# Secondary Stochastic Processes and Storage Reservoir Optimization

#### András Prékopa, Tamás Szántai, István Zsuffa

Department of Differential Equations, Budapest University of Technology and Economics, Műegyetem rkp. 3-9, 1111 Budapest, Hungary e-mail: szantai@math.bme.hu

Abstract: In the first two sections of the paper, stream flow is investigated on a probability theoretical basis. We will show that under some realistic conditions its probability distribution is of gamma type. In the model of the third section the optimal capacity of a storage reservoir is determined. In the model of the fourth section optimal water release policy is sought, given that water demands should be met by a prescribed large probability. Finally, in the last fifth section, in addition to the before mentioned reliability type constraint an upper bound is imposed on the number of days when demands may not be met and the cost of the intake facility is to be minimized<sup>1</sup>.

Keywords: reservoir capacity; release policy; stochastic programming

# 1 Secondary Stochastic Processes Derived by a Poisson Process

The use of Poisson type stochastic processes is frequent in hydrology. Presently, we assume that the sequence of rainfall events follows a Poisson process. That is, if  $\xi(I)$  denotes the (random) number of rainfalls in a time interval *I*, then

- a) for all  $I_1, K, I_n$  interval systems, where any two intervals have no common inner points, the random variables  $\xi(I_1), K, \xi(I_n)$  are independent,
- b)  $\xi(I)$  has Poisson distribution with parameter  $\lambda(I) \ge 0$ .

<sup>&</sup>lt;sup>1</sup> The problem of finding storage reservoir capacity was formulated by István Zsuffa many years ago. The detailed elaboration of the problem is more recent and is due to the first two authors who offered the Hungarian version of this paper appeared in Alkalmazott Matematikai Lapok **27** (2010) 175-188 to the memory of their friend and co-worker, István Zsuffa. The first author many years ago planned to publish the paper in English, too. After András Prékopa passed away last year, this task remained to the second author, who offers this paper to the memorial volume of Acta Polytechnica Hungarica.

A secondary process derived by a Poisson process means that to the random events of the Poisson process, in our case to the time points of rainfalls, a random secondary phenomenon is ordered, which is now a random flood wave. Let denote the random field of the secondary events Y. On this random field more probability measures are defined. For discussing secondary processes an appropriate tool is the so called product space method, see [3]. This consists of regarding the secondary process in the set of the element pairs (t, y), with other words in the product space  $(T \times Y)$ , where T is a subset of the time axis and t is one of its elements. A special run of the secondary process, that is its realization means a random point system in the space  $T \times Y$ . Indeed, if K,  $t_{-1}$ ,  $t_0$ ,  $t_1$ ,  $t_2$ , K is the Poisson-type point process and K,  $y_{-1}$ ,  $y_0$ ,  $y_1$ ,  $y_2$ , K is the series of the appropriate secondary phenomena then the realization of the secondary process can be characterized by the

K, 
$$(t_{-1}, y_{-1}), (t_0, y_0), (t_1, y_1), (t_2, y_2), K$$

random point system in the space  $T \times Y$ .

The main theorem of the product space theory on secondary processes [3] claims the following.

If the selection from the space *Y* of the secondary phenomena belonging to different points of the Poisson process is serially independent and identically distributed with the same probability measure  $\mu$ , then the random point system in the space  $T \times Y$  is also of Poisson type with parameter measure  $\lambda \times \mu$ .

It may occur that the secondary phenomena belonging to the points of the Poisson process are serially independent but their probability distribution depends on *t*. This means that the recession of a flood wave depends on the time when the flood wave was initiated. In this case, one has to use measures  $\mu_t$  instead of the single measure  $\mu$ . Then the parameter measure belonging to a set *D* of the random Poisson type point system in the product space, is determined by the following integral

$$\int_{C} \mu_{t}(D_{t})\lambda(\mathrm{d}t), \qquad (1)$$

where *C* is the projection of *D* on the set *T* and *D<sub>t</sub>* is the intersection of the set *D* with that subset of  $T \times Y$  on which *t* is constant i.e.  $D_t = \{y \mid (t, y) \in D\}$ .

The number of random points belonging to the set *D* of the product space  $T \times Y$  can be denoted as  $\eta(D)$ . So integral (1) equals to  $E(\eta(D))$ .

For simplicity we suppose in the following that  $\mu_t$  is independent of t.

A different treatment of the theory of secondary processes can be found in [10].

#### 2 Streamflow Probability Model Based on the Theory of Secondary Processes

Let the flow response to rainfall depth  $\kappa$  be characterized by function  $f(t,\kappa)$ , where  $\kappa$  is a random variable. One possible empirical version of this function is

$$f(t,\kappa) = \kappa t^{\alpha-1} e^{-\beta t}, t \ge 0, \qquad (2)$$

where  $\alpha$  and  $\beta$  are positive parameters depending on watershed characteristics.

Let the relationship between rainfall and runoff at time point  $t_i$  be described by the function

$$f(t-t_i,\kappa_i), t \ge t_i , \qquad (3)$$

where the random variables  $\kappa_i$  are serially independent. Streamflow  $\eta_i$  is described by the superposition of the functions (3), i.e. the function:

$$\sum_{t_i \leq t} f(t - t_i, \kappa_i) = \xi_t.$$

We determine the probability distribution of the random variable  $\xi_t$  for the case of function (2).

From our main theorem it follows that the number of runoff events between limits (a,b) follows Poisson distribution with a parameter given by the integral

$$\int_{-\infty}^{t} P\left(a \le \kappa (t-x)^{\alpha-1} e^{-\beta(t-x)} \le b\right) \lambda(\mathrm{d}x).$$
(4)

In the case of a=y, b=y+dy and supposing that  $\lambda(dx) = \lambda dx$ , where  $\lambda \ge 0$  constant, we get for this:

$$\int_{-\infty}^{t} P(y \le \kappa (t-x)^{\alpha-1} e^{-\beta(t-x)} \le y + dy)) \lambda dx$$
$$= \int_{0}^{\infty} P(y \le \kappa v^{\alpha-1} e^{-\beta v} \le y + dy) \lambda dv$$
$$= \int_{0}^{\infty} \frac{d}{dy} \left[ 1 - e^{-y \hat{\infty}^{\beta v} v^{1-\alpha}} \right] dy \lambda dv$$
$$= \lambda \int_{0}^{\infty} \hat{\infty} e^{\beta v} v^{1-\alpha} e^{-y \hat{\infty}^{\beta v} v^{1-\alpha}} dv dy$$

where  $\kappa$  is exponentially distributed with expected value  $1/\delta$ . It is not essential to suppose the exponential distribution; we may use any other probability distribution, too.

The probability distribution function of streamflow at time point *t* can be determined in the following way. Let denote  $\eta(I)$  the number of individual runoff events in interval *I*. Then according to the earlier results  $\eta(I)$  is Poisson distributed with parameter (4) in the case of I = (a,b). Accordingly, the characteristic function of the probability distribution we are looking for is:

$$\mathbf{e}_{0}^{\overset{\circ}{\int}\left(\mathbf{e}^{iuy}-1\right)E\left(\eta(\mathrm{d}\mathbf{y})\right)} = \mathbf{e}_{0}^{\overset{\circ}{\int}\left(\mathbf{e}^{iuy}-1\right)\left[\overset{\circ}{\underset{0}{\int}}\overset{\circ}{\partial}\mathbf{e}^{\partial \mathbf{y}}v^{1-\alpha}\mathbf{e}^{-y\overset{\circ}{\partial}\mathbf{e}^{\partial \mathbf{y}}v^{1-\alpha}}\mathrm{d}v\right]\mathrm{d}\mathbf{y}}$$
(5)

In the case of  $\alpha = 1$  we get as result:

$$e^{\lambda \int_{0}^{\infty} \left(e^{iuy}-1\right)\frac{e^{-\delta y}}{\beta y}dy}$$

which is the characteristic function of a gamma distribution. Namely, if  $\alpha = 1$ , then the equation (5) can be continued as

$$e^{\lambda \int_{0}^{\infty} \left(e^{iuy}-1\right) \left[\int_{0}^{\infty} de^{\beta v} v^{1-\alpha} e^{-ydv} dv\right] dy} = e^{\lambda \int_{0}^{\infty} \left(e^{iuy}-1\right) \frac{1}{\beta y} e^{-\delta y} dy}$$
(6)

The form (6) of the characteristic function of gamma distribution can be found on page 92 of book [2].

Considerations applied in this section can be transferred to different, possibly more complicated  $f(t,\kappa)$  functions that include rainfall-runoff relationships too. The result not necessarily can be expressed by a formula; however, it always can be calculated numerically. As a result we can always provide the probability distribution of  $\xi_t$ .

### 3 A Stochastic Programming Model for Determining the Optimal Capacity of Irrigation Reservoirs

Let us regard consecutive time sections (periods) and introduce the following notations:

 $\eta_k$  water demand in period k:  $\eta_k = h_k - \gamma_k$ , where  $h_k$  is constant meaning the total amount of demand,  $\gamma_k$  is the amount of rainfall in period k

$\xi_k$	streamflow in period k
т	storage capacity, the decision variable
М	reasonable upper bound for the capacity $m$
c(m)	cost of the reservoir as a function of its capacity
$\varsigma_k = \min(m, \xi_k)$	amount of water released in period $k$
c <sub>k</sub>	benefit per water unit in period $k$
Κ	number of periods
Ν	number of years
p > 0	inflation rate supposed to be constant up to year $N$

Let us suppose that the damage in period k is proportional to the amount of water shortage.

The model to be discussed can be formulated for the case of nonlinearly increasing penalty, too.

The random amount of damage generated in period k is described by the random variable

$$\chi_{k} = c_{k} [(\eta_{k} - \varsigma_{k})]_{+} = \begin{cases} c_{k} (\eta_{k} - \varsigma_{k}), & \text{if } \eta_{k} > \varsigma_{k} \\ 0, & \text{otherwise.} \end{cases}$$

Regarding the number of *K* consecutive periods, the expected value of the total amount of generated damages will be  $\sum_{k=1}^{K} E(\chi_k)$ . If we want to minimize the expected value of the total amount of generated damages summarized over the current and the next consecutive *N* years, then regarding the expected present value of the damages, we have to solve the following optimization problem:

$$\min\left[c(m) + \sum_{i=0}^{N} \left(\sum_{k=1}^{K} E(\chi_k)\right) \frac{1}{(1+p)^i}\right], \quad \text{supposing} \quad 0 \le m \le M$$
(7)

Problem (7) is a single variable optimization problem, and the minimum of the objective function is sought on the interval [0, M]. We show that the sum in the objective function is a convex function of m. It is enough to show the convexity for one term of the sum. Let  $G_k$  and  $F_k$  denote the probability distribution function of the random variables  $\eta_k$  resp.  $\xi_k$ , and  $f_k$  the probability density function according to  $F_k$ . Then by definition of the random variable  $\varsigma_k$  we get:

$$\frac{1}{c_{k}}E(\chi_{k}) = E([\eta_{k} - \varsigma_{k}]_{+})$$

$$= \int_{0}^{m} E([\eta_{k} - z]_{+})f_{k}(z)dz + \int_{m}^{\infty} E([\eta_{k} - m]_{+})f_{k}(z)dz$$

$$= \int_{0}^{m} \left(\int_{z}^{\infty} (1 - G_{k}(x))dx\right)f_{k}(z)dz + \int_{m}^{\infty} (1 - G_{k}(x))dx(1 - F_{k}(m))$$

$$= \int_{0}^{m} \int_{0}^{\infty} (1 - G_{k}(y + z))f_{k}(z)dydz + \int_{m}^{\infty} (1 - G_{k}(x))dx(1 - F_{k}(m))$$
(8)

Here we used the fact that if a random variable  $\xi$  has probability density function f(x), probability distribution function F(x) and its expected value exists, then it is easy to check by partial integration that for any real number z we have

$$E([\xi-z]_+) = \int_{z}^{\infty} (x-z)f(x)dx = \int_{z}^{\infty} (1-F(x))dx$$

One can check the convexity of the function  $(1/c_k)E(\chi_k)$  by differentiating twice the formula (8). As  $c_k > 0$ , k = 1, K, N, it follows that  $E(\chi_k)$  and the sum of these is also convex. As p > 0 it is clear that the expected damage summed for N years and transformed to present value is also a convex function of the variable *m*. If the function c(m) is also convex then the whole objective function is convex. However, if c(m) is not convex, then the convexity of the objective function cannot be proved, but in some special cases it may be convex as it can be seen also in our example. The optimization can be done relatively simply. The distribution of streamflow can be selected to be gamma and the distribution of water demands can be supposed to be normal or gamma, too.

The model (7) can be extended by prescribing reliability type constraints for the random water demand to be met with a high probability.

We will illustrate the model (7) with an example provided in [5] including the stochastic programming model applied to a serially linked water reservoir system. Now we regard only the first reservoir out of the two serially linked reservoirs for three consecutive periods (June, July and August). We suppose that the random variables  $\eta_1, \eta_2, \eta_3$ , describing the random water demands, are independent of each other and the random streamflow is gamma distributed with the following parameters:

	expected value ( $m^3$ )	standard deviation( $m^3$ )	9	λ
$\eta_{\scriptscriptstyle 1}$	215 760	327 120	0.000 002 016	0.435 038 479
$\eta_2$	433 608	243 600	0.000 007 307	3.168 400 000
$\eta_3$	484 416	214 368	0.000 010 541	5.106 426 041

Table 1 Parameters of the gamma distributed random water demands

Similarly we suppose that the random variables  $\xi_1, \xi_2, \xi_3$ , describing random streamflow values are independent of each other and of the random water demands and have gamma distribution with the following parameters:

Table 2 Parameters of the gamma distributed random streamflow values

	expected value ( $m^3$ )	standard deviation( $m^3$ )	9	λ
$\xi_1$	464822	186984	0.000013295	6.179658245
$\xi_2$	320576	266040	0.000004529	1.452005071
$\xi_3$	266040	234040	0.000004857	1.292152284

The cost in HUF of a reservoir with capacity m let be the following piecewise linear function

$$c(m) = \begin{cases} 100m, & \text{if } m \le 500000\\ 50000000 + 150(m - 500000), & \text{if } m > 500000 \end{cases}$$

and let us suppose that we cannot build up any reservoir with capacity greater than 25 000 000  $m^3$ .

The benefit of water/ $m^3$  in the consecutive periods let be  $c_1 = 200$  HUF,  $c_2 = 300$  HUF,  $c_3 = 250$  HUF. Let N = 10 and the constant inflation rate p = 0.05. Then the single variable optimization problem (7) can be solved by some standard Matlab routines (gamma, gammainc, quad, dblquad, fminbnd).

The optimal solution of the above described test problem is  $m = 580\ 391\ m^3$  and the optimum value according to this solution equals to 523 146 000 HUF. Fig. 1 shows the objective function values of the optimization problem (7) for its whole domain.



Figure 1 Diagram of the objective function values of optimization problem (7)

# 4 Optimization of Reservoir Release Policy

Let us regard consecutive periods and introduce the following notations:

$\xi_0$	amount of water in the reservoir at beginning the first period
$\xi_k$	amount of streamflow in period k
$a_k(b_k)$	smallest (largest) allowed amount of water in the reservoir in period $\boldsymbol{k}$
$z_k$	amount of release in period $k$ , the decision variable
Ν	number of periods
$f(z_1,\mathbf{K},z_N)$	present value of the benefit of released water $z_1, K, z_N$ in consecutive periods
m	reservoir capacity
c(m)	cost of the reservoir as a function of its capacity
Κ	upper bound - the cost of building the reservoir with capacity $m$
p	reliability level prescribed, close to one

The optimization problem is formulated as

$$\max[f(z_1, \mathbf{K}, z_N) - c(m)] \quad \text{supposing that}$$

$$P\left\{a_k \le \xi_0 + \sum_{j=1}^k \xi_j - \sum_{j=1}^k z_j \le b_k, \, k = 1, \mathbf{K}, N\right\} \ge p \tag{9}$$

 $0 \le z_k \le m, \quad k = 1, \mathbf{K}, N.$ 

If *m* is given then we don't regard it as a variable, otherwise the problem remains unchanged. If we want to build into the model the random water demands  $\eta_k$ , it may be done without any further as in Section 3 was discussed. The numerical solution of problem (9) is possible if we put some special assumptions on the random variables  $\xi_1$ , K,  $\xi_N$ , see the papers [7], [8], [9]. The model (9) can be successfully applied to scaling the capacity value *m*.

A further variant of model (9) is when the decision-maker may give an upper bound K on the cost of building the reservoir with capacity m, c(m). In this case it is not necessary to subtract the value c(m) from the objective function and the problem of the modified model can be formulated as

 $\max[f(z_1, K, z_N) - c(m)]$  supposing, that

$$P\left\{a_{k} \leq \xi_{0} + \sum_{j=1}^{k} \xi_{j} - \sum_{j=1}^{k} z_{j} \leq b_{k}, k = 1, K, N\right\} \geq p$$
(10)

$$c(m) \leq K, \quad 0 \leq z_k \leq m, \quad k = 1, K, N$$

It's worth mentioning that if the probability distribution of the random variables  $\xi_1$ , K,  $\xi_N$  is continuous and their density function is logarithmically concave, then the  $m, z_1$ , K,  $z_N$  feasible domain of problems (9) and (10) is convex (see for example Prékopa [6]). So if  $f(z_1, K, z_N)$  and c(m) are convex functions, then the problems (9) and (10) are convex.

Let us regard a reservoir for four consecutive months, say April, May, June and July, as an example of Problem (10). Let streamflow data follow joint normal probability distribution with the following parameters:

Parameters of joint normal distribution of the random streamflow values

	expected value	lue standard deviation		correlation coefficients		
	$(10^6 m^3)$	$(10^6 m^3)$				
$\xi_1$	79.74	83.51	1.000	0.284	-0.017	0.047
$\xi_2$	29.78	63.11	0.284	1.000	0.333	0.198

$\xi_3$	-4.52	73.98	-0.017	0.333	1.000	0.579
$\xi_4$	-43.44	73.96	0.047	0.198	0.579	1.000

Create the aggregated random variables

$$\begin{split} & \zeta_1 = \xi_1 \\ & \zeta_2 = \xi_1 + \xi_2 \\ & \zeta_3 = \xi_1 + \xi_2 + \xi_3 \\ & \zeta_4 = \xi_1 + \xi_2 + \xi_3 + \xi_4 \,. \end{split}$$

These random variables as linear transforms of  $\xi_1, \xi_2, \xi_3, \xi_4$  have also normal distribution with the transformed expected values, standard deviations and correlation coefficients:

Table 4 Parameters of joint normal distribution of the random stream flow values

	expected value	xpected value standard deviation		correlation coefficients		
	$(10^6 m^3)$	$(10^6 m^3)$				
$\varsigma_1$	79.740	83.510	1.000	0.859	0.670	0.542
$\varsigma_2$	109.520	118.112	0.859	1.000	0.873	0.736
$\varsigma_3$	105.000	149.408	0.670	0.873	1.000	0.935
$\varsigma_4$	61.560	191.201	0.542	0.736	0.935	1.000

Let us suppose that optimization in the problem (10) $f(z_1, z_2, z_3, z_4) = 40z_1 + 70z_2 + 80z_3 + 50z_4$ , i.e. the total benefit of released water is a linear function. Let the cost of the reservoir of capacity m also be linear function: c(m) = 50m. For the smallest water level of the reservoir let be prescribed  $a_k = 100$  in all periods k = 1, 2, 3, 4; for the largest water level of the reservoir let be prescribed  $b_k = 1000$  in all periods k = 1, 2, 3, 4; and let us suppose that at the beginning of the first period the season starts with full reservoir. If we solve the arising optimization problem with different bounds on the building cost then the decision-maker can select the economically reasonable capacity. Introducing new variables for simplifying the terms inside the probability expressing the reliability-type constraint, we solved the following optimization problem for different building up cost bounds K:

 $\max(40z_1 + 70z_2 + 80z_3 + 50z_4)$  supposing that

 $l_1 = 100 + z_1 - 1000$ 

$$\begin{split} l_2 &= 100 + z_1 + z_2 - 1000 \\ l_3 &= 100 + z_1 + z_2 + z_3 - 1000 \\ l_4 &= 100 + z_1 + z_2 + z_3 + z_4 - 1000 \\ u_1 &= 1000 + z_1 - 1000 \\ u_2 &= 1000 + z_1 + z_2 - 1000 \\ u_3 &= 1000 + z_1 + z_2 + z_3 - 1000 \\ u_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_4 &= 1000 + z_1 + z_2 + z_3 + z_4 - 1000 \\ l_5 &= 1000 \\ l_6 &= 10000 \\ l_6 &= 1000 \\ l_6 &= 10000 \\ l_6 &= 10000 \\ l_6 &= 10$$

 $50m \le K, z_1 \le m, z_2 \le m, z_3 \le m, z_4 \le m$ .

Notice that the probabilistic constraint has been multiplied by 100. As a result, the problem can be solved numerically in a more stable way. Then the only difficulty is the calculation of the probability values and its partial derivatives. For this we can write the probability value in the following form:

$$P\begin{pmatrix} l_{1} \leq \zeta_{1} \leq u_{1} \\ l_{2} \leq \zeta_{2} \leq u_{2} \\ l_{3} \leq \zeta_{3} \leq u_{3} \\ l_{4} \leq \zeta_{4} \leq u_{4} \end{pmatrix} = F(u_{1}, u_{2}, u_{3}, u_{4}) - F(l_{1}, u_{2}, u_{3}, u_{4})$$
$$- F(u_{1}, l_{2}, u_{3}, u_{4}) - F(u_{1}, u_{2}, l_{3}, u_{4})$$
$$- F(u_{1}, u_{2}, u_{3}, l_{4}) + F(l_{1}, l_{2}, u_{3}, u_{4})$$
$$+ F(l_{1}, u_{2}, l_{3}, u_{4}) + F(l_{1}, u_{2}, u_{3}, l_{4})$$
$$+ F(u_{1}, l_{2}, l_{3}, u_{4}) + F(u_{1}, l_{2}, u_{3}, l_{4})$$
$$+ F(u_{1}, u_{2}, l_{3}, u_{4}) - F(l_{1}, l_{2}, l_{3}, u_{4})$$
$$- F(l_{1}, l_{2}, l_{3}, l_{4}) - F(l_{1}, u_{2}, l_{3}, u_{4})$$
$$- F(l_{1}, l_{2}, l_{3}, l_{4}) - F(l_{1}, u_{2}, l_{3}, u_{4})$$
$$- F(u_{1}, l_{2}, l_{3}, l_{4}) + F(l_{1}, l_{2}, l_{3}, u_{4})$$

where  $F(x_1, x_2, x_3, x_4)$  denotes the joint normal probability distribution function of the random variables  $\zeta_1, \zeta_2, \zeta_3, \zeta_4$  with parameters given in Table 4. This means that for the calculation of one probability value we have to calculate  $2^4 = 16$  four dimensional normal probability distribution function values.

We prepared the AMPL model of the above defined nonlinear programming problem. To calculate values of the multivariate normal probability distribution functions the numerical integration code developed by A Genz ([1]) has been added to the AMPLE model. Then the problem was solved by the solver LOQO for different values of cost bound K. The results are summarized in Table 5.

Table 5
---------

Here *m* is the optimal capacity of the reservoir and  $z_k$  is the optimal amount of water release in period *k*. All of them are given in  $10^6 m^3$ . *K* and the total benefit values are given in millions of HUF.

Number	K	total benefit	т	$Z_1$	$z_2$	$z_3$	$z_4$
1	10000	36634.493	200.001	200.001	180.665	199.848	0.000
2	10500	39682.114	210.011	210.011	206.903	209.975	0.010
3	11000	41250.695	220.003	220.003	212.155	219.996	0.001
4	11500	42270.302	230.004	230.004	209.583	229.990	0.003
5	12000	42948.048	240.010	240.010	202.120	239.990	0.003
6	12500	43378.927	250.012	250.012	191.146	249.973	0.009
7	13000	43615.155	259.999	259.999	177.390	259.973	0.003
8	13500	43741.739	269.943	268.574	162.960	269.943	0.002
9	14000	43792.337	279.971	268.973	151.969	279.971	0.001
10	14500	43836.424	289.895	268.947	141.354	289.895	0.000
11	15000	43861.877	299.046	268.963	132.222	299.046	0.002

Figure 2 represents the possible benefit values according to different cost bounds K. This graph may be useful for decision-makers when deciding how much should be spent for providing a given reservoir capacity. The graph shows the benefit increase of water releases if the amount of money spent for building the reservoir is increased. The decision should take into account, of course, that the cost of reservoir occurs only once and the benefit of released water can be realized for many years.



Figure 2 The benefit of irrigation water in function of the money spent for building a reservoir

## 5 Probability Constrained Stochastic Programming Model for an Intake Facility

The capacity of an intake facility, say a pumping station is considered to satisfy random water demands (e.g. irrigation) utilizing the available streamflow. Let us regard a given time period which can be a month, say August of the year. We will prescribe that the number of days with unsatisfied water demands should not exceed a given value with a high probability. The model will be described for a time interval of n days. We introduce the notations:

$\xi_1, \mathbf{K}, \xi_n$	daily available stream flows
$\gamma_1, \mathbf{K}, \gamma_n$	daily rainfalls
$\eta_1, \mathbf{K}, \eta_n$	daily water demands
m	daily capacity of the intake facility, the decision var
Μ	upper bound for capacity m
c(m)	cost of the facility
b	maximum number of days with unsatisfied water given time period
р	reliability level prescribed, close to one

There is enough water on the  $k^{\text{th}}$  day if and only if the following relation holds

$$\min\left(\xi_k, m\right) + \gamma_k \ge \eta_k. \tag{11}$$

Let  $x_1$ , K,  $x_n$  be deterministic variables which take on values 0 and 1, only. The following relation doesn't mean any constraint if  $x_k = 0$ , but if  $x_k = 1$  it is equivalent to the constraint (11):

$$\min(\xi_k, m) + \gamma_k \ge x_k \eta_k. \tag{12}$$

Beside (12) for all k = 1, K, n prescribing the constraint

$$x_1 + \Lambda + x_n \ge n - b$$

we require that at least n-b out of the constraints (11) be met, i.e. at least  $x_1$ , K,  $x_n$  times the opposite of the constraint (11) be met. Then our model can be formulated as

min c(m) supposing, that

$$P\{\min(\xi_k, m) + \gamma_k \ge x_k \eta_k, k = 1, K, n - b\} \ge p$$

$$x_1 + \Lambda + x_n \ge n - b$$

$$x_k = 0 \text{ or } 1, k = 1, K, n, 0 \le m \le M.$$
(13)

Like the earlier models, this model also has more variants. Among others, one can build into the objective function, a cost factor, that depends on the number of days, *b*. If the random variables  $\xi_k, \eta_k, \gamma_k, k = 1, K$ , *N* have continuous joint probability distribution and their joint density function is logarithmically concave then the constraints of problem (13) except the constraints  $x_k = 0$  or 1, define a convex feasible set.

#### References

- [1] A. Genz, Numerical Computation of the Multivariate Normal Probabilities, Journal of Computational and Graphical Statistics, **1**, 141-150, 1992
- [2] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Translated and annotated by K. L. Chung, Addison-Wesley, Cambridge, 1954
- [3] A. Prékopa, On secondary processes generated by random point distributions of Poisson type, Annales Univ. Sci. R. Eötvös, Sectio Math., 2, 139-146, 1959
- [4] A. Prékopa, Contributions to the theory of stochastic programming, Mathematical Programming, 4, 202-221, 1973

- [5] A. Prékopa, T. Rapcsák and I. Zsuffa, A new method for serially linked reservoir system design using stochastic programming (in Hungarian), *Alkalmazott Matematikai Lapok*, 2, 189-2001, 1976
- [6] A. Prékopa, *Stochastic Programming*, Kluwer Scientific Publishers, Dordrecht, Boston, 1995
- [7] A. Prékopa and T. Szántai, On Optimal Regulation of a Storage Level with Application to the Water Level Regulation of a Lake, *European Journal of Operational Research*, **3**, 175-189, 1979
- [8] T. Szántai, Evaluation of a Special Multivariate Gamma Distribution, Mathematical Programming Study, 27, 1-16, 1986
- [9] T. Szántai, A Computer Code for Solution of Probabilistic Constrained Stochastic Programming Problems, In: *Numerical Techniques for Stochastic Optimization* (Yu. Ermoliev and R. J.-B. Wets, eds.), Springer, New York, 229-235, 1988
- [10] L. Takács, Secondary processes generated by Poisson process and their applications in physics (in Hungarian), *MTA Matematikai és Fizikai Tudományok Osztályának Közleményei*, **4**, 473-504, 1954