

# On probabilistic constraints induced by rectangular sets and multivariate normal distributions

Wim Van Ackooij · René Henrion ·  
Andris Möller · Riadh Zorgati

Received: 9 March 2010 / Accepted: 25 May 2010 / Published online: 9 June 2010  
© Springer-Verlag 2010

**Abstract** In this paper, we consider optimization problems under probabilistic constraints which are defined by two-sided inequalities for the underlying normally distributed random vector. As a main step for an algorithmic solution of such problems, we prove a derivative formula for (normal) probabilities of rectangles as functions of their lower or upper bounds. This formula allows to reduce the calculus of such derivatives to the calculus of (normal) probabilities of rectangles themselves thus generalizing a similar well-known statement for multivariate normal distribution functions. As an application, we consider a problem from water reservoir management. One of the outcomes of the problem solution is that the (still frequently encountered) use of simple individual probabilistic constraints can completely fail. By contrast, the (more difficult) use of joint probabilistic constraints, which heavily depends on the derivative formula mentioned before, yields very reasonable and robust solutions over the whole time horizon considered.

**Keywords** Stochastic programming · Probabilistic constraints ·  
Chance constraints · Derivative of probabilities of rectangles · Water reservoir  
management

**Mathematics Subject Classification (2000)** MSC 90C15 · MSC 90C90

---

This work was supported by the OSIRIS Department of *Electricité de France R&D* and by the DFG Research Center MATHEON “Mathematics for key technologies” in Berlin.

---

W. Van Ackooij · R. Zorgati  
Électricité de France R&D, 1, avenue du Général de Gaulle, 92141 Clamart CEDEX, France

R. Henrion (✉) · A. Möller  
Weierstrass Institute, Mohrenstraße 39, 10117 Berlin, Germany  
e-mail: henrion@wias-berlin.de

## 1 Introduction

Probabilistic constraints have turned out to be an important tool for modeling optimization problems with uncertain data (see Prékopa 1995, 2003). Often, they are induced by two-sided stochastic inequalities bounding a normally distributed random vector by some decision-dependent functions. More precisely the probabilistic constraint may take the form

$$\mathbb{P}(\alpha(x) \leq \xi \leq \beta(x)) \geq p$$

Here,  $\xi$  is a random vector having a regular multivariate normal distribution,  $\mathbb{P}$  denotes the probability measure,  $p \in (0, 1)$  is a probability level and  $x$  refers to a decision vector. In geometric terms, it is required that the probability of some  $x$ -dependent rectangle be not smaller than  $p$ . In order to determine an optimal decision  $x^*$  in the context of an optimization problem, one has to have access to values and derivatives of this probability function. As far as values are concerned, one may employ numerical algorithms designed for the calculus of normal distribution functions (Szántai 2000), of normal probabilities of general convex sets (Deák 1980) or directly of rectangles (Genz 1992; Genz and Bretz 2009). However, none of these algorithms provides gradients of the probability function with respect to changes of the lower and upper limit of the rectangle. In case of one-sided constraints (i.e.  $\alpha$  is missing so that one is dealing with distribution functions), there is no problem to reduce the computation of the gradient to that of a value of a distribution function (see Lemma 1 below). Formally, one could also do so with gradients of two-sided constraints by exploiting a representation of rectangle probabilities in terms of distribution functions (see (1)) and then taking derivatives of the latter ones term by term. We note that such representation allowing for reduction of derivatives to those of distribution functions is available even for general polyhedra (Henrion and Römisch 2005). This approach, however, becomes impractical already in small dimension. For example in the case of an  $n$ -dimensional rectangle, the number of terms in the representation equals  $2^n$ . For a general but difficult to apply derivative formula for probability functions we refer to Uryasev (1995).

The aim of this paper is threefold: first we present a simple formula for the derivative of the normal probability of rectangles with respect to their lower and upper limits. In particular, this formula allows to reduce the problem to the same calculus of probabilities of rectangles (in one dimension less). Consequently, the same algorithm in Genz (1992) can be used for computing at a time values and derivatives of the probability function introduced above. Second, we apply this formula in the context of a problem from water reservoir management. This problem involves a 64-dimensional random vector and its solution is based on their joint distribution. Doing so, we demonstrate that joint probabilistic constraints with two-sided bounds can be successfully treated even in rather large dimension. Third, it is shown that treating probabilistic constraints individually—as it is still often done in water management problems in order to drastically simplify the algorithmic solution—typically fails completely as far as robustness over the whole time horizon is concerned.

## 2 A derivative formula for probabilities of rectangles under a multivariate normal distribution

Let  $\xi$  be some  $n$ -dimensional random vector having a nondegenerate multivariate normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$ . We will write  $\xi \sim \mathcal{N}(\mu, \Sigma)$  for short. Denote by

$$\Phi_\xi(z) := \mathbb{P}(\xi \leq z) \quad \forall z \in \mathbb{R}^n$$

its cumulative distribution function (with  $\mathbb{P}$  referring to the underlying probability measure). We further introduce the rectangle probability function

$$F_\xi(a, b) := \mathbb{P}(a \leq \xi \leq b) \quad \forall a, b \in \mathbb{R}^n : a \leq b.$$

The following relation is well known to hold whenever  $a \leq b$ :

$$F_\xi(a, b) = \sum_{i_1, \dots, i_n \in \{0,1\}} (-1)^{[n + \sum_{j=1}^n i_j]} \Phi_\xi(y_{i_1}, \dots, y_{i_n}), \tag{1}$$

where

$$y_{i_j} := \begin{cases} a_j & \text{if } i_j = 0 \\ b_j & \text{if } i_j = 1 \end{cases}.$$

For instance, if  $n = 2$ , the probability of the rectangle  $[a, b]$  calculates via the distribution function as

$$F_\xi(a, b) = \Phi_\xi(a_1, a_2) - \Phi_\xi(a_1, b_2) - \Phi_\xi(b_1, a_2) + \Phi_\xi(b_1, b_2).$$

The following lemma can be found (in its equivalent form for standard normal distributions) in Prékopa (1995). It shows how the derivative of a multivariate normal distribution can be reduced to values of a different multivariate normal distribution (in one dimension less):

**Lemma 1** *Assume that  $\xi \sim \mathcal{N}(\mu, \Sigma)$  with some positive definite covariance matrix  $\Sigma = (\sigma_{ij})$ . Then,  $\Phi_\xi$  is continuously differentiable and*

$$\frac{\partial \Phi_\xi}{\partial z_i}(z) = f_{\xi_i}(z_i) \cdot \Phi_{\tilde{\xi}(z_i)}(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n) \quad (i = 1, \dots, n).$$

Here,  $f_{\xi_i}$  denotes the one-dimensional probability density of the component  $\xi_i$ ,  $\tilde{\xi}(z_i)$  is an  $n - 1$ -dimensional random vector distributed according to  $\tilde{\xi}(z_i) \sim \mathcal{N}(\hat{\mu}, \hat{\Sigma})$ ,  $\hat{\mu}$  results from the vector  $\mu + \sigma_{ii}^{-1}(z_i - \mu_i)\sigma_i$  by deleting component  $i$  and  $\hat{\Sigma}$  results from the matrix  $\Sigma - \sigma_{ii}^{-1}\sigma_i\sigma_i^T$  by deleting row  $i$  and column  $i$ , where  $\sigma_i$  refers to column  $i$  of  $\Sigma$ .

In the next theorem, we generalize the result of Lemma 1 to the case of probability functions  $F_\xi$  defined by rectangles. In particular, the presented formula allows to again reduce the derivative of  $F_\xi$  to the calculus of values of a similar function induced by a different normally distributed random vector. Though easy to prove, this formula has not been presented so far to the best of our knowledge and has an enormous impact on the dimensionality of problems which can be solved. Recall that the naive approach discussed in the introduction (which amounts to applying Lemma 1 to a representation of rectangles by means of orthants) reduces the dimension  $n$  of the considered random vector to very small numbers like  $n = 5$  to  $n = 10$ . This is because the number of terms in (1) equals  $2^n$ . In our application we shall deal, however, with a random vector having dimension  $n = 64$ .

**Theorem 1** *Assume that  $\xi \sim \mathcal{N}(\mu, \Sigma)$  with some positive definite covariance matrix  $\Sigma$ . Then, for  $i = 1, \dots, n$ ,*

$$\frac{\partial}{\partial b_i} F_\xi(a, b) = f_{\xi_i}(b_i) F_{\tilde{\xi}(b_i)}(\tilde{a}, \tilde{b}) \tag{2}$$

$$\frac{\partial}{\partial a_i} F_\xi(a, b) = -f_{\xi_i}(a_i) F_{\tilde{\xi}(a_i)}(\tilde{a}, \tilde{b}). \tag{3}$$

Here,  $f_{\xi_i}$  is as in Lemma 1,  $\tilde{\xi}(b_i), \tilde{\xi}(a_i)$ , are  $n - 1$ -dimensional random vectors distributed according to  $\tilde{\xi}(b_i), \tilde{\xi}(a_i) \sim \mathcal{N}(\hat{\mu}, \hat{\Sigma})$  such that  $\hat{\mu}$  results from the vector  $\mu + \sigma_{ii}^{-1}(b_i - \mu_i)\sigma_i$  (in case of  $b_i$ ) or from the vector  $\mu + \sigma_{ii}^{-1}(a_i - \mu_i)\sigma_i$  (in case of  $a_i$ ) by deleting component  $i$  and  $\hat{\Sigma}$  is defined as in Lemma 1. Moreover  $\tilde{a}$  and  $\tilde{b}$  result from  $a$  and  $b$  by deleting the respective component  $i$ .

*Proof* It suffices to prove (2), the proof of (3) running along the same lines. In order to show (2), we may assume without loss of generality and for notational convenience that  $i = 1$ . According to (1), we have

$$\frac{\partial}{\partial b_1} F_\xi(a, b) = \sum_{i_1, \dots, i_n \in \{0, 1\}} (-1)^{[n + \sum_{j=1}^n i_j]} \frac{\partial}{\partial b_1} \Phi_\xi(y_{i_1}, \dots, y_{i_n}).$$

If, in the sum on the right hand side,  $i_1 = 0$  then  $b_1 \notin \{y_{i_1}, \dots, y_{i_n}\}$ , i.e., the variable  $b_1$  does not belong to the set of variables  $\{y_{i_1}, \dots, y_{i_n}\}$ . Hence,

$$\frac{\partial}{\partial b_1} \Phi_\xi(y_{i_1}, \dots, y_{i_n}) = 0.$$

Otherwise, if  $i_1 = 1$ , then

$$\frac{\partial}{\partial b_1} \Phi_\xi(y_{i_1}, \dots, y_{i_n}) = \frac{\partial}{\partial b_1} \Phi_\xi(b_1, y_{i_2}, \dots, y_{i_n}) = f_{\xi_1}(b_1) \Phi_{\tilde{\xi}(b_1)}(y_{i_2}, \dots, y_{i_n})$$

by Lemma 1 with the notation used there. It follows that

$$\begin{aligned} \frac{\partial}{\partial b_1} F_\xi(a, b) &= \sum_{i_2, \dots, i_n \in \{0,1\}} (-1)^{\left[n+1+\sum_{j=2}^n i_j\right]} f_{\xi_1}(b_1) \Phi_{\tilde{\xi}(b_1)}(y_{i_2}, \dots, y_{i_n}) \\ &= f_{\xi_1}(b_1) \sum_{i_2, \dots, i_n \in \{0,1\}} (-1)^{\left[n-1+\sum_{j=2}^n i_j\right]} \Phi_{\tilde{\xi}(b_1)}(y_{i_2}, \dots, y_{i_n}) \\ &= f_{\xi_1}(b_1) F_{\tilde{\xi}(b_1)}(a_2, \dots, a_n, b_2, \dots, b_n), \end{aligned}$$

with the last equality resulting once more from (1) but in one dimension less. Now, the asserted formula (2) (for  $i = 1$ ) follows from the respective definitions.  $\square$

In order to demonstrate the impact of the derived formula, we consider an optimization problem

$$\min\{c^T x \mid \mathbb{P}(Ax + a \leq L\xi \leq Bx + b) \geq p\}, \tag{4}$$

where  $A, B, L$  and  $a, b, c$  are matrices and vectors, respectively, of appropriate orders. Given that  $\xi$  (and so  $L\xi$  too) has a multivariate normal distribution, we know from Prékopa (1995) that the function

$$x \mapsto \log \mathbb{P}(Ax \leq L\xi \leq Bx) \tag{5}$$

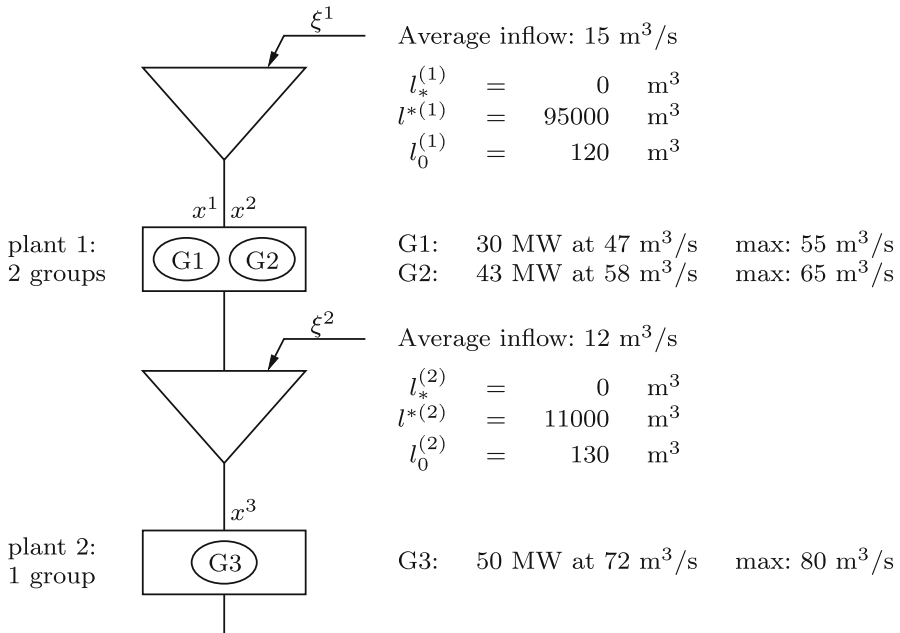
is concave. This allows to rewrite the optimization problem as a convex one:

$$\min\{c^T x \mid -\log \mathbb{P}(Ax + a \leq L\xi \leq Bx + b) \leq -\log p\}$$

Now one can apply, for instance, a supporting hyperplane type method as described in Prékopa (1995) in order to solve this problem. This requires, apart from functional values, also to calculate gradients of the function (5) which amounts to determine partial derivatives of the function  $F_\xi(Ax, Bx)$  introduced above. The latter task can efficiently be realized with the aid of our formula proved in Theorem 1 by relying on the same algorithm as used for determining values of such functions.

### 3 Application to a problem of hydro power management

In this section, we are going to illustrate the application of probabilistic programming described above to a problem of hydro power management involving two connected reservoirs each of them supplied with independent random inflow. The aim is to find optimal (in the sense of energy production) release policies for the two reservoirs over a given future time horizon. Optimal release is mainly driven by different efficiencies and production levels of the turbines involved, by some time-dependent price for selling energy as well as by filling levels for the reservoirs which have to be satisfied with high probability.



**Fig. 1** Sketch of a problem in hydro power management (for details see text)

There are two different turbines producing energy from the release out of the first reservoir and one turbine producing energy from the release out of the second reservoir. Each turbine has some maximum operating level and they all are different in efficiency. Both of the reservoirs have upper and lower levels to be respected during operation. Initial levels in the reservoirs define the starting conditions. The situation is sketched in Fig. 1. As the realizations of the future inflows to the reservoirs are not known, satisfaction of reservoir levels is modeled by means of probabilistic constraints. The important point is that we insist on joint probabilistic constraints which means that keeping the levels with high probability is required for the whole time horizon. We shall see later that the still frequently applied model with individual probabilistic constraints while being much simpler and therefore more appealing from the algorithmic point of view only guarantees keeping the levels with high probability at each time step separately whereas violation at least once in the whole interval can also occur with high probability. We are assuming here a static decision process, i.e., the release policies are determined in the very beginning without future reactions on previously observed inflows. For a recent study of a dynamic model with joint probabilistic constraints in this context, we refer to [Andrieu et al. \(2009\)](#). This model, however, is too demanding for the purposes of this paper because it allows only for a small number of time steps at the present level of knowledge whereas we are going to consider two inflow processes with a subdivision of the time horizon into 32 intervals. This leads to joint probabilistic constraints based on a 64-dimensional random vector. One may of course imbed the solution approach presented here into a pseudo-dynamic context by considering a moving window technique as generally made in real-life hydro policies.

### 3.1 Description of the model

In order to provide a model of the problem, we introduce the following notation:

- $l_*^{(i)}, l^{*(i)}$  = lower and upper level to be respected in reservoir  $i$  ( $i = 1, 2$ ).
- $l_0^{(i)}$  = initial water level in reservoir  $i$  ( $i = 1, 2$ ).
- $\xi_t^{(i)}$  = random inflow to reservoir  $i$  during time interval  $t$  ( $i = 1, 2$ ;  
 $t = 1, \dots, T$ ).
- $m^{(j)}$  = upper operating level of turbine  $j$  ( $j = 1, 2, 3$ ).
- $x_t^{(j)}$  = water release used for turbine  $j$  during time interval  $t$   
( $j = 1, 2, 3$ ;  $t = 1, \dots, T$ ).
- $\lambda^{(j)}$  = efficiency for turbine  $j$  ( $j = 1, 2, 3$ )
- $\pi_t$  = price for one unit of water released during time interval  $t$  ( $t=1, \dots, T$ ).
- $\omega_i$  = evaluation coefficient for end level of reservoir  $i$  ( $i = 1, 2$ )
- $p$  = probability required for keeping reservoir levels.

The objective function of the problem is to maximize the profit obtained from selling energy produced by water release and at the same time to keep a possibly high water level in both reservoirs at the end of the time horizon. In other words, one wishes to maximize the quantity

$$\sum_{j=1}^3 \sum_{t=1}^T \lambda^{(j)} \pi_t x_t^{(j)} + \omega_1 l_T^{(1)} + \omega_2 l_T^{(2)}, \tag{6}$$

where  $l_T^{(i)}$  refers to the water level in reservoir  $i$  after the last time interval  $T$ .

### 3.2 The optimization problem under joint probabilistic constraints

The filling level  $l_t^{(1)}$  of reservoir 1 at the end of time interval  $t$  is given by

$$l_t^{(1)} = l_0^{(1)} + \sum_{\tau=1}^t \xi_{\tau}^{(1)} - \sum_{\tau=1}^t x_{\tau}^{(1)} - \sum_{\tau=1}^t x_{\tau}^{(2)} \tag{7}$$

whereas the filling level  $l_t^{(2)}$  of reservoir 2 calculates as

$$l_t^{(2)} = l_0^{(2)} + \sum_{\tau=1}^t \xi_{\tau}^{(2)} + \sum_{\tau=1}^t x_{\tau}^{(1)} + \sum_{\tau=1}^t x_{\tau}^{(2)} - \sum_{\tau=1}^t x_{\tau}^{(3)}. \tag{8}$$

The filling level constraints are given by

$$l_*^{(i)} \leq l_t^{(i)} \leq l^{*(i)} \quad (i = 1, 2; t = 1, \dots, T). \tag{9}$$

Using the lower triangular matrix

$$\Delta := \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \dots & 1 & 0 \\ 1 & \dots & 1 & 1 \end{pmatrix},$$

and putting

$$\xi^{(i)} := (\xi_1^{(i)}, \dots, \xi_T^{(i)}), \quad x^{(j)} := (x_1^{(j)}, \dots, x_T^{(j)}) \quad (i = 1, 2; j = 1, 2, 3),$$

one can rewrite the the whole system (9) of level constraints for the two reservoirs using (7) and (8) in the more compact forms

$$l_*^{(1)} \mathbf{1} \leq l_0^{(1)} \mathbf{1} + \Delta \xi^{(1)} - \Delta x^{(1)} - \Delta x^{(2)} \leq l^{*(1)} \mathbf{1}$$

and

$$l_*^{(2)} \mathbf{1} \leq l_0^{(2)} \mathbf{1} + \Delta \xi^{(2)} + \Delta x^{(1)} + \Delta x^{(2)} - \Delta x^{(3)} \leq l^{*(2)} \mathbf{1},$$

where  $\mathbf{1}$  refers to the vector with all entries equal to 1. Rewriting this and combining the two systems to a single one, we arrive at

$$\begin{aligned} l_*^{(1)} \mathbf{1} - l_0^{(1)} \mathbf{1} + \Delta x^{(1)} + \Delta x^{(2)} &\leq \Delta \xi^{(1)} \leq l^{*(1)} \mathbf{1} - l_0^{(1)} \mathbf{1} + \Delta x^{(1)} + \Delta x^{(2)} \\ l_*^{(2)} \mathbf{1} - l_0^{(2)} \mathbf{1} - \Delta x^{(1)} - \Delta x^{(2)} + \Delta x^{(3)} &\leq \Delta \xi^{(2)} \leq l^{*(1)} \mathbf{1} - l_0^{(2)} \mathbf{1} \\ &\quad - \Delta x^{(1)} - \Delta x^{(2)} + \Delta x^{(3)}. \end{aligned}$$

On a more abstract level, this system is of the form

$$Ax + a \leq L\xi \leq Bx + b \tag{10}$$

with

$$\begin{aligned} x &= (x^{(1)}, x^{(2)}, x^{(3)}), \quad \xi = (\xi^{(1)}, \xi^{(2)}) \\ A = B &= \begin{pmatrix} \Delta & \Delta & 0 \\ -\Delta & -\Delta & \Delta \end{pmatrix} \quad L = \begin{pmatrix} \Delta & 0 \\ 0 & \Delta \end{pmatrix} \\ a &= \begin{pmatrix} l_*^{(1)} \mathbf{1} - l_0^{(1)} \mathbf{1} \\ l_*^{(2)} \mathbf{1} - l_0^{(2)} \mathbf{1} \end{pmatrix}, \quad b = \begin{pmatrix} l^{*(1)} \mathbf{1} - l_0^{(1)} \mathbf{1} \\ l^{*(1)} \mathbf{1} - l_0^{(2)} \mathbf{1} \end{pmatrix}. \end{aligned}$$



Consequently, turning (10) into probabilistic reservoir level constraints, one arrives exactly at the probabilistic constraint considered in problem (4). Given the fact that the components of the inequality system (10) refer to time steps, the interpretation of the associated probabilistic constraint in (4) is as follows: a release policy  $x$  for feeding the three turbines is declared to be feasible if it leads to a satisfaction of upper and lower levels in the two reservoirs **over the whole time interval** at least with probability  $p$ .

In addition to the probabilistic constraint we also have the deterministic constraints

$$0 \leq x_t^{(j)} \leq m^{(j)} \quad (j = 1, 2, 3; t = 1, \dots, T).$$

Concerning the objective in (6), we must take into account that the end levels  $l_T^{(1)}, l_T^{(2)}$  of the two reservoirs are random variables, so we take their expected values. By (7) and (8) one gets that

$$\begin{aligned} \mathbb{E}l_T^{(1)} &= l_0^{(1)} + \sum_{t=1}^T \mathbb{E}\xi_t^{(1)} - \sum_{t=1}^T x_t^{(1)} - \sum_{t=1}^T x_t^{(2)} \\ \mathbb{E}l_T^{(2)} &= l_0^{(2)} + \sum_{t=1}^T \mathbb{E}\xi_t^{(2)} + \sum_{t=1}^T x_t^{(1)} + \sum_{t=1}^T x_t^{(2)} - \sum_{t=1}^T x_t^{(3)}. \end{aligned}$$

Consequently, according to (6), we wish to maximize the quantity

$$\sum_{j=1}^3 \sum_{t=1}^T \lambda^{(j)} \pi_t x_t^{(j)} - \omega_1 \left( \sum_{t=1}^T x_t^{(1)} + \sum_{t=1}^T x_t^{(2)} \right) + \omega_2 \left( \sum_{t=1}^T x_t^{(1)} + \sum_{t=1}^T x_t^{(2)} - \sum_{t=1}^T x_t^{(3)} \right),$$

where we left out, without loss of generality, terms not depending on the decision vector  $x$ . This profit function is linear in  $x$ , so the coefficient vector  $c$  in the optimization problem (4) is easily found as  $c = (c^{(1)}, c^{(2)}, c^{(3)})$ , where

$$c_t^{(1)} = \omega_1 - \omega_2 - \lambda^{(1)} \pi_t, \quad c_t^{(2)} = \omega_1 - \omega_2 - \lambda^{(2)} \pi_t, \quad c_t^{(3)} = \omega_2 - \lambda^{(3)} \pi_t \quad (t = 1, \dots, T)$$

and where the reversed signs are due to minimizing  $-c^T x$  rather than maximizing  $c^T x$ .

### 3.3 Numerical case study

In this section, we are going to provide a numerical case study, where we solve for the purpose of comparison four different optimization problems, namely

- The expected-value problem

$$\min\{c^T x \mid Ax + a \leq L\mathbb{E}\xi \leq Bx + b\},$$

where the random vector is simply replaced by its expectation

- The problem with individual probabilistic constraints

$$\min \left\{ c^T x \mid \begin{array}{l} \mathbb{P}(A_i x + a_i \leq L_i \xi) \geq p, \\ \mathbb{P}(L_i \xi \leq B_i x + b_i) \geq p \end{array} \quad (i = 1, \dots, 2T) \right\},$$

where the reservoir level constraints are required to hold with probability  $p$  at **each time step and for each reservoir level separately**.

- The problem (4) with joint probabilistic constraints according to the specifications in the previous section.
- The robust ‘max  $p$ ’-problem

$$\max\{p \mid \mathbb{P}(Ax + a \leq L\xi \leq Bx + b) \geq p\},$$

where the objective is to maximize the probability level under which level satisfaction can still be guaranteed.

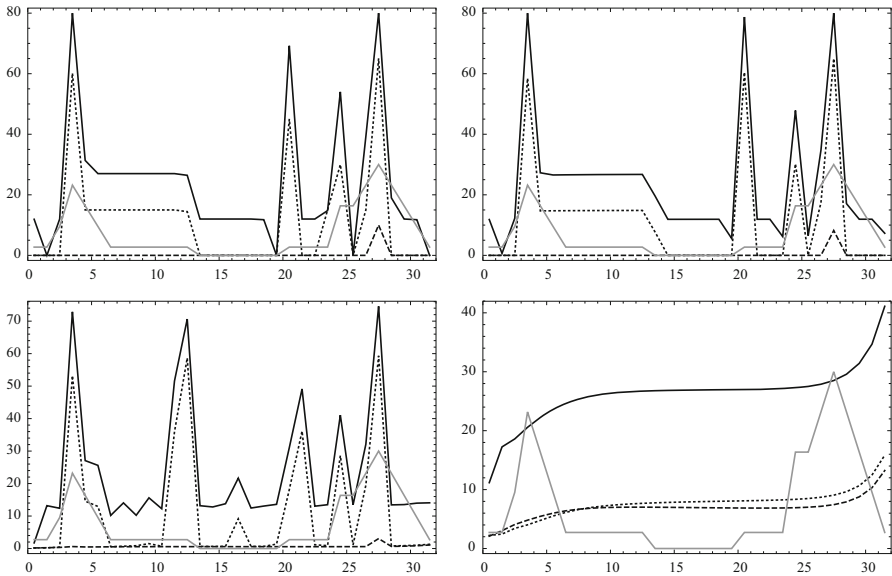
The reason to include the first two problems into the study is that they are simple alternatives to overcome algorithmic difficulties with joint probabilistic constraints. Both models allow a reduction of the original stochastic inequality system to a linear one. Hence, linear programming is all what is needed to find a solution of the respective problems. At the same time, the release policies found in both models lack robustness with respect to the uncertain inflow process. While this is well known for the expected-value problem, it seems to be less recognized for the case of individual constraints. The consideration of the ‘max- $p$ ’-problem is useful for two reasons. Its optimal solution  $(x^*, p^*)$  provides us at the same time with a Slater point  $x^*$  needed for the cutting plane algorithm in the solution of (4) and with the maximum possible probability level  $p$  which would still allow for a nonempty set of feasible release policies. This information is important for the decision maker for fixing an appropriate (not too large) level  $p$  in his/her original problem (4).

We shall consider a time horizon of  $T = 32$  intervals of 15 min each. For the discrete random inflow processes to the two reservoirs we rely on the quantities indicated in Fig. 1 as expected values. We assume that all components of the whole random process  $(\xi^{(1)}, \xi^{(2)})$  are normally distributed with the indicated expected values and with a relative standard deviation of 2.5% for the first and 2% for the second reservoir. Within each reservoir, inflows are assumed to be correlated over time by a band structure. They are assumed to be independent between the two reservoirs. Hence, we are considering a 64-dimensional random vector having a multivariate normal distribution with non-diagonal correlation matrix. We fix the probability level as  $p = 0.9$ . For the efficiencies of the turbines, we rely on the ratios between values for the three turbines as indicated in Fig. 1:

$$\lambda^{(1)} = 0.638, \quad \lambda^{(2)} = 0.741, \quad \lambda^{(3)} = 0.694.$$

Moreover, the time-dependent price vector  $\pi$  was determined as

$$\pi = (2.5, 2.5, 3.75, 6.25, 5.0, 3.75, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.0, 2.0, 2.0, 2.0, 2.0, 2.0, 2.5, 2.5, 2.5, 2.5, 5.0, 5.0, 6.25, 7.5, 6.25, 5.0, 3.75, 2.5).$$

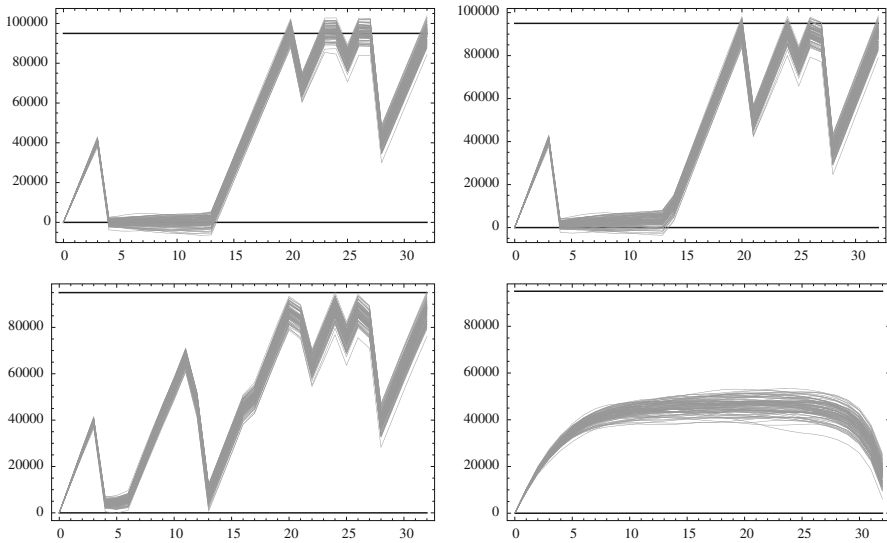


**Fig. 2** Illustration of the optimal release policies for the three turbines (*dashed* = 1, *dotted* = 2, *solid* = 3) in case of the expected-value problem (*top left*), the problem with individual probabilistic constraints (*top right*), the problem with joint probabilistic constraints (*bottom left*) and the ‘max-p’ problem (*bottom right*). The gray line represents the profile of the price vector but with a large offset cut off for graphical reasons

Finally, for the evaluation of end levels, we have chosen the coefficients

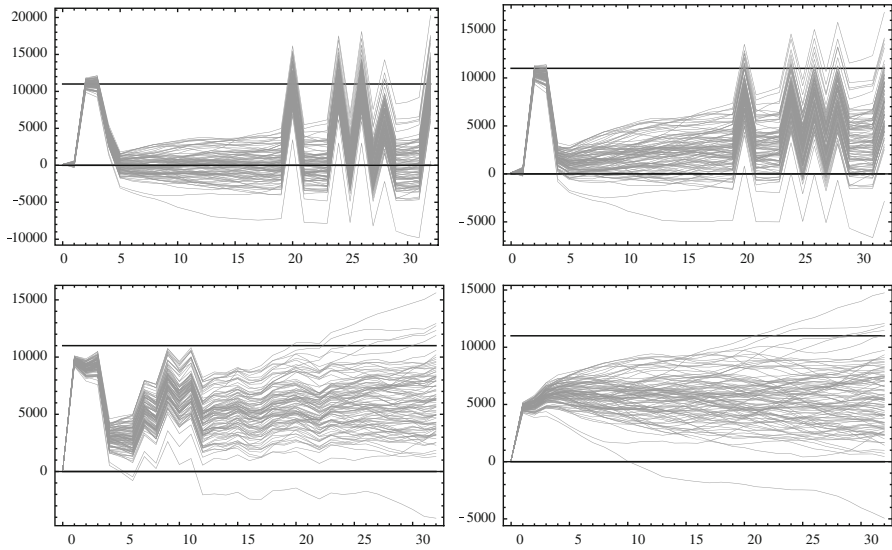
$$\omega_1 = 20.0, \quad \omega_2 = 15.0.$$

Calculation of normal probabilities of rectangles as well as of their derivatives (again requiring normal probabilities of rectangles via Theorem 1) was carried out by using Genz’ Fortran code MVNDST (Genz 1992). Figure 2 shows the optimal release policies for the three turbines for all four problems solved. It can be seen that in the first three problems the second and third turbine have their peaks at high prices as a result of profit maximization. Additional peaks guarantee the feasibility of the level constraints. The first turbine remains almost inactive because it is less efficient and so the second one can do a better job unless it reaches its operational limit. The profiles of the first two problems are very similar. A more detailed inspection reveals that different operational levels can be afforded e.g. in the intervals 5, 20, 21, 24, 25, 26 and 32. The release policy under joint probabilistic constraints exhibits some differences: the peaks are a bit smaller (recall different scales in the figures), while there is an additional peak in intervals 12 and 13 and the base level for the third turbine is higher than in the previous two cases. Not surprisingly, the most robust release policy looks completely different, because it does not aim at maximizing profit and, hence, does not follow the pattern of the price vector. On the other hand, its maximum probability level was found as  $p^* = 0.95$ , a value that cannot be exceeded by any release policy.



**Fig. 3** Illustration of 100 filling level scenarios for the first reservoir with the arrangement of figures corresponding as in Fig. 2 to the 4 problems investigated

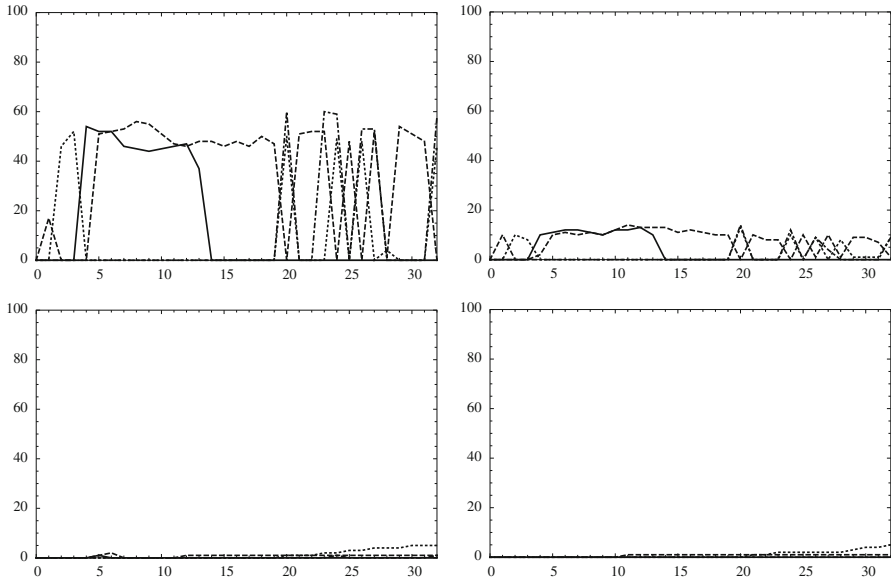
In order to validate our results and, in particular, to check the probabilities of keeping the reservoir levels, we generated a set of 100 inflow scenarios for both reservoirs according to the chosen distribution parameters of the two inflow processes (see above). Now, given these simulated inflow profiles and the calculated optimal release policies, one can visualize the resulting corresponding filling levels in the two reservoirs. Figure 3 shows the results for the upper reservoir for the four different problems. As one can see, the filling levels are very well respected in the problem with joint probabilistic constraints and in the ‘max-p’ problem. In contrast, there are several violations in the problem with individual probabilistic constraints (at time intervals 4–13, 20, 24, 26, 27, 32) and even many more violations in the expected-value problem (at time intervals 4–13, 20, 23, 24, 26, 27, 32). This is not surprising because in the latter problem, constraint satisfaction is required in expectation only. Things become a lot more significant for the second (much smaller) reservoir (see Fig. 4). The filling levels are violated in all intervals by the problem with individual probabilistic constraints as well as by the expected-value problem. It is obvious that the expected-value solution cannot be used in practice. The problem under individual probabilistic constraints seems to lead to reasonable level satisfaction consistent with the chosen probability  $p = 0.9$ . However, this holds true only if time intervals and reservoir levels are considered **separately**. It is hard to visually detect from Figs. 3 to 4 how many of these 100 scenarios really pass through the feasible range during the whole time horizon in case of individual probabilistic constraints. There is no doubt about that in the remaining two cases (joint probabilistic constraints and ‘max-p’) only very few trajectories ever violate a level constraint during the whole time horizon. More insight about the probability of level constraint violation can be gained by plotting the numbers of scenarios violating any of the four constraints (first



**Fig. 4** Illustration of 100 filling level scenarios for the second reservoir with the arrangement of figures corresponding as in Fig. 2 to the 4 problems investigated

reservoir—upper level, first reservoir—lower level, second reservoir—upper level, second reservoir—lower level) as functions of time. This corresponds to an empirical probability of constraint violation. The resulting plots are provided in Fig. 5, where different plot styles are used for different level constraints. The already observed uselessness of the expected-value solution is clearly demonstrated by the large number of time intervals at which violations of the level constraints occur with (empirical) probability of up to 60%. In contrast, the solution based on individual probabilistic constraints keeps what it promises: at each time interval, any of the level constraints is violated at most in approximately 10% of the scenarios which fits very well the chosen probability level  $p = 0.9$ . The remaining two solutions (joint probabilistic constraints and ‘max- $p$ ’) are almost at zero level.

However, at this point one might get curious to know how many of the scenarios really pass through the whole time horizon without violating any of the level constraints. These results are summarized in Table 1 which also contains the strictly calculated probability of satisfying the joint probabilistic constraints and the corresponding value of the objective. The resulting failure numbers are: 100 (!) in case of expected values, 68 in case of individual probabilistic constraints, 10 in case of joint probabilistic constraints and 6 in case of the robust solution provided by the ‘max- $p$ ’ problem. This means that almost surely (in the literal and probabilistic sense) at least one of the filling level constraints will be violated at least once during the time horizon when using the release policies under expected-value constraints. More interestingly, release policies based on individual probabilistic constraints are not much better. While for each time interval separately, the probability of satisfying level constraints fits well the chosen level  $p = 0.9$ , the probability of satisfying them simultaneously over the whole time horizon reduces to approximately 0.32 which is highly undesirable and



**Fig. 5** Plot of the empirical probabilities for level constraint violation as a function of time (*solid* = lower limit of upper reservoir, *dotdashed* = upper limit of upper reservoir, *dashed* = lower limit of lower reservoir, *dotted* = upper limit of lower reservoir). The arrangement of figures corresponds as in Fig. 2 to the 4 problems investigated

**Table 1** Probability levels and optimal values

Problem	#Violating scenarios	Probability level	Optimal value
Max $p$	6	0.95	3135163
Joint constr.	10	0.90	4803072
Individual constr.	68	0.32	4954845
Expected value	100	0.00	5118385

explains, why individual probabilistic constraints are not the right model in our application though they are appealing due to their easy to obtain solution. The number of 10 violating scenarios in the model with joint probabilistic constraints corresponds well now to the chosen probability level  $p = 0.9$ . The same is true in case of the ‘max-p’ problem, where the maximum possible probability level  $p^*$  was calculated as 0.95.

Finally, the optimal values recorded in Table 1 allow the following two important observations: first, a robust release policy with a probability level as high as 0.9 can be guaranteed by accepting a negligible loss of profit as compared to the two simplified models which are not robust at all. Second, insisting on maximum robustness (0.95 here) may come at the price of a significant loss of profit.

**References**

Andrieu L, Henrion R, Römisch W (2009) A model for dynamic chance constraints in hydro power reservoir management, Preprint 536, DFG Research Center MATHEON “Mathematics for key technologies”

- Deák I (1980) Three digit accurate multiple normal probabilities. *Numerische Mathematik* 35:369–380
- Genz A (1992) Numerical computation of multivariate normal probabilities. *J Comp Graph Stat* 1:141–149
- Genz A, Bretz F (2009) Computation of multivariate normal and t probabilities. *Lecture Notes in Statistics*, vol. 195. Springer, Dordrecht
- Genz A (1992) <http://www.math.wsu.edu/faculty/genz/homepage>
- Henrion R, Römisch W (2005) Lipschitz and differentiability properties of quasi-concave and singular normal distribution functions, to appear in: *Ann Oper Res*, appeared electronically under doi:[10.1007/s10479-009-0598-0](https://doi.org/10.1007/s10479-009-0598-0)
- Prékopa A (1973) On logarithmic concave measures and functions. *Acta Scientiarum Mathematicarum (Szeged)* 34:335–343
- Prékopa A (1995) *Stochastic programming*. Kluwer, Dordrecht
- Prékopa A (2003) Probabilistic programming, Chap 5. In: Ruszczyński A, Shapiro A (eds) *Stochastic programming*. *Handbooks in Operations Research and Management Science*, vol 10. Elsevier, Amsterdam
- Szántai T (2000) Improved bounds and simulation procedures on the value of the multivariate normal probability distribution function. *Ann Oper Res* 100:85–101
- Uryasev S (1995) Derivatives of probability functions and some applications. *Ann Oper Res* 56:287–311