Optim Eng (2014) 15:509–531
DOI 10.1007/s11081-013-9236-4

Joint chance constrained programming for hydro reservoir management

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

Received: 22 September 2011 / Accepted: 5 July 2013 / Published online: 18 October 2013

Abstract In this paper, we deal with a cascaded reservoir optimization problem with uncertainty on inflows in a joint chance constrained programming setting. In particular, we will consider inflows with a persistency effect, following a causal time series model with Gaussian innovations. We present an iterative algorithm for solving similarly structured joint chance constrained programming problems that requires a Slater point and the computation of gradients. Several alternatives to the joint chance constraint problem are presented. In particular, we present an individual chance constraint problem and a robust model. We illustrate the interest of joint chance constrained programming by comparing results obtained on a realistic hydro valley with those obtained from the alternative models. Despite the fact that the alternative models often require less hypothesis on the law of the inflows, we show that they yield conservative and costly solutions. The simpler models, such as the individual chance constraint one, are shown to yield insufficient robustness and are therefore not useful. We therefore conclude that Joint Chance Constrained programming appears as an approach offering a good trade-off between cost and robustness and can be tractable for complex realistic models.

W. van Ackooij (✉) · R. Zorgati
EDF R&D. OSIRIS, 1, avenue du Général de Gaulle, 92141 Clamart Cedex, France
e-mail: wim.van-ackooij@edf.fr

R. Zorgati
e-mail: riadh.zorgati@edf.fr

R. Henrion · A. Möller
Weierstrass Institute Berlin, Mohrenstr. 39, 10117 Berlin, Germany

R. Henrion
e-mail: henrion@wias-berlin.de

A. Möller
e-mail: moeller@wias-berlin.de

Keywords Chance constrained programming · Hydro reservoir management · Joint chance constraints · Stochastic inflows

1 Introduction

An important optimization problem in energy management, known as the "unit-commitment Problem", aims at computing the production schedule that satisfies the offer-demand constraint at minimal cost. That schedule indicates production levels for each production unit in a hydro-thermal system. Each unit is subject to many complex technical constraints. This fact, together with the fact that the offer-demand constraints are coupling constraints and link all these various and numerous units together, leads to the conclusion that the unit-commitment problem is often large-scale and difficult to solve. In order to tackle these large scale problems, the coupling constraints are often dualized, using Lagrangian techniques, leading to an effective price decomposition scheme (Cohen and Zhu 1983; Lemaréchal and Sagastizábal 1994). Since the global unit-commitment problem is already challenging to solve in a deterministic setting due to its non-convex feasible sets and large scale, uncertainty is often neglected, even though decisions are taken at least one day in advance. Uncertainty in unit-commitment problems comes at least from the following sources: customer load, renewable generation, inflows, unit availability. Integrating uncertainty in global unit-commitment will be quite challenging for the reasons outlined above. Hence, as a first necessary step, we will focus on hydro valley optimization. In the Lagrangian dualization setting of a unit-commitment problem, hydro valley optimization can be seen as a sub-problem. Alternatively, one can interpret this problem as an optimization against market-prices. Complex dynamic constraints on watershed controls introduce combinatorial aspects in this sub-problem, making it difficult to solve. The focus of this paper will therefore be on integrating uncertainty in hydro valley management.

The aforementioned combinatorial aspects result from formulating smoothness requests on watershed. From an engineering perspective it is undesirable to have turbining output increase and decrease rapidly over short time spans as this induces a strain on material. Other combinatorial elements can arise when modelling very realistic efficiency curves. We refer to Diniz and Maceira (2008) for an approach to deal (i.e., remove) the latter combinatorial elements. In hydro dominated systems, such as in Brazil, Scandinavia and Canada, the emphasis of accurate modelling lies on hydro generation and combinatorial optimization is common for cascaded reservoir management. We refer to Belloni et al. (2003), Finardi and Da Silva (2006), Ponrajah et al. (1998), Nilsson and Sjelvgren (1997) for more details on such models. In thermal dominated systems, such as the French system, the modelling emphasis lies more on thermal generation. In these large-scale unit-commitment problems, such additional combinatorial elements are often neglected in order to have an acceptable computional burden (see Dubost et al. 2005). We will make the same assumption. Integrating uncertainty and combinatorial elements in a cascaded reservoir model is quite challenging and will be investigated in future work. A potential entry point for such an approach would be the decomposition idea investigated in Finardi and Da Silva (2006).

Uncertainty in cascaded reservoir management results from uncertainty on inflows and impacts the physical constraints of the system. Since decisions are taken prior to the observation of uncertainty, appropriate modelling approaches for integrating uncertainty have to be considered. The two main approaches are chance constrained programming and robust optimization. The main focus of this paper is on the former, for the latter we refer to Ben-Tal et al. (2009). In dynamic decision processes, i.e., when decisions in later time periods are allowed to adapt to earlier observed uncertainty, the main approaches are Stochastic Dynamic Programming and SDDP (see Pereira and Pinto 1991). Often the convenient hypothesis is made that uncertainty within the transition problem is known. This essentially makes the transition problem a deterministic problem. The latter choice is especially questionable when the transition problem covers a time span of a week such as assumed in Philpott et al. (2011).

Introduced by Charnes and Cooper (1960), probability constraints are quite an appealing tool for dealing with uncertainty. In particular, when uncertainty arises in physical constraints, since they also offer a simple interpretation. A classical introduction to the theory and numerical treatment of chance constraints can be found in Prékopa (1995). In the same monograph, one can find convexity results, for uncertainty separated from decisions and for a large class of distributions including the multivariate Gaussian one. Since their first introduction, chance constraints have become quite common in hydro valley management (Loucks et al. 1981; Duranyildiz et al. 1999; Edirisinghe et al. 2000; Loiaciga 1988; Morgan et al. 1993; Zorgati et al. 2009; Zorgati and van Ackooij 2011; van Ackooij et al. 2010), but often individual chance constraints are used and not joint chance constraints. Though a very appealing approximation, individual chance constraints unfortunately do not offer sufficient robustness (see van Ackooij et al. 2010). Hydro reservoir models with joint chance constraints have been considered, for instance, in Prékopa and Szántai (1978a, b), van Ackooij et al. (2010). In Andrieu et al. (2010) even a dynamic approach has been developed in this context. However, these models were comparatively simple from their structure (no serially linked reservoirs, no delay time between reservoirs, no realistic water value condition, no time series modelling of statistical data, small dimension). The main focus of van Ackooij et al. (2010) is on deriving an efficient gradient formula for joint chance constraints of a specific form. The latter form arises naturally in hydro reservoir management. The interest of the formula is then illustrated on a stylized hydro reservoir optimization problem.

The contributions of this paper to the field of applied chance constrained programming are the following. We consider a realistically sized hydro valley from the industry, wherein water is valued according to a structure arising in practice. Uncertainty is modelled according to a causal time series model integrating a persistency effect on inflows. This offers a wide range of modelling choices. We compare the chance constrained programming approach with other typical approaches tempting to solve the same problem: the well established practice in the industry consisting of replacing the random vector by its expectation, approximating joint chance constraints by individual chance constraints and robust optimization. We show that the former two approaches are insufficient since they offer no robustness guarantee and that the latter one often yields over-conservative solutions. Finally, we derive an efficient formula for the Hessian of a joint chance constraint that could be exploited in smooth non-linear optimization algorithms for joint chance constrained programming.

This paper is organized as follows. In Sect. 2, we present our model for hydro reservoir management, where combinatorial constraints are neglected and random inflows are introduced. We give a detailed description of a real hydro valley, and present the main optimization problem. As the uncertainty on inflows is concerned, many statistical models are based on a deterministic trend (potentially dependent on explanatory variables) and a causal noise process. Since convexity results exist for specific classes of randomness and in particular Gaussian ones, it seems tempting to place ourselves in such a setting. Restricting uncertainty laws to such a setting, might seem restrictive at first. However, we will show that a large class of models is available, i.e., the class of causal time series models with Gaussian innovations (Shumway and Stoffer 2000).

In Sect. 3, we derive algorithms for dealing with our model. One difficulty for solving joint chance-constraint models is to be able to compute gradients of such constraints efficiently. Gradient formulæ for multivariate Gamma, Dirichlet, and Gaussian distributions can be found in Prékopa and Szántai (1978b), Prékopa (1995), Gouda and Szántai (2010), Szántai (1985), Henrion and Möller (2012), van Ackooij et al. (2010), respectively. Similarly to evaluating the chance constraint, these formulæ involve computing a probability. We present here an iterative algorithm for solving, in the convex case, joint chance constrained programming problems that requires a Slater point and the computation of gradients. Several alternatives to the joint chance constraint problem are then presented. In particular, we present an individual chance constraint problem and a robust model.

In Sect. 4, we report results obtained when solving these various models on a realistic instance of a hydro valley management problem. The interest of joint chance constrained programming is illustrated by comparing results obtained on this hydro valley with those obtained from the alternative models.

An algorithmic perspective and some auxiliary lemmas are given in Appendix A. We also provide a modest extension of the theoretical results obtained in van Ackooij et al. (2010), since it presents an efficient formula for the Hessian of a joint chance constraint. Finally, conclusions are drawn in Sect. 5.

2 Problem description

In this section we will give a description of the hydro reservoir management problem. We will consider a discretized time horizon. To this end let $\tau = \{1, \dots, T\}$ denote the set of (homogeneous) time steps, where T denotes the last time step. Let Δt be the time step size expressed in hours. We will begin by providing problem constraints and the objective function. We will conclude with a paragraph highlighting the structure of the problem.

2.1 Problem constraints

2.1.1 Topology

A hydro valley can be seen as a set of connected reservoirs and associated turbines. We can therefore represent this with a directed graph. Let $\mathcal{N}$ be the set of nodes and let A (of size $|\mathcal{N}| \times |\mathcal{N}|$) be the connection matrix, i.e., $A_{n,m} = 1$ whenever water released from reservoir n will flow into reservoir m. We will assume that D is the flow duration vector, i.e., D_m is the amount of time (measured in time steps) it takes for water to flow from upper reservoir m to its unique child. It is assumed that pumping is (nearly) instantaneous. Let $\mathcal{T} := \{g^i, i = 1, \dots, N_{\mathcal{T}}\}$ denote the set of turbines and $\mathcal{P} := \{p^i, i = 1, \dots, N_{\mathcal{P}}\}$ denote the set of pumping stations. We furthermore introduce the mapping $\sigma_{\mathcal{T}} : \{1, \dots, N_{\mathcal{T}}\} \to \mathcal{N}$ ($\sigma_{\mathcal{P}} : \{1, \dots, N_{\mathcal{P}}\} \to \mathcal{N}$) attributing to each turbine (pumping station) the reservoir number to which it belongs. We will also introduce the sets $\mathcal{A}(n) = \{m \in \mathcal{N} : A_{m,n} = 1\}$ and $\mathcal{F}(n) = \{m \in \mathcal{N} : A_{n,m} = 1\}$. The set $\mathcal{A}(n)$ is empty for uphill reservoirs and the set $\mathcal{F}(n)$ for downhill reservoirs. To each reservoir $n \in \mathcal{N}$ and for each time step $t \in \tau$ we associate its volume $V^n(t)$ in cubic hectometers hm^3. The initial volume of each reservoir $n \in \mathcal{N}$ is denoted by $V^n(0)$, lower and upper bounds are $V^n_{min}(t)$ and $V^n_{max}(t)$ respectively.

2.1.2 Controls

We will assume that each turbine (and pumping station) can be controlled for each time step. To this end we introduce the variables $x^i(t)$ for each $t \in \tau$ and $i = 1, \dots, N_{\mathcal{T}}$. In a similar way we introduce the variables $y^i(t)$ for the pumping stations. The units are in cubic meters per hour, i.e., m^3/h. Furthermore we assume that each of these variables are bounded from below by zero and from above by $\overline{x}^i$ ($\overline{y}^i$ respectively).

2.1.3 Random inflows

We will assume that inflows (in m^3/h) in reservoirs are the result of some stochastic process. Let $A^n(t)$ denote this stochastic process for reservoir n. Not all reservoirs will have stochastic inflows, some of them will have deterministic inflows. This can be explained by the fact that top reservoirs have random inflows due to the melting of snow in the high mountains, whereas rain can be neglected for lower reservoirs. Let $\mathcal{N}^r \subseteq \mathcal{N}$ denote the set of reservoirs receiving random inflows. We will assume that the stochastic inflow process is the sum of a deterministic trend s^n_t and a causal process (Shumway and Stoffer 2000) generated by Gaussian innovations. To this end, let $\zeta^n(t)$ be a Gaussian white noise process, where $(\zeta^{k_1}(t), \dots, \zeta^{k_l}(t))$ is a Gaussian random vector of zero average and variance-covariance matrix $\Sigma(t)$ ($\{k_1, \dots, k_l\} = \mathcal{N}^r$). We will assume independence between time steps of the ζ vector. Since $A^n(t)$ is a causal process, we can write it as follows

$$A^n(t) = s^n_t + \sum_{j=0}^{\infty} \psi^n_j \zeta^n(t-j) = s^n_t + \sum_{j=t}^{\infty} \psi^n_j \zeta^n(t-j) + \sum_{j=0}^{t-1} \psi^n_j \zeta^n(t-j),$$
$$\forall n \in \mathcal{N}^r, t \in \tau$$

for some coefficient vector ψ^n and infinite past before $t = 0$ (the beginning of the optimization horizon). We will assume that randomness before (and including) $t = 0$ is known and as such we can assume w.l.o.g. that the random inflow process can be written as

$$A^n(t) = s_t^n + \sum_{j=0}^{t-1} \psi_j^n \zeta^n(t-j), \quad \forall n \in \mathcal{N}^r, t \in \tau. \tag{1}$$

For reservoirs $n \in \mathcal{N} \setminus \mathcal{N}^r$, we simply have $A^n(t) = s_t^n$.

2.1.4 Flow constraints and volume bounds

Each reservoir is subject to flow constraints induced by pumping and turbining. The following balance constraint applies

$$\begin{aligned} V^n(t) = V^n(t-1) &+ \sum_{m \in \mathcal{A}(n)} \sum_{i \in \sigma_{\mathcal{T}}^{-1}[m]} x^i(t - D_m)\Delta t - \sum_{i \in \sigma_{\mathcal{T}}^{-1}[n]} x^i(t)\Delta t \\ &+ \sum_{m \in \mathcal{F}(n)} \sum_{i \in \sigma_{\mathcal{P}}^{-1}[m]} y^i(t)\Delta t - \sum_{i \in \sigma_{\mathcal{P}}^{-1}[n]} y^i(t)\Delta t + s_t^n \Delta t \\ &+ \sum_{j=0}^{t-1} \psi_j^n \zeta^n(t-j)\Delta t, \quad \forall t \in \tau, n \in \mathcal{N}. \end{aligned} \tag{2}$$

The above equation is entirely deterministic except for the reservoirs $n \in \mathcal{N}^r$. In order to deal with this randomness and reservoir bounds we will therefore add the following constraints

$$\mathbb{P}\big[V_{min}^n(t) \le V^n(t) \le V_{max}^n(t)\ \forall t \in \tau, n \in \mathcal{N}^r\big] \ge p \tag{3}$$

$$V_{min}^n(t) \le V^n(t) \le V_{max}^n(t)\ \forall t \in \tau, n \in \mathcal{N} \setminus \mathcal{N}^r, \tag{4}$$

where $\mathbb{P}$ is a probability measure and p a security level. Constraint (3) is a joint chance constraint. This means that we wish to satisfy all linear inequalities of the stochastic system simultaneously with high enough probability. This can be compared to a model with individual chance constraints, which is a model wherein we wish to satisfy each inequality with high enough probability, but taken separately. We will see in this paper that the latter model offers insufficient robustness.

2.1.5 Water values

In short term optimization problems (with time horizons ranging from several days up to a month) water values provide a way to associate a cost with used water. Incorporating no such cost in a short term optimization problem would inevitably lead to a maximum use of water on this specific time horizon, whereas water might be needed in later time periods. Water might be used to reduce the use of costly thermal generation or as a security to avoid “black-outs” in difficult situations. Water values are obtained as the by product of (stochastic) dynamic programming approaches in mid term (time horizons ranging from 1 to 5 years).

In full generality water values depend on time, a multivariate random vector, the current water levels in all reservoirs and other quantities that can be considered as inventories or stocks (such as customer interruption options (see Zorgati and van Ackooij 2011 for more details), i.e., an inventory globally very similar to the number of remaining exercise rights in swing options). As the effect of uncertainty is concerned, it is often averaged out on a set of reasonable scenarios in order to integrate unconditional water values in short term optimization. The stochastic dynamic programming algorithms typically deal with uncertainty effects rarely integrated in short term optimization such as stochastic fuel prices.

The multivariate stock dependency is only known approximately, if at all, since one quickly hits the curse of dimensionality of dynamic programming. In such cases, approaches such as approximate dynamic programming (ADP) (de Farias and Van Roy 2002), approximate dual dynamic programming (ADDP) (Girardeau 2010), SDDP (Pereira and Pinto 1991; Philpott and Guan 2008) or aggregation approaches (Turgeon 1980; Torrion and Leveugle 1985) are applied in order to approximately solve the dynamic programming problem. In the ADP approach, it is commonly assumed that the continuation function of dynamic programming decomposes as a sum of 1 dimensional functions. Each function depending on a unique stock only. This then automatically results in single stock dependent water values. Even if water values would be available as multivariate functions, they would only be known on a set of grid points. If this is to be incorporated in short term optimization one surely needs interpolation techniques very similar to those explained in d'Ambrosio et al. (2010). This interpolation approach leads to the introduction of binary variables in the optimization problem. Since multivariate effects in water values are only rarely known and integrating them induces combinatorial aspects, we will focus on single stock dependent water values in this paper.

As the temporal dependency is concerned it is often daily or intra-daily. Due to the average effect of climate on unit-commitment, some specific weeks are far more costly than surrounding weeks. Such weeks have peaking customer load and high risk of black outs. Such effects get reflected in the water values as well. These effects are moreover strengthened by averaging out stochastically dependent water values as explained above.

If we wish to incorporate water values in short term optimization, the latter temporal effect can either be neglected or taken into account. In the first case, we would value the differential between the end and the initial volume of a reservoir against water values at that time step. In the second approach we would either value volumes against water values at each time step or value local volumetric differences. The first approach would consider indifferently any two storage paths leading to the same end volume. When the short term time horizon is close to a month and one of the above difficult weeks is within this time horizon, from an operational view point two paths leading to the same end volume are not necessarily considered equivalent. It is therefore of interest to integrate the temporal dependence in order to reflect this feature. A second reason for integrating this effect is to provide a model that fits better with current practice. In practice, in order to control the storage path, a selection of time steps $\tau^s \subseteq \tau$ is made where artificially we force $V^n_{min}(t) \approx V^n_{max}(t)$ for $t \in \tau^s$. Integrating the temporal dependency of water values in short term optimization is a natural way to have control over the storage path without risking to have an empty feasible set.

In this section we present a model for incorporating water values without reflecting temporal dependencies as the focus of the paper is on Chance Constrained programming for hydro reservoir management. Upon valuing the volume at each time step against local water values, the presented model allows for a straightforward extension for incorporating the above discussed temporal effect of water values.

Volume dependent water values Our aim is to set up a model which evaluates the expected amount of water in the reservoir at the end of the optimization horizon[1]. This is necessary in order not to carry out the optimization at the expense of later periods of time. A possible way to do so is to subdivide the levels of each reservoir into a finite number of values from bottom to top as follows:

$$V^n_0, \dots, V^n_{K_n} \quad \forall n \in \mathcal{N}.$$

Each compartment $[V^n_{i-1}, V^n_i)$ is assigned a water value W^n_i (in €/m^3) such that

$$W^n_{i-1} > W^n_i \geq 0 \quad \forall n \in \mathcal{N} \ \forall i = 1, \dots, K_n. \tag{5}$$

The value of the expected final water level $\mathbb{E}(V^n(T))$ of reservoir n is then simply the cumulative value of water in the compartments below:

$$\sum_{i \leq i^*} W^n_i \big(V^n_i - V^n_{i-1}\big) + W^n_{i^*}\big(\mathbb{E}\big(V^n(T)\big) - V^n_{i*}\big), \quad i^* := \max\big\{i \,|\, \mathbb{E}\big(V^n(T)\big) \geq V^n_i\big\}.$$

Note that this value is an increasing function of the expected final level $\mathbb{E}V^n(T)$ despite the fact that water values are strictly decreasing from bottom to top.

Now, in order to avoid combinatorial arguments concerning the index i^*, we introduce auxiliary variables z^n_i indicating for each reservoir n the amount of water in compartment $[V^n_{i-1}, V^n_i)$. Of course, since all compartments have to be completely filled up to i^*, one has that

$$z^n_i = \begin{cases} V^n_i - V^n_{i-1} & i = 1, \dots, i^* \\ \mathbb{E}(V^n(T)) - V^n_{i*} & i = i^* + 1 \\ 0 & i = i^* + 2 \dots, K_n \end{cases} \qquad \forall n. \tag{6}$$

Then, the value of the final water level in reservoir n equals

$$\sum_{i=1}^{K_n} W^n_i z^n_i \quad \forall n. \tag{7}$$

[1] In practice, one would evaluate the difference of the final and initial volume. The latter adds a constant to the objective function and can theoretically be omitted. In practice, it may generate some numerical difficulties, especially when large volumes are valued and turbining/pumping capacity is small compared to the volume. In that case, relative changes in valuation induced by the controls are easily considered negligible. Moreover, the constant can easily be added.

We claim that the relations (6) for variables z_i^n can be replaced by the following relations in which the crucial index i^* is absent:

$$\sum_{i=1}^{K_n} z_i^n = \mathbb{E}\big(V^n(T)\big) - V_0^n \quad \forall n \tag{8}$$

$$0 \leq z_i^n \leq V_i^n - V_{i-1}^n \quad \forall n \forall i = 1, \ldots, K_n. \tag{9}$$

The argument is as follows: as part of the overall objective function in our problem, we shall maximize the value of the final water level (7). Given the strictly decreasing order of water levels in (5) (from bottom to top), it is clear from (8) that the upper inequality in (9) will be satisfied as an equality as long as possible and that only the most upper compartment may not be completely filled. This of course is equivalent with (6) but avoiding the explicit description of that most upper compartment.

Since the initial volume $V^n(0)$ is known in advance, one can define variables $z_{0,i}^n$ in a similar way as z_i^n. It then follows that

$$\sum_{i=1}^{K_n} W_i^n \big(z_{0,i}^n - z_i^n\big) \tag{10}$$

is the cost of used water for reservoir $n \in \mathcal{N}$. The valuation induced by $\sum_{i=1}^{K_n} W_i^n z_{0,i}^n$ is in fact a constant and can be omitted.

2.2 Objective function

Often, in reality, each reservoir only has a single turbine. The power output of turbining x, in cubic meters per second m^3/s, is given by a function $\rho(x)$. This function is strictly increasing and concave, i.e., $\rho'(x) \geq 0$ and $\rho''(x) \leq 0$. In our model we have split this range into several subsections (hence several turbines), each with efficiency $\rho_i = \rho'(s_i^*)/3600$ ($\mathrm{MWh}/\mathrm{m}^3$) for some s_i^* in each section. We can thus remark that for any two turbines i_1 and i_2 belonging to the same reservoir we have $\rho_{i_1} \geq \rho_{i_2}$ whenever $i_1 \leq i_2$. This approximation comes down to approximating $\rho(x)$ by a piece-wise linear function.

We assume given a time dependent price signal $\lambda(t)$ (in €/MWh). The following objective function has to be minimized, when integrating the cost of used water according to (10):

$$\sum_{n\in\mathcal{N}} \sum_{i=1}^{K_n} \big(W_i^n \big(z_{0,i}^n - z_{F,i}^n\big) - \sum_{t\in\tau} \lambda(t)\Delta t \left(\sum_{i=1}^{N_\mathcal{T}} \rho_i(t) x^i(t) - \sum_{i=1}^{N_\mathcal{P}} \frac{1}{\theta_i(t)} y^i(t)\right), \tag{11}$$

where, $\theta^i(t)$ is the efficiency of pumping and the auxiliary variables z_i^n satisfy (8), (9).

2.3 Matrix formulation

In this section we show that (3) can be written as bilateral joint chance constraint. This means that the model we are interested in is a bilateral joint chance constrained program with linear objective function and some polyhedral constraints.

Let us consider (2) and apply it recursively to establish the identity

$$
\begin{aligned}
V^n(t) = V^n(0) &+ \sum_{u=1}^{t} \sum_{m \in \mathcal{A}(n)} \sum_{i \in \sigma_{\mathcal{T}}^{-1}[m]} x^i(u - D_m)\Delta t - \sum_{u=1}^{t} \sum_{i \in \sigma_{\mathcal{T}}^{-1}[n]} x^i(u)\Delta t \\
&+ \sum_{u=1}^{t} \sum_{m \in \mathcal{F}(n)} \sum_{i \in \sigma_{\mathcal{P}}^{-1}[m]} y^i(u)\Delta t - \sum_{u=1}^{t} \sum_{i \in \sigma_{\mathcal{P}}^{-1}[n]} y^i(u)\Delta t \\
&+ \sum_{u=1}^{t} s_u^n \Delta t + \sum_{u=1}^{t} \sum_{j=0}^{u-1} \psi_j^n \zeta^n(u - j)\Delta t,
\end{aligned} \tag{12}
$$

holding for all $t \in \tau$ and $n \in \mathcal{N}$. In what follows we will denote with $V_n \in \mathbb{R}^T$ the vector $V_n = (V^n(1), \ldots, V^n(T))$. It is of interest to explicitly establish the way in which V_n depends on the vector ζ^n in order to identify the correlation structure of the global underlying uncertainty vector. One easily observes that V_n depends linearly on x and y. In order to establish the correlation structure of the vector ζ, we introduce the matrix mapping $\mathfrak{C} : \mathbb{R}^T \to \mathcal{M}_{T \times T}$. Here $\mathcal{M}_{T \times T}$ stands for the set of $T \times T$ real matrices and $\mathfrak{C}$ as applied to the sequence $\psi := (\psi_0, \ldots, \psi_{T-1}) \in \mathbb{R}^T$ is defined as:

$$
\mathfrak{C}(\psi) = \begin{pmatrix} \psi_0 & 0 & 0 & \cdots & 0 \\ \psi_0 + \psi_1 & \psi_0 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \sum_{j=0}^{T-1} \psi_j & \cdots & & \cdots & \psi_0 \end{pmatrix}.
$$

It will be convenient to extend the definition of $\mathfrak{C}$ to a sequence ψ shorter than T by appending with zero entries.

Following (12) for each $n \in \mathcal{N}$ we can find a $T \times T N_{\mathcal{T}}$ matrix $M_{\mathcal{T}}^n$ and $T \times T N_{\mathcal{P}}$ matrix $M_{\mathcal{P}}^n$ such that

$$
V_n = V_0 - \Delta t M_{\mathcal{T}}^n x + M_{\mathcal{P}}^n y + \Delta t \mathfrak{C}(1) s^n + \Delta t \mathfrak{C}\big(\psi^n\big)\zeta^n, \tag{13}
$$

where s^n is the vector formed from the deterministic trend s_t^n of (1). Equations (8), (9) can be written easily in linear form by extracting the last line from (13) without the term in ζ^n.

3 Models for dealing with uncertainty

In this section we will provide our main model, which is a joint chance constrained programming problem (JCCP). We will also provide several alternative models.

3.1 Expectation model

In a classic version of cascaded reservoir management in short term optimization, uncertainty is assumed to be absent or sufficiently characterized by a forecast. This amounts to the choice of replacing ζ^n in (13) or equivalently (1) by its expectation, i.e., $\zeta^n(t) = \mathbb{E}(\zeta^n(t)) = 0\ \forall t \in \tau$. This substitution in turn impacts equation (3).

When combining (8), (13) and relations (9), (4) we know that we can find some extended decision vector (also noted $x \in \mathbb{R}^n$) containing (x, y, z) and some matrix A, vector b such that the system $Ax \leq b$ models all the deterministic constraints (including bounds on x) found in Sect. 2. One can moreover find a matrix A^r and vectors a^r, b^r such that equation (3) wherein we have substituted the expectation of ζ for ζ is reflected by $a^r + A^r x \leq 0 \leq A^r x + b^r$.

Combined, this gives the following linear program:

$$\begin{aligned} \min_{x \in \mathbb{R}^n, x \geq 0} \quad & c^\mathsf{T} x \\ \text{s.t.} \quad & Ax \leq b \\ & -A^r x \leq b^r \\ & A^r x \leq -a^r. \end{aligned} \tag{14}$$

This model can be identified with the model considered in a classical deterministic unit-commitment setting.

3.2 A joint chance constraint model (JCCP)

In contrast to the expectation model wherein the effect of uncertainty is neglected, incorporating uncertainty fully in equation (3) leads to a joint chance constrained program. Indeed, by combining (13) with (3), we can see that the problem of Sect. 2 can be cast into the following form, where $\eta \in \mathbb{R}^m$ is a Gaussian random vector with variance-covariance matrix Σ and zero mean (we have explicitly extracted the non-zero average in (13)):

$$\begin{aligned} \min_{x \in \mathbb{R}^n, x \geq 0} \quad & c^\mathsf{T} x \\ \text{s.t.} \quad & Ax \leq b \\ & p \leq \mathbb{P}\big[a^r + A^r x \leq \eta \leq b^r + A^r x\big]. \end{aligned} \tag{15}$$

In fact the feasible set of (15) is convex due to the Gaussian character of $\eta \in \mathbb{R}^m$ and a theorem by Prékopa (1995). This makes the previous optimization problem a convex one. For convenience we define $\varphi : \mathbb{R}^n \to [0, 1]$ as $\varphi(x) = \mathbb{P}[a^r + A^r x \leq \eta \leq b^r + A^r x]$.

3.2.1 Link with the expectation problem

The chance constrained model can be seen as an extension of the expectation model since it takes into account the available stochastic information on the distribution of randomness, whereas model (14) only uses a single parameter. The following Lemma shows that any feasible solution of (15) is feasible for (14). Physically this can be explained by the fact that a “robust” control has to work well in the average situation.

Lemma 1 *Assume that $p > 0.5$ and that $\eta \in \mathbb{R}^m$ is a symmetric random variable, i.e., $\mathbb{P}[\eta \in A] = \mathbb{P}[\eta \in -A]$ for any measurable set $A \subseteq \mathbb{R}^m$. The feasible set of* (15) *is contained in the feasible set of* (14). *As a consequence the optimal value of* (14) *is lower than that of* (15).

Proof Assume that $x \in \mathbb{R}^n$ is not feasible for (14), for instance not $a^r + A^r x \leq 0$, i.e., there is at least one strictly positive component. By rearranging we may assume that this is the first one. Now

$$\begin{aligned}\mathbb{P}\big[a^r + A^r x \leq \eta \leq b^r + A^r x\big] &\leq \mathbb{P}\big[a^r + A^r x \leq \eta\big] \\ &\leq \mathbb{P}\big[e_1^{\mathsf{T}}\big(a^r + A^r x\big) \leq e_1^{\mathsf{T}} \eta\big] \leq \mathbb{P}[0 < \xi] < 0.5,\end{aligned}$$

where ξ is a centered one dimensional Gaussian random variable, and e_1 is a standard unit-vector of $\mathbb{R}^m$. This shows that x can't be feasible for (15). □

As mentioned the expectation model is a simple linear program. It is therefore much easier to solve than problem (15). Despite this fact and the fact that it yields solutions with low optimal values, it will be shown later in this paper that the solutions are useless since they violate constraints almost surely.

3.2.2 An algorithm for solving JCCP

In order to solve problem (15) we will use the supporting hyperplane method. This method was originally introduced by Veinott (1967) and adapted to the context of joint chance constrained programming by Prékopa and Szántai (1978a), Szántai (1988). This algorithm converges in a finite number of steps as shown in Prékopa and Szántai (1978a). We repeat the algorithm for completeness.

1. (Initialization) Let x_0 be the solution of (14), x_s a Slater point for (15). Set $A_0 = A$, $b_0 = b$ and $k = 0$ and pick some tolerance tol, e.g., tol $= 10^{-2}$. Let $\varepsilon > 0$ be a tolerance on the evaluation of φ.
2. (Interpolation) Find λ^* such that $x_k^* = (1 - \lambda^*)x_k + \lambda^* x_s$ and $p - \varepsilon \leq \varphi(x_k^*) \leq p$.
3. (Add Cut) Add constraint $-\nabla\varphi(x_k^*)^{\mathsf{T}} x \leq -\nabla\varphi(x_k^*)^{\mathsf{T}} x_k^*$ to the matrix system $A_k x \leq b_k$.
4. (Solve LP) Solve

$$\begin{aligned}&\min_{x \in \mathbb{R}^n, x \geq 0} \quad c^{\mathsf{T}} x \\ &\text{s.t.} \quad A_k x \leq b_k\end{aligned}$$

   to find x_{k+1}.
5. (Stopping Test) If $\frac{c^{\mathsf{T}}(x_k^* - x_{k+1})}{c^{\mathsf{T}} x_{k+1}} <$ tol then stop, x_k^* is sufficiently optimal, else set $k = k + 1$ and go to step 2.

For the previous algorithm to function we require a Slater point, i.e., some x_s such that $Ax_s \leq b$, and $\varphi(x_s) > p$. It can be obtained by solving the "max-p" problem (see Sect. 3.3). Moreover, we should be able to efficiently evaluate φ and $\nabla\varphi$. As shown in Corollary 1 below and Theorem 1 of van Ackooij et al. (2010), evaluating the gradient can be analytically reduced to computing function values in smaller dimension. Finally computing function values such as $\varphi(x)$ can be done by using the code of Genz (1992). Evaluating φ and $\nabla\varphi$ requires $2n+1$ calls to Genz' code.

3.3 Max-P problem

We define the "max-p" problem as the following optimization problem:

$$\begin{aligned} &\max_{x\in\mathbb{R}^n, x\geq 0} \varphi(x) := \mathbb{P}\big[a^r + A^r x \leq \eta \leq b^r + A^r x\big] \\ &Ax \leq b. \end{aligned} \tag{16}$$

Clearly any solution x_s of the previous problem with objective function value strictly bigger than p is a Slater point for problem (15). This "max-p" problem is not only an auxiliary problem for obtaining Slater points, but can also be interpreted as the problem of a decision-maker looking for maximum robustness, regardless of the costs. As a matter of fact if the optimal solution of (16) is strictly below one, then almost surely satisfying the "random" physical constraints (3) is not possible. The "max-p" problem therefore also provides us with information on the maximum robustness level p that is "possible".

3.4 Individual chance constraint model (ICCP)

We consider a simplification of the joint chance constrained model (15) by transforming each stochastic inequality into individual chance constraints of type $\mathbb{P}[d_1 + \langle a_1, x\rangle \leq \chi] \geq p$ and $\mathbb{P}[\chi \leq \langle a_2, x\rangle + d_2] \geq p$ for well chosen vectors $a_1, a_2 \in \mathbb{R}^n$, scalars d_1, d_2 and a standard Gaussian random variable $\chi \in \mathbb{R}$. An exact formulation is:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n, x\geq 0} c^\mathsf{T} x \\ &\text{s.t.} \quad Ax \leq b \\ &\qquad \mathbb{P}\big[e_i^\mathsf{T}\big(a^r + A^r x\big) \leq \eta_i\big] \geq p \quad \forall i = 1, \dots, m \\ &\qquad \mathbb{P}\big[\eta_i \leq e_i^\mathsf{T}\big(b^r + A^r x\big)\big] \geq p \quad \forall i = 1, \dots, m, \end{aligned} \tag{17}$$

where $e_i \in \mathbb{R}^m$ is the i-th standard unit vector.

As a matter of fact, model (17) can be reduced to a simple linear program since the inverse of $F_{\eta^i}(z) = \mathbb{P}[\eta_i \leq z]$ can be evaluated easily. It also offers improved robustness with respect to the expectation model (14) that offered none. However it can't guarantee a probability level of p for the whole stochastic inequality system and therefore offers far less robustness than the joint model (15) (van Ackooij et al. 2010). This will become apparent in the numerical experience.

3.5 A robust model

We would like to identify an uncertainty set $\mathcal{E}_p \subseteq \mathbb{R}^m$ for our random inflow process $\eta \in \mathbb{R}^m$ in such a way that the probability of η falling in this set is approximately p. We will then enforce the constraints of problem (15) to hold for all η in this set rather than in probability. We will use a specific ellipsoidal form for the uncertainty set and show that the thus obtained robust optimization problem then boils down to a linear program, once two conic quadratic problems have been solved.

In order to determine $\mathcal{E}_p$, let $LL^\mathsf{T} = \Sigma$ be the Cholesky decomposition of Σ. Let $y \in \mathbb{R}^m$ be defined as $y = L^{-1}\eta$ and assume that we dispose of a statistical estimate of $\mathbb{E}(y_i^4)$ (in the Gaussian case these are known exactly) for $i = 1, \dots, m$. Whenever the law of η is unknown, we can use the variance covariance matrix Σ obtained from statistic estimates. By construction, y is uncorrelated, we will make the (wrong) approximation that this is the same as independence. Now by the Lindeberg-Feller Central Limit Theorem (Prokhorov and Statulevičius 2000) we obtain that $y^\mathsf{T} y$ is approximately normally distributed with mean m and standard deviation σ_C, i.e., $y^\mathsf{T} y \approx \mathcal{N}(m, \sigma_C)$, with $\sigma_C = \sqrt{\sum_{i=1}^m \mathbb{E}(y_i^4) - m}$.

We now define $\mathcal{E}_p = \{z \in \mathbb{R}^m : z^\mathsf{T} \Sigma^{-1} z \le m + \Phi^{-1}(p)\sigma_C\}$. It follows in the case that η follows a multivariate Gaussian law that $\mathbb{P}[\eta \in \mathcal{E}_p] = p$. This will be true approximately when η follows another multivariate law.

We therefore consider the following robust version of problem (15):

$$\begin{aligned} \min_{x \in \mathbb{R}^n, x \ge 0} \quad & c^\mathsf{T} x \\ \text{s.t.} \quad & Ax \le b \\ & a^r + A^r x \le \inf \mathcal{E}_p \\ & b^r + A^r x \ge \sup \mathcal{E}_p, \end{aligned} \tag{18}$$

where $\inf \mathcal{E}_p \in \mathbb{R}^m$ denotes the vector whose components are the coordinate-wise minima of $\mathcal{E}_p$ ($\sup \mathcal{E}_p \in \mathbb{R}^m$ is defined similarly). Both $\inf \mathcal{E}_p$ and $\sup \mathcal{E}_p$ are solutions of a conic quadratic optimization problem. Indeed model (18) is equivalent with

$$\begin{aligned} \min_{x \in \mathbb{R}^n, x \ge 0} \quad & c^\mathsf{T} x \\ \text{s.t.} \quad & Ax \le b \\ & a^r + A^r x \le \xi \le b^r + A^r x \quad \forall \xi \in \mathcal{E}_p. \end{aligned}$$

Since model (18) basically looks at the smallest rectangle containing $\mathcal{E}_p$ and requires satisfaction of constraints for all elements in the rectangle, one could also look at alternative ways to obtain such a rectangle. Basically, we are looking for some $\overline{\eta}$ and $\underline{\eta}$ such that $\mathbb{P}[\underline{\eta} \le \eta \le \overline{\eta}] \approx p$. These would then give better bounds than $\inf \mathcal{E}_p$ and $\sup \mathcal{E}_p$ as above, since in general $\mathbb{P}[\inf \mathcal{E}_p \le \eta \le \sup \mathcal{E}_p] > p$. In the Gaussian case considered here we can exactly evaluate the probabilistic contents of such rectangles and hence fine-tune the rectangle. Clearly any feasible solution of problem (15) will also provide such vectors. This last way of obtaining those vectors offers no computational advantage to (18) other than prematurely ending the algorithm that solves (15). An alternative would be to take some $q < p$, such that $\mathbb{P}[\inf \mathcal{E}_q \leq \eta \leq \sup \mathcal{E}_q] \approx p$. This is computationally not intensive, but requires evaluations of probabilistic contents. In order to investigate the impact of the choice of this rectangle we have made some runs with model (18) wherein the rectangle was made to fit perfectly. In practice, we have obtained $\underline{\eta}$ and $\overline{\eta}$ by taking some ad-hoc convex combinations between the Slater point and the solution of (14). These results will be referred to as Robust-Calibrated (Robust-Calib) or (18)-Calib.

4 Numerical example

In this section we consider a numerical example from the industry. The instance size is moderate but realistic. The nominal inflows, i.e., s_t^n in (1), are considered constant through time. Finally, the water values are not assumed to depend on the volume, and thereby correspond to the V_0 level. It was shown in Sect. 2.1.5 that adding the volume dependency induces no substantial difficulties. The focus of this numerical example is the impact of uncertainty. We will consider 24 time steps of 2 hours each. Figures 1 and 2 show further data of our example. This implies the following dimensions for our problem: the Gaussian vector dimension m is 48, the decision vector has 700 elements and the polyhedral constraints are defined by about 1000 linear inequalities.

As uncertainty is concerned we will assume that reservoirs 1 (“Vouglans”) and 2 (“Saut Mortier”) have random inflows. The standard deviations of the innovations ζ of the inflow process A^n in (1) are taken to be equal to 20 % of the nominal inflow values (0.3 m^3/s for reservoir 2). We will consider two instances, one wherein inflows on both reservoirs follow an AR(1) process with coefficient 0.9. A second instance is one wherein we assume that inflows on reservoir 2 follow an AR(3) process with

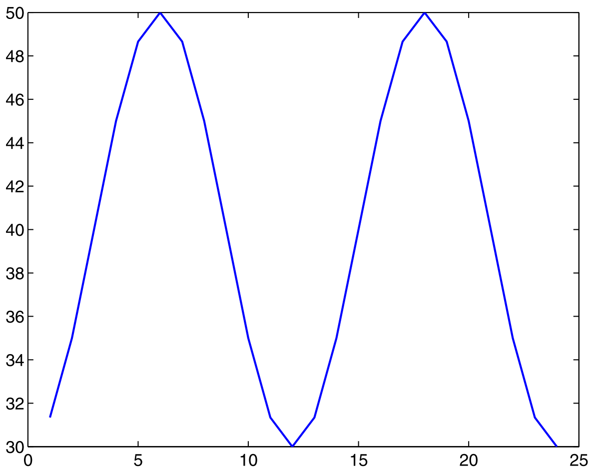

Fig. 1 The price signal

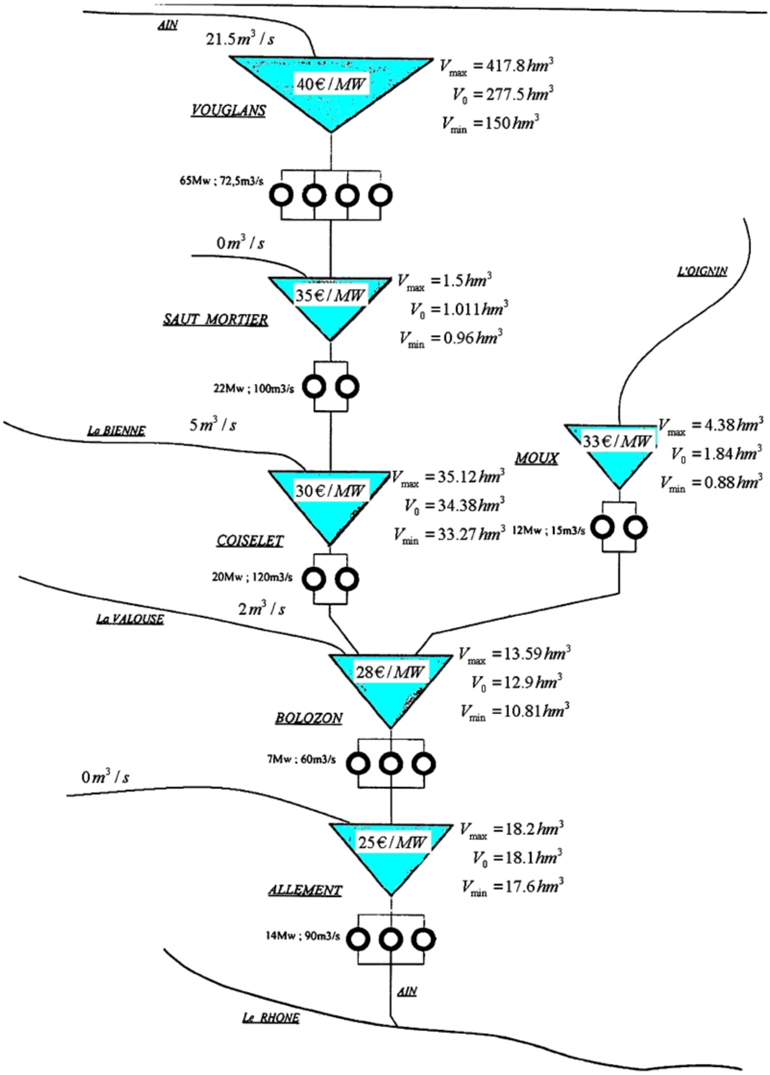

Fig. 2 The hydro valley

coefficients $(0.9, 0.7, -0.7)$. In this instance inflows on reservoir 1 still follow an AR(1) process. The required probability level p in (15) is taken to be 0.8.

Solving the problems introduced in Sect. 3, we obtain the results as given in Table 1, Figs. 3 and 4. We have set a tolerance of 10^{-2} for the supporting hyperplane algorithm for joint chance constrained programming. It should be stated that the true optimal solution of problem (15) for instance 2 gives a cost, only 0.6 % away from the deterministic cost. Indeed the price of chance-constrained robustness is cheap here.

Table 1 shows optimal costs and number of violations. In order to compute the latter information, we have made an a posteriori check of empirical probabilities by

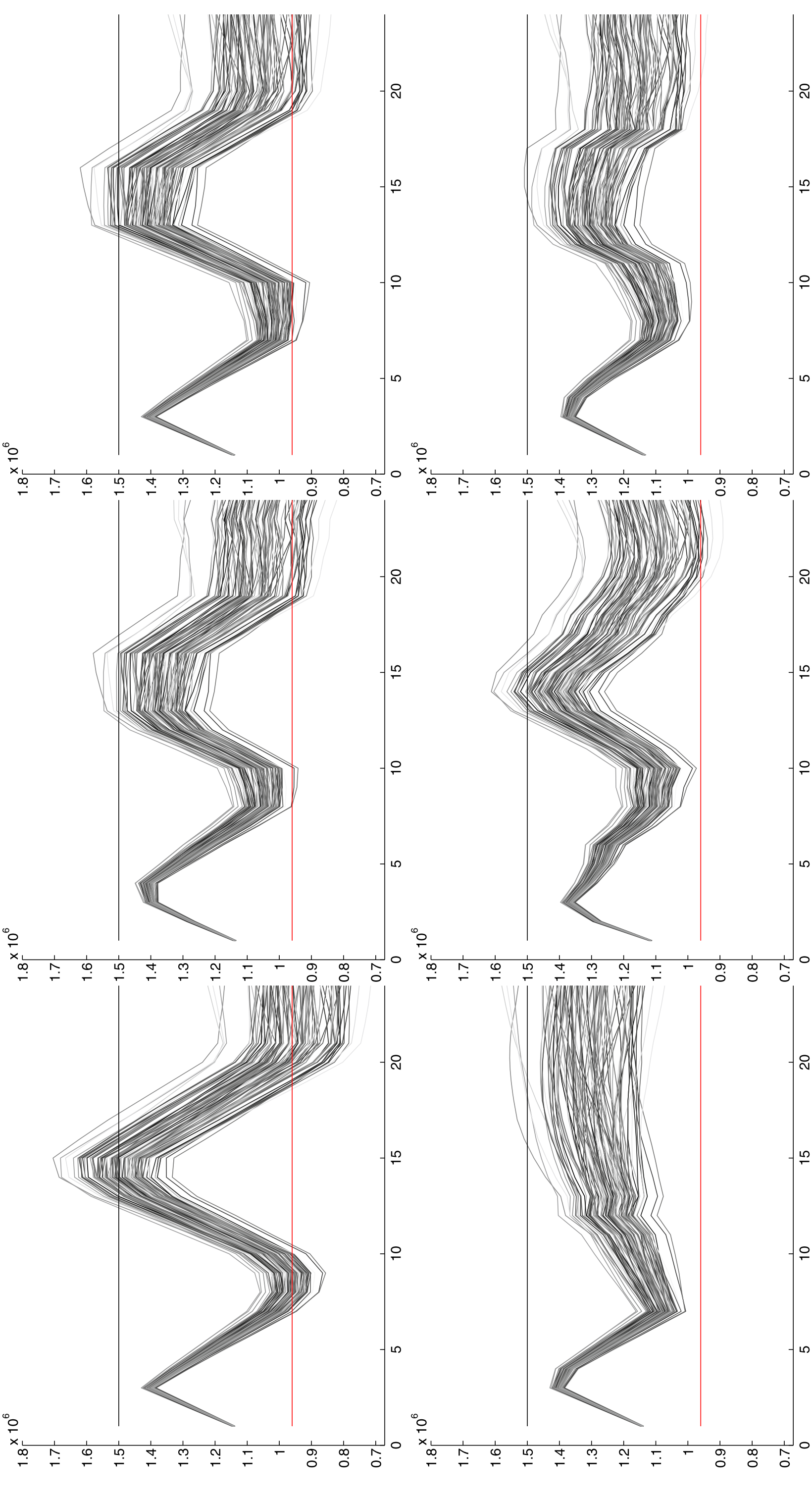

Fig. 3 Trajectories of filling levels in reservoir "Saut Mortier" and instance 2 for 100 simulated inflow scenarios. From *top left* to *bottom right*, solutions of problems (14), (15), (17), (18), (18)-Calib and (16)

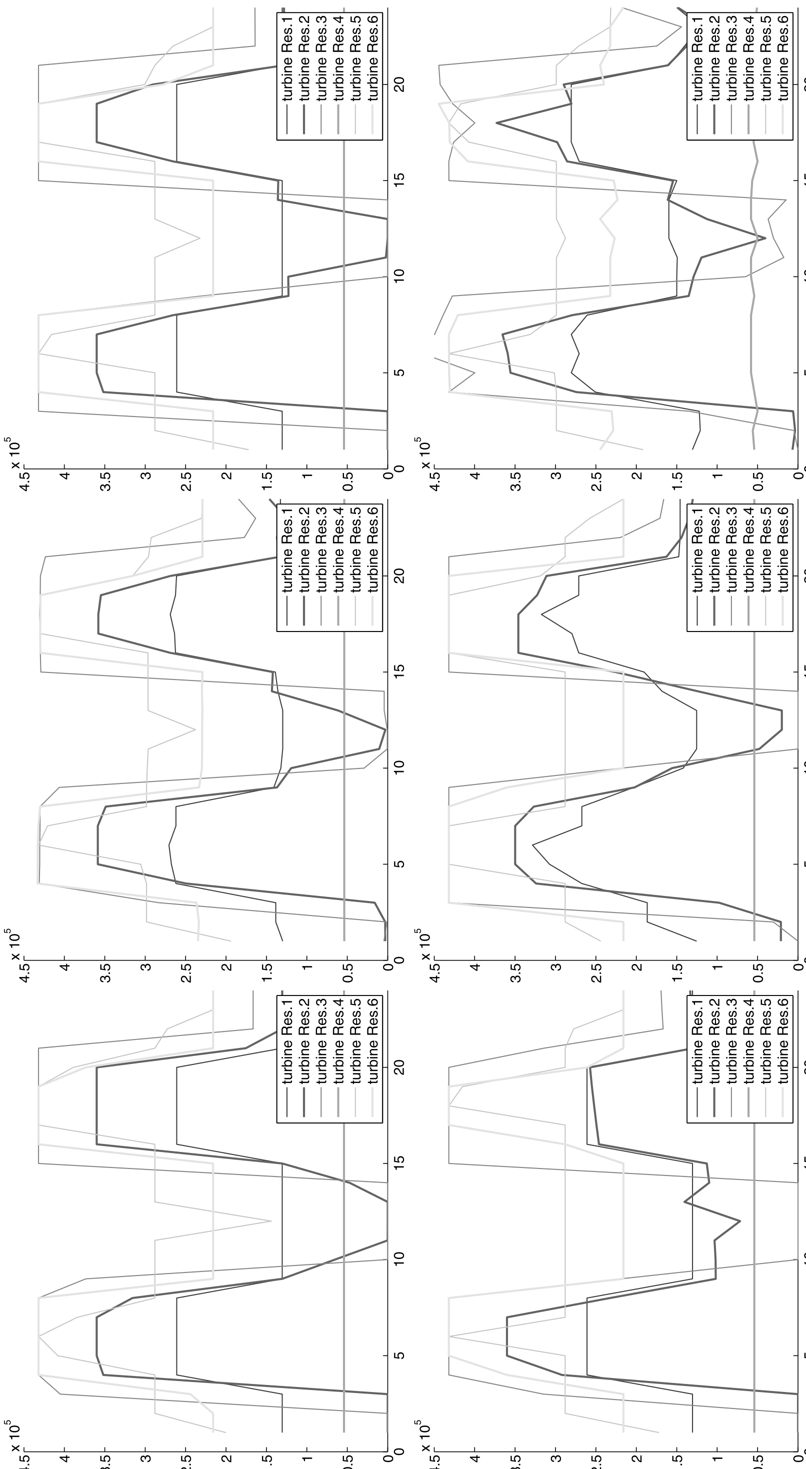

Fig. 4 Turbined volumes (m^3) for instance 2. From *top left* to *bottom right*, solutions of problems (14), (15), (17), (18), (18)-Calib. and (16)

Table 1 Comparison of costs and number of violations

Inst.	Item/Problem	Det (14)	JCCP (15)	ICCP (17)	Robust ($\mathcal{E}_p$) (18)-1	Robust (Calib) (18)-Calib	MaxP (16)
1	nbViolation	100	20	29	0	1	0
1	Cost (€)	$-1.0478e^5$	$-1.0395e^5$	$-1.0443e^5$	$-1.0355e^5$	$-1.0099e^5$	$-9.9176e^4$
2	nbViolation	100	20	35	4	21	2
2	Cost (€)	$-1.0478e^5$	$-1.0340e^5$	$-1.0422e^5$	$-1.0282e^5$	$-1.0251e^5$	$-9.9176e^4$

generating 100 scenarios and counting the number of violations. The volume trajectories resulting from these scenarios are shown in Fig. 3. Clearly we observe the advantage of using joint chance constrained programming. The additional cost with respect to the deterministic solution is only small, but robustness can be fine tuned. A full robust solution turns out quite costly. Finally individual chance constrained programming can not be used to mimic joint chance constraints as we have no control over the number of violations over a period of time.

When comparing the turbined volumes in Fig. 4, one can observe that they are quite similar for most solutions (except for max-p which does not see the cost vector and is hence only incited to turbine if this allows us to improve robustness) and most reservoirs, except for "Saut Mortier". This reservoir has tight volume bounds and is most heavily impacted by the stochastic inflows. The solution (15) turbines a bit less in the beginning to avoid violations in time steps 8–10, a bit more during time steps 12–15 to avoid violations there and stops earlier to avoid violations for the last time steps. Solution (17) offers an intermediate solution. The solution (18) heavily increases turbining during steps 10–15 and drastically reduces during steps 15–20 for additional robustness. Indeed, even though the uncertainty $\mathcal{E}_p$ is very well calibrated, the solution is over-robust. Unfortunately for larger values of p (in fact $p > 0.85$) this will lead to an empty feasible set of problem (18), whereas solutions of (15) can be found. It also shows the difficulty of getting the robust rectangle well calibrated for problem (18)-Calib. Indeed, even though the rectangle is calibrated to give exactly the same probabilistic contents in both instances, one gives over-robust results (3.6 % away from deterministic solution), whereas the other gives more reasonable results as the number of violations is concerned, but still at a large cost (2.2 % away from deterministic solution).

5 Conclusions

In this paper, we have set up a joint chance constrained programming approach for dealing with uncertainty on inflows in hydro valley optimization. We have derived a detailed model for a real hydro valley, but one wherein combinatorial constraints are neglected. In order to have a more realistic description of inflows, we have considered a causal time series setting with Gaussian innovations. The latter choice allows us to preserve convexity of the optimization problem and have a more realistic model on inflows. In order to solve this JCCP problem we have used a supporting hyperplane method that requires a Slater point and gradients. The probability functions and gradients can be efficiently computed using Genz' code.

In order to highlight the interest of joint chance constraint programming, we have also investigated alternative models. Indeed, we have considered a model based on individual chance constraints and a robust model. The obtained results have been compared on a realistic hydro valley. Hence, despite the fact that the alternative models often require less hypothesis on the law of the inflows, they provide conservative and costly solutions. The simpler models, such as the individual chance constraint one, are shown to yield insufficient robustness. The robust model induces an important extra cost, despite the well calibrated "uncertainty set" and moreover often leads to empty feasible sets. Joint Chance Constrained programming appears as an approach offering a good trade-off between cost and robustness and can be tractable for complex realistic models. In addition, we have shown that in principle we can handle a real size valley.

Future perspectives consist in working on model realism and on the algorithm for solving the chance constraint problem. Indeed, from a modeling perspective, we could integrate the combinatorial constraints on the decision variables, potentially without many difficulties. From an algorithmic perspective, instead of using a supporting hyperplane idea, one could use a bundle method to hopefully improve computation times and stability. A second point that needs investigations is an improved use of Genz' code. We could combine the use of Genz' code with Prekopa's linear programming estimation method of probability measures in order to either increase the size of the model or improve the speed.

Acknowledgements The second author would like to thank the DFG Research Center MATHEON "Mathematics for Key Technologies" in Berlin for supporting his work.

Appendix A: Algorithmic perspectives/Second order methods

If one is interested in applying second order solution methods to increase the efficiency of the solution process, one has to work out second derivatives of the probability function φ on the basis of the gradients obtained in Theorem 1 of van Ackooij et al. (2010). This is done in the following lemma.

Lemma 2 *Let ξ be an n-dimensional Gaussian random vector with mean μ and variance-covariance matrix Σ. We define the mapping $F_\xi(a,b) = \mathbb{P}[a \le \xi \le b]$ for any rectangle, i.e., $a \le b$. Let D_n^i denote the n-th order identity matrix from which the ith row has been deleted. For each $y \in \mathbb{R}^n$, $1 \le i \le n$ and $z \in \mathbb{R}$ we define $y^{c^{i,n}(z,\Sigma_i)} = D_n^i(y + \Sigma_{i,i}^{-1}(z - y_i)\Sigma_i) \in \mathbb{R}^{n-1}$, where Σ_i is the ith column of Σ. We will occasionally abbreviate this with $y^{c_1^i(z)}$. We also define*

$$y^{c_2^{i,j}(z,w)} = \left(y^{c^{i,n}(z,\Sigma_i)}\right)^{c^{j,n-1}(w,\Sigma_j^{c^n(i)})},$$

where we have defined $\Sigma^{c^n(i)} = D_n^i(\Sigma - \Sigma_{i,i}^{-1}\Sigma_i\Sigma_i^{\mathsf{T}})(D_n^i)^{\mathsf{T}}$. *We define* $\xi^{c_1^i(z)}$ *as the Gaussian random variable with mean* $\mu^{c_1^i(z)}$ *and covariance matrix* $\Sigma^{c^n(i)}$. *In a similar way, we define* $\xi^{c_2^{i,j}(z,w)}$ *as the Gaussian random variable with mean* $\mu^{c_2^{i,j}(z,w)}$ *and covariance matrix* $\Sigma^{c_2^{i,j}} := D_{n-1}^j(\Sigma^{c^n(i)} - (\Sigma_{j,j}^{c^n(i)})^{-1}\Sigma_j^{c^n(i)}(\Sigma_j^{c^n(i)})^{\mathsf{T}})(D_{n-1}^j)^{\mathsf{T}}$, *where* $\Sigma_j^{c^n(i)}$ *denotes the* j*-th column of* $\Sigma^{c^n(i)}$. *The following holds, for* $j = \hat{j}$ *if* $\hat{j} < i$ *and* $j = \hat{j} - 1$ *if* $\hat{j} > i$:

$$
\begin{aligned}
&\frac{\partial^2}{\partial a_{\hat{j}}\partial a_i}F_\xi(a,b)\\
&\quad= f_{\mu_j^{c_1^i(a_i)},\Sigma_{j,j}^{c^n(i)}}(a_j)f_{\mu_i,\Sigma_{i,i}}(a_i)F_{\xi^{c_2^{i,j}(a_i,a_j)}}\big(D_{n-1}^jD_n^ia,\, D_{n-1}^jD_n^ib\big) \quad \forall\hat{j}\neq i\\
&\frac{\partial^2}{\partial b_{\hat{j}}\partial a_i}F_\xi(a,b)\\
&\quad= \begin{cases} -f_{\mu_j^{c_1^i(a_i)},\Sigma_{j,j}^{c^n(i)}}(b_j)f_{\mu_i,\Sigma_{i,i}}(a_i)F_{\xi^{c_2^{i,j}(a_i,b_j)}}(D_{n-1}^jD_n^ia,\, D_{n-1}^jD_n^ib) & \forall\hat{j}\neq i\\ 0 & \hat{j}=i\end{cases}\\
&\frac{\partial^2}{\partial b_{\hat{j}}\partial b_i}F_\xi(a,b)\\
&\quad= f_{\mu_j^{c_1^i(b_i)},\Sigma_{j,j}^{c^n(i)}}(b_j)f_{\mu_i,\Sigma_{i,i}}(b_i)F_{\xi^{c_2^{i,j}(b_i,b_j)}}\big(D_{n-1}^jD_n^ia,\, D_{n-1}^jD_n^ib\big) \quad \forall\hat{j}\neq i,
\end{aligned}
$$

where $f_{\mu,\sigma}(x)$ *is the Gaussian density with mean* μ *and variance* σ. *Moreover, whenever* $j = i$ *and* z *is* a *or* b *we have*:

$$
\begin{aligned}
&\frac{\partial}{\partial z_i}(b_i - a_i)\frac{\partial^2}{\partial z_i^2}F_\xi(a,b)\\
&\quad= -\frac{z_i - \mu_i}{\Sigma_{i,i}}f_{\mu_i,\Sigma_{i,i}}(z_i)F_{\xi^{c_1^i(z_i)}}\big(D_n^ia,\, D_n^ib\big)\\
&\qquad - f_{\mu_i,\Sigma_{i,i}}(z_i)\big(D_n^i\Sigma_{i,i}^{-1}\Sigma_i\big)^{\mathsf{T}}\big(\nabla_{\tilde{a}}F_{\tilde{\xi}^{c_1^i(z_i)}}(\tilde{a},\tilde{b}) + \nabla_{\tilde{b}}F_{\tilde{\xi}^{c_1^i(z_i)}}(\tilde{a},\tilde{b})\big),
\end{aligned}
$$

where $\tilde{a} = D_n^ia - \mu^{c_1^i(z_i)}$, $\tilde{\xi}^{c_1^i(z_i)} = \xi^{c_1^i(z_i)} - \mu^{c_1^i(z_i)}$ *and* $\tilde{b}$ *is defined similarly.*

Proof The formula for the cross derivatives follow from a straight-forward second application of Theorem 1 in van Ackooij et al. (2010). The diagonal terms are more subtle to derive and require the following reformulation:

$$
\begin{aligned}
F_{\xi^{c_1^i(z_i)}}\big(D_n^ia,\, D_n^ib\big) &= \mathbb{P}\big(D_n^ia \le \xi^{c_1^i(z_i)} \le D_n^ib\big)\\
&= \mathbb{P}\big(D_n^ia - \mu^{c_1^i(z_i)} \le \xi^{c_1^i(z_i)} - \mu^{c_1^i(z_i)} \le D_n^ib - \mu^{c_1^i(z_i)}\big)\\
&= F_{\tilde{\xi}^{c_1^i(z_i)}}(\tilde{a},\tilde{b}).
\end{aligned}
$$

In particular one obtains for $\tilde{a}(z_i) = \tilde{a}$

$$\begin{aligned}\tilde{a} &= D_n^i a - \mu^{c_1^i(z_i)} = D_n^i a - D_n^i\big(\mu + \Sigma_{i,i}^{-1}(z_i - \mu_i)\Sigma_i\big)\\ &= D_n^i\big(a - \mu + \Sigma_{i,i}^{-1}\mu_i\Sigma_i\big) - D_n^i\Sigma_{i,i}^{-1}z_i\Sigma_i,\end{aligned}$$

which together with the following identity

$$\frac{\partial}{\partial z_i}F_{\xi^{c_1^i(z_i)}}\big(\tilde{a}(z_i), \tilde{b}(z_i)\big) = \nabla_{\tilde{a}}F_{\xi^{c_1^i(z_i)}}(\tilde{a}, \tilde{b})D_{z_i}\tilde{a}(z_i) + \nabla_{\tilde{b}}F_{\xi^{c_1^i(z_i)}}(\tilde{a}, \tilde{b}))D_{z_i}\tilde{b}(z_i),$$

an application of the chain-rule and the already established formula for 1st derivatives gives the proposition. □

The following corollary deals with gradients and Hessians of our probability function $\varphi : x \in \mathbb{R}^n \mapsto \mathbb{P}[a + Ax \leq \xi \leq Bx + b] \in [0, 1]$. These follow easily from Lemma 2 and Theorem 1 of van Ackooij et al. (2010) upon noting that $\varphi(x) = F_\xi(Ax, Bx)$ with F_ξ as introduced in Lemma 2.

Corollary 1 *Let ξ be a Gaussian Random variable of dimension n. Let x, A,B,a,b be vectors and matrices of appropriate dimension. Now consider the mapping $\varphi : x \in \mathbb{R}^n \mapsto \mathbb{P}[a + Ax \leq \xi \leq Bx + b] \in [0, 1]$. We have*:

$$\begin{aligned}\nabla\varphi &= \nabla_a F_\xi(a, b)^\mathsf{T} A + \nabla_b F_\xi(a, b)^\mathsf{T} B\\ \nabla^2\varphi &= A^\mathsf{T}\nabla^2_{aa}F_\xi(a, b)A + A^\mathsf{T}\nabla^2_{ab}F_\xi(a, b)B + B^\mathsf{T}\nabla^2_{ba}F_\xi(a, b)A + B^\mathsf{T}\nabla^2_{bb}F_\xi(a, b)B.\end{aligned}$$

References

Andrieu L, Henrion R, Römisch W (2010) A model for dynamic chance constraints in hydro power management. Eur J Oper Res 207:579–589

Belloni A, Diniz AL, Maceira ME, Sagastizábal C (2003) Bundle relaxation and primal recovery in unit-commitment problems. The Brazilian case. Ann Oper Res 120(1–4):21–44

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) Robust optimization. Princeton University Press, Princeton

Charnes A, Cooper W (1960) Chance-constrained programming. Manag Sci 6:73–79

Cohen G, Zhu DL (1983) Decomposition-coordination methods in large-scale optimization problems. The non-differentiable case and the use of augmented Lagrangians. In: Large scale systems. Theory and applications, vol 1

d'Ambrosio C, Lodi A, Martello S (2010) Piecewise linear approximation of functions of two variables in Milp models. Oper Res Lett 38:39–46

de Farias DP, Van Roy B (2002) Approximate dynamic programming via linear programming. In: Advances in neural information processing systems, vol 14

Diniz AL, Maceira MEP (2008) A four-dimensional model of hydro generation for the short-term hydrothermal dispatch problem considering head and spillage effects. IEEE Trans Power Syst 23(3):1298–1308

Dubost L, Gonzalez R, Lemaréchal C (2005) A primal-proximal heuristic applied to French unitcommitment problem. Math Program 104(1):129–151

Duranyildiz I, Önöz B, Bayazit M (1999) A chance-constrained LP model for short term reservoir operation optimization. Turk J Eng 23:181–186

Edirisinghe NCP, Patterson EI, Saadouli N (2000) Capacity planning model for a multipurpose water reservoir with target-priority operation. Ann Oper Res 100:273–303

Finardi EC, Da Silva EL (2006) Solving the hydro unit commitment problem via dual decomposition and sequential quadratic programming. IEEE Trans Power Syst 21(2):835–844

Genz A (1992) Numerical computation of multivariate normal probabilities. J Comput Graph Stat 1:141–149

Girardeau P (2010) Résolution de grands problèmes en optimisation stochastique dynamique et synthèse de lois de commande. PhD thesis, Université Paris Est

Gouda A, Szántai T (2010) On numerical calculation of probabilities according to Dirichlet distribution. Ann Oper Res 177:185–200

Henrion R, Möller A (2012) A gradient formula for linear chance constraints under Gaussian distribution. Math Oper Res 37:475–488

Lemaréchal C, Sagastizábal C (1994) An approach to variable metric bundle methods. In: Lecture notes in control and information science, vol 197. Springer, Berlin, pp 144–162

Loiaciga HA (1988) On the use of chance constraints in reservoir design and operation modeling. Water Resour Res 24:1969–1975

Loucks DP, Stedinger JR, Haith DA (1981) Water resource systems planning and analysis. Prentice-Halls, New York

Morgan DR, Eheart JW, Valocchi AJ (1993) Aquifer remediation design under uncertainty using a new chance constraint programming technique. Water Resour Res 29:551–561

Nilsson O, Sjelvgren D (1997) Variable splitting applied to modelling of start-up costs in short term hydro generation scheduling. IEEE Trans Power Syst 12(2):770–775

Pereira MVF, Pinto LMVG (1991) Multi-stage stochastic optimization applied to energy planning. Math Program 52(2):359–375

Philpott AB, Dallagi A, Gallet E (2011) On cutting plane algorithms and dynamic programming for hydroelectricity generation. Available online at http://www.epoc.org.nz/papers/MORGANEvsDOASA.pdf, pp 1–20

Philpott AB, Guan Z (2008) On the convergence of stochastic dual dynamic programming and related methods. Oper Res Lett 36(4):450–455

Ponrajah RA, Witherspoon J, Galiana FD (1998) Systems to optimise conversion efficiencies at Ontario hydro's hydroelectric plants. IEEE Trans Power Syst 13(3):1044–1050

Prékopa A (1995) Stochastic programming. Kluwer, Dordrecht

Prékopa A, Szántai T (1978a) Flood control reservoir system design using stochastic programming. Math Program Stud 9:138–151

Prékopa A, Szántai T (1978b) A new multivariate gamma distribution and its fitting to empirical streamflow data. Water Resour Res 14:19–24

Prokhorov YV, Statulevičius V (eds) (2000) Limit theorems of probability theory. Springer, Berlin

Shumway RH, Stoffer DS (2000) Time series analysis and its applications, 1st edn. Springer, Berlin

Szántai T (1985) Numerical evaluation of probabilities concerning multi-dimensional probability distributions. PhD thesis, Hungarian Academy of Sciences

Szántai T (1988) A computer code for solution of probabilistic-constrained stochastic programming problems. In: Ermoliev Y, Wets RJ-B (eds) Numerical techniques for stochastic optimization, pp 229–235

Torrion P, Leveugle J (1985) Comparaison de différentes méthodes d'optimisation appliquées à la gestion annuelle du système offre-demande d'électricité français. M.85.108, p 20

Turgeon A (1980) Optimal operation of multi-reservoir power systems with stochastic inflows. Water Resour Res 16(2):275–283

van Ackooij W, Henrion R, Möller A, Zorgati R (2010) On probabilistic constraints induced by rectangular sets and multivariate normal distributions. Math Methods Oper Res 71(3):535–549

Veinott AF (1967) The supporting hyperplane method for unimodal programming. Oper Res 15:147–152

Zorgati R, van Ackooij W (2011) Optimizing financial and physical assets with chance-constrained programming in the electrical industry. Optim Eng 12(1):237–255

Zorgati R, van Ackooij W, Apparigliato R (2009) Supply shortage hedging: estimating the electrical power margin for optimizing financial and physical assets with chance-constrained programming. IEEE Trans Power Syst 24(2):533–540