 = 309.9$, the condition on acceptable recreation storage will be fulfilled. The corresponding reliability

$$\hat{\alpha}_3 = P\{\zeta_3 \geq 309.9\} = \phi(0.874) = 0.809$$

is sufficiently high again.

The comparison of results is given in Table 8.

7. DISCUSSION

The comparison of the goals of the water resources system with the results obtained shows that all the target values cannot be achieved by the operation of the total reservoir storage 334 mil. m^3 of the reservoir V. The relationships between the goals and possibilities of the reservoir V were as follows:

1 The goal of maintaining the minimum reservoir storage for environmental conservation and technological purposes is met in all alternatives.

2 The flood control goal is met in the first alternative in cooperation with the levels along the river and it is mostly met in the third alternative too. In the second alternative the flood control storage is too small and cannot be accepted by the decision makers.

3 The minimum storage requirement in case of recreation pool cannot be met fully as the substantially increased values of m_3 give an empty solution set. This fact can be easily interpreted in water resources system analysis as too much water is released in the first and second period and the inflow during the third period is not sufficient. An ex post analysis shows that the solution of the second and third alternative gives the desired recreation pool with probability greater than 0.8 which may be acceptable.

TABLE 8

Alternative	A	B	C	D	E
Model/Code	(17)/STO	(17)/STO	(17)/STO	(17)/STO	(11)/PCSP
Parameters	$m_k = 57\ \forall k$ $v_k = 70\ \forall k$ $u_0 = 334$	$m_k = 131\ \forall k$ $v_k = 10\ \forall k$ $u_0 = 334$	$m_k = 57,\ k = 1, 2, 4,\ m_3 = 137$ $v_k = 70,\ k = 1, 2, 4,\ v_3 = 10$ $u_0 = 334$	$m_k = 57\ \forall k$ $v_k = 70\ \forall k$ $u_0 = 500$	$m_k = 57\ \forall k$ $v_k = 70\ \forall k$ $u_0 = 500$
Optimal solution					
x_0	291.6	304.1	334.0	494.9	494.9
x_1	107.9	109.4	67.55	40.7	39.8
x_2	69.6	69.6	67.55	57.3	66.9
x_3	69.8	65.1	37.8	74.8	76.1
x_4	35.7	38.9	110.10	52.5	42.5
Reliabilities					
α (releases)	0.979	0.987	0.412	0.99	0.99
$\hat{\alpha}_3$ (recreation)	0.633	0.814	0.809	0.81	0.79
γ (freeboard)	0.4	0.4	0.4	0.75	0.75
Goals (met)					
min storage	YES (0.9)	YES (0.9)	YES (0.9)	YES (0.9)	YES (0.9)
recreation pool	NO	YES (0.8)	YES (0.8)	YES (0.8)	YES (0.79)
freeboard volume	YES	NO	YES	YES	YES
water supply	YES	YES	NO	YES	YES

4 The joint probability constraint for water supply was fulfilled in the first two alternatives on a surprisingly high level. To satisfy the fixed demand requirements, deterministic constraints (25) were added. In the alternative *C*, the joint probability constraint was not fulfilled because of the low release in the third period.

5 The comparison of the three alternatives *A*, *B*, *C* shows that the requirements for flood control, maintaining the recreation pool and the water supply are antagonistic and cannot be met in the water resource system by the reservoir *V* only with the realistic total reservoir storage 334 mil. m^3.

6 Alternatives *D*, *E* that have used the higher reliability of the freeboard storage for long periods, $\gamma = 0.75$, required the non-realistic total reservoir capacity approx. 500 mil. m^3. As this total reservoir capacity cannot be reached, the construction of the reservoir *K* will be necessary.

8. CONCLUSION

Stochastic programming models were used for identification of those management plans of water resources system development which best meet the required objectives. For this purpose the economic objective and physical environmental and economic probability constraints were expressed mathematically. These management plans involved the choice of the design and operation variables, i.e. the total reservoir capacity and the releases in the investigated periods. The analysis of the design alternatives shows the contradictory character of the main goals of the water resources system – water supply for industry and irrigation flood control, environmental conservation and recreation. As the optimum alternatives do not meet all these goals, the water resources system has to be enlarged by the reservoir *K*. As an screening aggregate model was used and the multidimensional distribution and marginal distributions were approximated by the multinormal and normal distributions respectively, the optimum design and operation variables derived by this model are rough approximations. However, the more precise values that could be derived using a more sophisticated model cannot differ to such degree that the main result (i.e. the necessity to plan a new reservoir) be altered.

The comparison of the different models of stochastic programming has proved a good agreement of the technological results. The method of multi modelling (i.e. the use of several models and programs for solution of the same problem) proved to be of use in planning of water resources systems development.

REFERENCES

Birge, J. and R.J-B. Wets (1986): Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse, WP-83-111, IIASA, Laxenburg, Austria (1983) and in A. Prékopa and R.J-B. Wets eds., *Mathematical Programming Study* **27**.

Dupačová , J. (1980): Water resources system modelling using stochastic programming with recourse. In: Kall, P. and A. Prékopa (eds.), Recent results in stochastic programming (Proc. Oberwolfach 1979). Lecture Notes in Economics and Math. Systems 179, 121–133. Springer, Berlin.

Dupačová, J. and Z. Kos (1979): Chance – constrained and simulation models of water resources systems. *Ekonomicko – matematicky' obzor* **15**, 178–191.

Edwards, J. ed. (1985): Documentation for the ADO/SDS collection of stochastic programming codes. WP-85-02, IIASA, Laxenburg.

Ermoliev, Yu. (1976): Methods of stochastic programming (in Russian). Nauka, Moskva.

Ermoliev, Yu. and A. Gaivoronski (1984): Stochastic quasigradient methods and their implementation. WP-84-55, IIASA, Laxenburg.

Gaivoronski, A. (1986): Stochastic quasigradient methods and their implementation. In: Ermoliev, Yu. and R. Wets eds. (1986) Numerical techniques for stochastic optimization problems. Springer, Berlin.

Gill, P.E., W. Murray, M.A. Saunders and M.H. Wright (1983): User's guide for SOL/QPSOL: A FORTRAN package for quadratic programming. SOL 83-7, Systems Optimization Laboratory, Stanford University.

Kall, P. (1979): computational methods for solving two-stage stochastic linear programs. *Z. Agnew.-Math. Phys.* **30** 261–271.

Kall, P. and D. Stoyan (1982): Solving stochastic programming problems with recourse including error bounds, *Math. Operations Forsch. Statist. Ser. Optimization* **13** 431–447.

Kos, Z. (1979): Operation of water resource systems in Czechoslovakia. In: The operation of multiple reservoir systems, Kaczmarek, S. – Kindler J. ed. CP-82-S3, IIASA, Laxenburg, p. 181–204.

Kos, Z. and V. Zeman (1979): The Odra River water resource system: a case study. In: The operation of multiple reservoir systems. Kaczmarek, Z. – Kindler J. ed., CP-82-S3, IIASA, Laxenburg, p. 363–381.

Kos, Z. (1982): Stochastic water requirements for supplementary irrigation in water resources systems. RR-81-34, IIASA, Laxenburg.

Loucks, D.P., J.R. Stedinger and D.A. Haith (1981): Water resource systems Planning and analysis, Prentice Hall, Inc. Englewood Cliffs, New Jersey.

Murtagh, B.A. and M.A. Saunders (1983): "MINOS 5.0 User's Guide", Report No. SOL 83-20, Systems Optimization Laboratory, Stanford University.

Prékopa, A. (1971): "Logrithmic concave measures with application to stochastic programming", *Acta Scientiarum Mathematicarum* **32**, 301–316.

Prékopa, A. (1973): Stochastic programming model for inventory control and water storage problems. In: Proceedings of the International Conference on Inventory Control and Water Storage held in Györ, Hungary, 1971, ed. A. Prékopa, Colloquie Mathematica Societatis János Bolyai 7, North Holland Publishing Company, 229–245.

Prékopa, A. ed. (1978): Studies in applied stochastic programming. Hungarian Academy of Sciences, MTA SzTAKI, Budapest.

Prékopa, A. (1986): Numerical solution of probabilistic constrained programming problems. In: Ermoliev, Yu. and R. Wets eds. (1986) Numerical techniques for stochastic optimization problems. Springer, Berlin.

Prékopa, A., T. Rapcsak and J. Zsuffa (1978): Senrially linked reservoir system design using stochastic programming. *Water Resources Research* **14**, 672–678.

Prékopa, A. and T. Szántai (1978): Flood control reservoir system design using stochastic programming. *Mathematical Programming Study* **9**, 138–151.

ReVelle, Ch., E. Joeres and W. Kirby (1969): The linear decision rule in reservoir management and design. *Wat. Res. Res.* **5**, 767–777.

Rockafellar, R.T. and R.J-B. Wets (1986): A Lagrangain finite generation technique for solving linear quadratic problems in stochastic programming. *Mathematical Programming Study* **28**, 63–93.

Szántai,T. (1986): Evaluation of a special multivariate gama distribution function. *Mathematical Programming Study* **27**, 1–16.

Veinott, A.F. (1967): Supporting hyperplane method. *Oper. Res.* **15**, 147–152.

Van Slyke, R. and R.J-B. Wets (1969): "L-shaped linear program with applications to optimal control and stochastic linear programs", *SIAM Journal on Applied Mathematics* **17** 638–663.

Wets, R. (1983): "Stochastic programming: solution techniques and approximation schemes", in: A. Bachem, M. Groetschel and B. Korte, eds., *Mathematical Programming: The State-of-the-Art*, 566–603, Springer-Verlag, Berlin.