A robust optimization model for the risk averse reservoir management problem

Charles Gauvin[1,2,3] , Erick Delage[2,4], Michel Gendreau[1,3]

[1]École polytechnique de Montréal, C.P. 6079, succursale Centre-ville, Montréal, (Canada), H3C 3A7
[2]GERAD, 3000 chemin de la Côte-Sainte-Catherine, Montréal, (Canada), H3T 2A7
[3]CIRRELT, C.P. 6128, succursale Centre-ville, Montréal, (Canada), H3C 3J7
[4]HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, (Canada), H3T 2A7

Abstract

This paper presents a new formulation for the risk averse stochastic reservoir management problem. Using recent advances in robust optimization and stochastic programming, we propose a dynamic, multi-objective model based on minimization of a multidimensional risk measure associated with floods and droughts for a hydro-electrical complex. We present our model and then identify approximate solutions using standard affine decision rules commonly found in the literature as well as lifted decision rules. Finally, we conduct thorough numerical experiments based on a real river system in Western Québec and conclude on the relative performance of families of decision rules.

Keywords: Stochastic programming, Multiple objective programming, OR in energy, Risk analysis, Robust optimization

1 Introduction

The problem of designing an optimal release schedule for a set of interconnected reservoirs is extremely challenging. Operators must often make decisions for various dams or other water control structures in each period of a given time horizon while taking into account the dynamical structure and topology of the system. It may also be important to consider complex nonlinear physical phenomenon such as the effect of water volumes on outflow or water delays. This generally leads to high dimensional dynamic and nonconvex problems.

Furthermore, operators must deal with various conflicting objectives of varying importance. Hydro-electrical complexes must namely balance criteria

such as recreational and environmental needs with electricity generation, irrigation and flood control (see (1)). The importance of multi-objective optimization for stochastic reservoir management problems is namely illustrated in (2; 3; 4).

There is also considerable uncertainty surrounding various factors such as price of electricity, turbine breakage, inflows and demand. It is virtually impossible to perfectly identify and represent the true multi-dimensional stochastic process. However, failure to take into account this stochasticity can lead not only to severely suboptimal solutions, but even disastrous consequences such as floods or droughts. Incorporating uncertainty also poses serious numerical limitations as this leads to multi-stage stochastic programs which are generally intractable (see (5)).

This paper addresses these issues by proposing a dynamic multi-reservoir model based on robust optimization. Robust optimization is a rapidly expanding paradigm which has gained in popularity over the years (6; 7). By restricting uncertainty to reside within a given parametrized and deterministic convex set, large stochastic models can be converted into their robust counterparts while maintaining tractability through conic programming duality.

Unlike methods traditionally used in stochastic reservoir management problems such as stochastic dynamic programming (SDP) (8; 9; 10; 11), stochastic dual dynamic programming algorithm (SDDP) (12; 3) and stochastic programs based on scenario trees (13; 14), robust optimization eliminates the need for distributional assumptions which increases the resilience of the solutions.

The paradigm has been extended to dynamic problems with affine decision rules (15) and later extended by various authors to provide more flexibility and precision (16; 17; 18). Although these decision rules provide a conservative approximation of the objective value, they remain tractable and provide rules that can easily be used in simulations, which is not necessarily the case of SDP, SDDP or tree-based stochastic programs.

Robust optimization models have recently been used for reservoir management problems. The authors of (19) and (20) namely use this framework

to derive operating policies to maximize the expected electric production for a multi-period and multi-reservoir hydro-electric complex. However, both approaches only consider stylized and simplified operating constraints.

This paper differs significantly from the previous work in that it considers realistic operating rules and focuses on a difficult problem instance without complete recourse where these tight constraints may be significantly violated under certain inflow scenarios. For this reason, we formulate a novel multi-objective problem that minimizes the risk of floods at various critical locations on the river.

Contrary to an overwhelming majority of models mentioned previously, we do not assume risk neutrality. We explicitly consider risk aversion through parametric families of *systemic* risk measures (see (21)). This allows us to easily interpret the risk of these natural disasters and perform sensitivity analysis.

Like (20), we evaluate the use of linear and piecewise affine decision rules. Our work draws heavily from both state-of-the art techniques in robust optimization and the work of (18) on stochastic programming. We conduct extensive numerical experiments based on a real river system to evaluate and compare these different policies and also consider a rolling horizon approach embedded in a realistic simulation environment.

Finally, our approach is extremely flexible, distribution-free, requires very few assumptions and does not suffer from the limitations of various competing methods such as SDP and SDDP. Indeed, our method can consider inflows delays without numerical difficulties, does not require discretization of the stochastic process and does not require time decomposition.

The rest of the paper is structured as follows. We introduce our model and the multi-objective nature of the problem in section 2. Section 3 focuses on the uncertainty surrounding inflows while section 4 discusses theoretical foundations from risk measures. Section 5 discusses affine and piecewise-affine decision rules. We then present a case study based on a real hydroelectric reservoir management problem in Québec in section 6 and offer concluding remarks in section 7.

2 Description of the deterministic reservoir management problem (RMP)

2.1 Basic model

We consider a discrete time horizon of T periods with decisions at each time $t = 1, ..., T$. If we completely omit all sources of uncertainty, the reservoir management problem we study consists in finding a feasible solution to the following program:

$$\text{(Vol. Bounds)} \quad \underline{v}_j \leq \mathcal{V}_{jt} \leq \bar{v}_j \qquad \forall j \in J, t \tag{1}$$

$$\text{(Flow cons.)} \quad \mathcal{V}_{jt} = v_{j0} + \sum_{\tau \leq t} \left(\sum_{i^- \in I^-(j)} \sum_{l=\delta_{i^-}^{min}}^{\min\{\delta_{i^-}^{max}, \tau-1\}} \lambda_{i^-l} \mathcal{R}_{i^-\tau-l} - \sum_{i^+ \in I^+(j)} \mathcal{R}_{i^+\tau} + \xi_{j\tau} \right) \beta \quad \forall j, t \tag{2}$$

$$\text{(Flow Bounds)} \quad \underline{r}_i \leq \mathcal{R}_{it} \leq \bar{r}_i, \qquad \forall i \in I, t \tag{3}$$

$$\text{(Evac. Curve)} \quad \mathcal{S}_{it} \leq \mathcal{C}_i(\mathcal{V}_{j^-(i)t}), \qquad \forall i \in I^{evac}, t \tag{4}$$

$$\text{(Var. Bound)} \quad |\mathcal{R}_{it} - \mathcal{R}_{it-1}| \leq \bar{\Delta}_i \qquad \forall i, t \geq 2 \tag{5}$$

$$\text{(Spill. Bounds)} \quad \underline{s}_i \leq \mathcal{S}_{it} \leq \bar{s}_i \qquad \forall i, t \tag{6}$$

$$\text{(Turb. Bounds)} \quad \underline{q}_i \leq \mathcal{Q}_{it} \leq \bar{q}_i \qquad \forall i, t \tag{7}$$

$$\text{(Flow def.)} \quad \mathcal{R}_{it} = \mathcal{Q}_{it} + \mathcal{S}_{it} \qquad \forall i, t \tag{8}$$

In this formulation, the decision variables $\mathcal{V}_t, \mathcal{S}_t, \mathcal{Q}_t, \mathcal{R}_t$, respectively represent average volumes (hm^3) during time t, average spillage (non-productive water discharge) (m^3/s) over time t, average turbined outflow (productive water discharge) (m^3/s) over time t and average total flow (m^3/s) over time t. The sum of the spillage and the turbined outflow is equal to the total flow.

We distinguish between the spillage and turbined controls which represent real implementable decisions and the volume and total flow which are analysis variables [1]used to enhance model readability (refer to table (1)) .

[1] Implementability indicates that decisions at a given time can be realistically controlled by human intervention and depend only on the available information. On the other hand, analysis variables are completely determined uniquely once the implementable decisions are fixed. They only serve to track the evolution of the system.

Notation	Name	Implementable	Unit
$\mathcal{V}_t$	Volume	No	(hm^3)
$\mathcal{S}_t$	Spillage	Yes	(m^3/s)
$\mathcal{Q}_t$	Turbined outflow	Yes	(m^3/s)
$\mathcal{R}_t$	Total flow	No	(m^3/s)

Table 1: Decision variables

We let J represent the set of reservoirs; some of which are simply portions of the river with little or no capacity. The set I represents the water controlling sites. Not all sites have plants that may produce electricity by letting water flow though turbines. The sets $I^-(j), I^+(j) \subset I$ represent the set of upstream (incoming) and downstream (outgoing) plants with respect to reservoir j.

For each reservoir j, lower volume bounds ($\underline{v}_j$) represents a minimum amount of water required namely for health, environmental and recreational purposes. These bounds are particularly important during the hot summer months when there are less precipitations and more recreational activities on the water. They are often enforced through contracts and agreements with local communities. The bounds may be fixed or variable over the entire year, but it is reasonable to keep them fixed since the time horizon is sufficiently small.

Upper bounds ($\bar{v}_j$) represent hard maximum levels and are based on the physical capacity of each river segment or reservoir. Those of reservoirs close to human habitations that display a high likelihood or flooding are particularly important. These bounds are also critical for large "head" reservoirs, since operators will lose the capacity to manage the river adequately if the volumes exceed these thresholds.

Flow conservation constraints ensure conservation of water across the system. Since plants are located at different distances from one another on heterogeneous terrain, a drop of water takes a different amount of time to flow down to the next reservoir depending on the plant from which it was released. This phenomenon is modelled with the parameter $\lambda_{i^-(j)l}$ which represents the fraction of the total water discharged from plant $i^-(j)$ reaching the unique downstream reservoir j after l time steps. $\delta^{min}_{i^-(j)}$ and $\delta^{max}_{i^-(j)}$ represent the minimum and maximum number of days required to transport one

hm^3 of water from plant $i^-(j)$ to the downstream reservoir. The parameter ξ_{jt} represents the inflows in reservoir j at time t due to precipitations, snow melt and natural water flow.

Operators realize the importance of considering these delays and planning several weeks or months ahead, even in the case of certain inflows. This is an important advantage over stochastic dynamic programming (SDP) which is one of the most popular frameworks used to solve stochastic reservoir management problem (see (1)).

The constant β allows conversion of $m^3/seconds$ to $hm^3/days$ while v_{j0} represents the fixed known amount of water (in hm^3) at the beginning of the time horizon in reservoir j. We reasonably neglect evaporation and other water losses by assuming a relatively cold and humid climate, small water surface area, relatively high pressure and non-porous soil.

Minimum flow bounds $\underline{r}_i$ represent the smallest quantity of water required to flow down site i to produce electricity, for environmental purposes, as well as to ensure navigation and recreation. The upper bound $\bar{r}_i$ represents a critical threshold based on historical observations and used for safety and operational purposes. For instance, operators may prefer to maintain relatively small water flows during the winter season to prevent ice coming into contact with the installations at high speed and to ensure a smooth ice formation.

Evacuation curves $C_i(\cdot)$ represent the maximal amount of water $\mathcal{C}_i(v)$ that can be physically discharged (spilled) from the evacuator at site $i \in I^{evac}$ for a given amount of (average) volume v during any given time t at the unique upstream reservoir where $I^{evac} \subset I$ represents the set of plants with such constraints. These functions are typically smooth, increasing, non-linear and non-convex. We choose to approximate them with affine functions as this provides a reasonable approximation while ensuring the linear structure of theses constraints.

Variation bounds Δ_i are added to ensure that the total flow of water at each plant i stays relatively constant from one time period to the next. Maintaining a relatively constant water flow is important to prevent turbine breakage as well as to ensure navigation.

2.2 Considering floods

Under very wet scenarios and given initial conditions, the tight operational constraints on flows, volumes, spillage and turbined outflow cannot be respected. Even with perfect foresight, numerical experimentation on a real river basin reveal that the wait-and-see solution may be infeasible if the initial volumes are sufficiently high (see section (6.1)). Hence, we modify the constraints of the original problem to allow controlled violations of volume constraints:

$$\text{(Vol. Bounds')} \qquad \underline{v}_j \leq \mathcal{V}_{jt} - \mathcal{E}^+_{jt} \leq \bar{v}_j \quad \forall j \in J, t \tag{9}$$

$$\text{(Floods)} \qquad 0 \leq \mathcal{E}^+_{jt} \qquad \forall j, \forall t \tag{10}$$

The overflows/floods $\mathcal{E}^+_{jt}$ represent the quantity of water (hm^3) exceeding the maximum volume thresholds. It is reasonably assumed that excess water will remain in the reservoirs and may eventually be released downstream. It is straightforward to extend this to droughts as well. However, for our particular application, violating lower volume bounds is highly undesirable and only very limited violations are physically possible. We therefore omit them in our formulation.

2.3 Multiobjective optimization

We seek a solution that minimizes the occurrence of these floods. This approach is tantamount to numerous hydro electrical complexes in the presence of nearby riparian human populations and infrastructure as well as fauna and flora. This leads to a multiobjective problem where we must simultaneously consider floods over all reservoirs and time periods. These objectives may be conflicting as variations in volumes upstream may have important cascading effects (pun intended) on downstream reservoirs.

It is also crucial to consider the relative importance of floods as they may have widely varying consequences depending on the segments of the river where they occur, the time of the year and their magnitude. We tackle the multi-objective nature of the problem by considering the positive linear combinations of floods $\hat{\Lambda} : \mathcal{E} \to \sum_{j \in J} \sum_{t=1}^{T} (\kappa^+_{jt} \mathcal{E}^+_{jt})$ with parameters $\kappa^+_{jt} > 0$, $\forall j, t$ and $\mathcal{E} = (\mathcal{E}^+_{11}, ..., \mathcal{E}^+_{|J|T})^\top \in \mathbb{R}^{|J|T}$.

To reflect the importance of reservoirs located near riparian populations and high risks of floods as well as those with critical importance, we define the set J^{crit} and fix $\kappa^+_{jt} = W\kappa^+$, $\forall j \in J^{crit}, t$ for some large $W \in \mathbb{N}$. Since the relative importance of upper volume bounds at each of the reservoirs in $J \setminus J^{crit}$ is similar, we fix $\kappa^+_{jt} = \kappa^+$, $\forall j \in J \setminus J^{crit}, t$. For our specific application, this categorization is clearly defined, but identifying the relative importance of violations may require using preference assessment tools such as MACBETH ((22)). We choose κ^+ such that the sum of weights is equal to 1 and keep it constant with respect to time since we consider a small time horizon.

3 Incorporating inflow uncertainty

In most river systems, there is a considerable amount of uncertainty surrounding inflows. This randomness is one of the main driving forces in the variability of the state of the system. For these reasons, we let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space where $\xi_t : \Omega \to \mathbb{R}$, $\forall t \in \{1, ..., T\}$ represents the *total inflows* (in m^3/s) for the river basin at time t, $\mathcal{F}$ is a σ-algebra and $\{\mathcal{F}_t\}$ is the natural filtration induced by the stochastic process $\{\xi_t\}_{t=1,...,T}$ where $\mathcal{F}_t = \sigma(\xi_s, s < t) \subset \mathcal{F} = \mathcal{F}_T$. The random vector $\xi_{[t]} = (\xi_1, ..., \xi_t)^\top$ represents the past inflows up time t and $\xi \equiv \xi_{[T]}$.

We denote E the expectation operator and assume that $E[\xi_t] = \mu_t < \infty$ and $E[(\xi_t - \mu_t)^2]^{\frac{1}{2}} = \sigma_t < \infty$ for all t. These moments are estimated with their corresponding non-parametric sample statistics. For the random variable $\xi_t : \Omega \to \mathbb{R}$, we denote $\operatorname{ess\,sup} \xi_t = \inf\{a \in \mathbb{R} : \mathbb{P}(\xi_t > a) = 0\}$. We assume $\zeta_t = \xi_t - \mu_t$ is essentially bounded and that the compact and convex set $\mathcal{Z}_\nu \subset \mathbb{R}^T$ represents the true support of the probability measure $P_\zeta = \mathbb{P} \circ \zeta^{-1}$ with fixed parameter $0 \leq \nu < \infty$. More precisely, we assume $\mathcal{Z}_\nu = \{\zeta \in \mathbb{R}^T : \zeta_t \in [-\min\{\nu\sigma_t, \mu_t\}, \nu\sigma_t], \forall t\}$ since $\xi_t = \zeta_t + \mu_t \geq 0$ with P_ζ a.s. $\forall t$.

We assume a relatively small river basin where inflows at each reservoir are perfectly correlated. Although our method can easily consider the case of independent or partially spatially correlated inflows, this would increase the computational burden and complicate the exposition of the problem. In the case of perfectly spatially correlated inflows, we can represent ξ_{jt}, the total water inflow in m^3/s from natural precipitations and spring thaw going in

reservoir j at time t as $\xi_{jt}(\zeta) = \mu_{jt} + \zeta_t \alpha_{jt}$ where $\alpha_{jt} = \mu_{jt}(\sum_j \mu_{jt})^{-1}$ is the expected proportion of the total inflows entering the reservoir at that time.

4 Risk analysis

We formalize the notion of risk associated with the floods over all reservoirs and times through *systemic* risk measures. Before formally defining this concept, we define one dimensional risk measures defined on $\mathcal{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}) \equiv \mathcal{L}^\infty$, the vector space of essentially bounded real valued random variables. In the context of flood management, these random variables can be interpreted as floods.

4.1 Measuring flood risk for one reservoir and one time period

Risk measures are functionals $\rho_0 : \mathcal{L}^\infty \to \mathbb{R}$ that evaluate the riskiness of uncertain floods scenarios. The risk of a flood $\mathcal{U} \in \mathcal{L}^\infty$ at a given reservoir and arbitrary time t is "acceptable" under ρ_0 if it lies within the set $\{\mathcal{U} \in \mathcal{L}^\infty : \rho_0(\mathcal{U} \leq 0)\}$ (23). For $\mathcal{U}, \mathcal{W} \in \mathcal{L}^\infty$, risk measures minimally respect the following axioms:

$$\mathcal{U} \geq \mathcal{W} \Rightarrow \rho_0(\mathcal{U}) \geq \rho_0(\mathcal{W}) \qquad \text{(Monotonicity)} \qquad (11)$$

$$\rho_0(\mathcal{U} + c) = \rho_0(\mathcal{U}) + c, \forall c \in \mathbb{R} \qquad \text{(Translation invariance)} \qquad (12)$$

Where the inequality $\mathcal{U} \geq \mathcal{W}$ means $\mathcal{U}(\omega) \geq \mathcal{W}(\omega)\, \forall \omega \in \Omega$. We restrict ourselves to coherent risk measures, which also respect the two following axioms:

$$\rho_0(\lambda\mathcal{U} + (1-\lambda)\mathcal{W}) \leq \lambda\rho_0(\mathcal{U}) + (1-\lambda)\rho_0(\mathcal{W}), \forall \lambda \in [0,1] \text{ (Convexity)} \qquad (13)$$

$$\rho_0(\lambda\mathcal{U}) = \lambda\rho_0(\mathcal{U}), \forall \lambda \geq 0 \qquad \text{(Positive homogeneity)} \qquad (14)$$

We consider the popular (one-dimensional) risk measure CVaR (24; 25). If $\mathcal{U}, \mathcal{W} \in \mathcal{L}^\infty$ are continuous random variables, then $\text{CVaR}_\alpha(\mathcal{U})$ represents the conditional expectation of $\mathcal{U}$ given it is greater than the $(1-\alpha)$ left quantile $\text{VaR}_\alpha(\mathcal{U}) = \inf\{u \in \mathbb{R} : \mathbb{P}(\mathcal{U} \leq u) \geq 1-\alpha\}$ for a given parameter $\alpha \in (0,1)$. CVaR_α is a coherent risk measure that respects the four basics

axioms described previously in 11 - 14 . For a continuous random variable $\mathcal{U}$, we also have:

$$\lim_{\alpha \to 1} \text{CVaR}_\alpha(\mathcal{U}) = E[\mathcal{U}] \qquad \text{and} \qquad \lim_{\alpha \to 0} \text{CVaR}_\alpha(\mathcal{U}) = \text{ess sup}\,\mathcal{U} \tag{15}$$

which correspond to the natural risk neutral approach and to the totally risk averse approach often used in robust optimization. Both E and ess sup also correspond to coherent risk measures. Furthermore, we have the relationship $E[\mathcal{U}] \leq \text{CVaR}_\alpha(\mathcal{U}) \leq \text{ess sup}\,\mathcal{U}$, $\forall \alpha \in (0, 1)$. CVaR therefore represents a very natural and flexible risk measure and offers the possibility to explore the impact of various risk tolerance and model conservativeness through sensitivity analysis on the α parameter.

In the case of $\rho_0 \equiv \text{CVaR}_\alpha$, we can interpret $\rho_0(\mathcal{W}) = E[\mathcal{W}|\mathcal{W} > \text{VaR}_\alpha(\mathcal{W})]$, the risk of the flood $\mathcal{W} \in \mathcal{L}^\infty$, as the average of the worst $\alpha 100\%$ of flood values for some $\alpha \in [0, 1]$(see (25)). As observed in financial applications, CVaR_α therefore offers the important property of considering all the realizations of a random variable in the "worst" α-tail of the distribution. As $\alpha \to 1$, we consider more realizations, but reduce the weight of extremely bad realizations.

The monotonicity of ρ_0 reasonably implies that guaranteed smaller floods are preferred. Positive homogeneity implies that for $\mathcal{W}, \mathcal{U} \in \mathcal{L}^\infty$, if the risk of $\mathcal{W}$ is larger than the risk of $\mathcal{U}$ then the risk of $\lambda\mathcal{W}$ is larger than the risk of $\lambda\mathcal{U}$ for all $\lambda > 0$. Translation invariance implies that increasing floods by a given deterministic amount increases the risk by the same deterministic amount. This makes sense since the risk of a future flood known exactly should be treated separately from the risk of an uncertain one. In addition. since $\rho_0(\mathcal{W} + c)$ represents the minimal deterministic amount of flood that must be removed in order to make the risk of $\mathcal{W} + c$ acceptable for a $c \in \mathbb{R}$ and $\mathcal{W} \in \mathcal{L}^\infty$, we know that we need to remove exactly c from the risk of $\mathcal{W}$ for $\mathcal{W} + c$ to be acceptable. Finally, the convexity of ρ_0 says that we prefer diversifying and that the perceived risk of a convex combination of floods is no larger than a convex combination of the individual flood risks. See (26) for more details on interpreting risk measures.

4.2 Measuring flood risk for a multi-reservoir complex over a fixed time horizon

Although there exists a sizeable literature on multidimensional risk measures (see namely (27; 21)), we chose to focus on the concept of *systemic* risk measures introduced by (21). Following their work, we consider the vector space $\mathcal{L}_O^\infty \equiv \mathcal{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^O)$ of essentially bounded $\mathbb{R}^O$ valued random vectors which represent floods over the different reservoirs of the river system at different time steps of the planning horizon.

We then define the multidimensional risk measure $\rho = (\rho_0 \circ \Lambda) : \mathcal{L}_O^\infty \to \mathbb{R}$ defined as the composition of the unidimensional risk measure $\rho_0 : \mathcal{L}^\infty \to \mathbb{R}$ and the aggregation function $\Lambda : \mathcal{L}_O^\infty \to \mathcal{L}^\infty$ for some $O \in \mathbb{N}$. We consider systemic risk measures that are convex, positively homogeneous and monotone in the sense that for $\mathcal{U}, \mathcal{W} \in \mathcal{L}_O^\infty$, $\mathcal{U}_i \geq \mathcal{W}_i, i = 1, ..., O \Rightarrow \rho(\mathcal{U}) \geq \rho(\mathcal{W})$ [2].

If we consider the composition of $\rho_0 \equiv \text{CVaR}_\alpha$ and $\Lambda : \mathcal{U} \to \sum_{j \in J} \sum_{t=1}^T (\kappa_{jt}^+ \mathcal{U}_{jt})$, for some $\mathcal{U} = \text{vec}(\mathcal{U}_1, ..., \mathcal{U}_T) \in \mathcal{L}_{|J|T}^\infty$ with $\mathcal{U}_t \in \mathcal{L}_{|J|}^\infty$, $\forall t$ and $O = |J|t$, which is basically the functional defined in section 2.3, then $\rho = \text{CVaR}_\alpha \circ \Lambda$ is a systemic risk measure respecting convexity, monotonicity and positive homogeneity. Indeed, the aggregation function Λ is non-decreasing in the sense that $\mathcal{U}_{jt} \geq \mathcal{W}_{jt}, \forall j, t \Rightarrow \Lambda(\mathcal{U}) \geq \Lambda(\mathcal{W})$, for all $\mathcal{U}, \mathcal{W} \in \mathcal{L}_O^\infty$. Λ is linear and therefore positively homogeneous and convex. The systemic risk measure $\rho = \rho_0 \circ \Lambda$ is therefore monotone, convex and positive homogeneous since $\rho_0 \equiv \text{CVaR}_\alpha$ is a coherent one-dimensional risk measure.

Given, $\mathcal{W} \in \mathcal{L}_O^\infty$, a random vector of floods at different periods and locations, $\rho(\mathcal{W})$ can naturally be interpreted as the risk associated to $\mathcal{W}' = \sum_{j \in J} \sum_{t=1}^T \kappa_{jt}^+ \mathcal{W}_{jt} \in \mathcal{L}^\infty$, the flood at a single aggregated artificial reservoir which is simply a weighted combination of the different reservoirs over the entire time horizon weighted by their relative importance. For the rest of the paper, we fix $\text{CVaR}_\alpha(\sum_{j \in J} \sum_{t=1}^T \kappa_{jt}^+ \mathcal{E}_{jt}^+)$ as the objective function for a fixed $\alpha \in [0, 1]$.

[2] Convexity, translation invariance and positive homogeneity of systemic risk measures are defined by directly adapting the corresponding definitions from scalar valued risk measures.

4.3 Dynamic risk measures and time consistency

Our optimization model evaluates the systemic risk measure ρ at the beginning of the entire horizon once in a static fashion. This approach may lead to solutions that violate the concept of time-consistency. This notion is related to Bellman's principle and can loosely be interpreted as the requirement that optimal decisions taken at a given time for a fixed horizon should remain optimal when the problem is solved again at any other latter date (see (28) for more details on dynamic risk measures). Although this issue is important in a dynamic decision making environment, we omit to address it in our model because there has been little work on tractable representations of time-consistent stochastic problems using dynamic risk measures except in simple cases such as serially independent random variables when considering the risk neutral expectation (see (29)).

5 Stochastic model and choice of decision rules

We consider a dynamic setting were the true realization of the random process is progressively revealed as time unfolds over the horizon of T days. A sequence of controls $\mathcal{X}_1, ..., \mathcal{X}_t$ must be fixed at each stage $t = 1, ..., T$ after observing the realized history $\zeta_{[t-1]}(\omega) = (\zeta_1, ..., \zeta_{t-1})^\top(\omega)$, but before knowing the future values $(\zeta_t, ..., \zeta_T)^\top(\omega)$. At each time $t = 1, ..., T$, we optimize over the class of bounded functions of the form $\mathcal{X}_t : \mathbb{R}^T \to \mathbb{R}^{n_t}$, $\sum_{t=0}^{T-1} n_t = n$. Hence, we convert the decision variables of our deterministic model[3]into more flexible functions.

Real implementable decisions must be non-anticipative such that $\mathcal{X}_t \circ \zeta$ must be $\mathcal{F}_{t-1}$ measurable. Once $\zeta_{[t-1]}(\omega)$ is known, they can be implemented to yield the actual decisions $\mathcal{X}_t(\zeta_{[t-1]}(\omega)) \in \mathbb{R}^{n_t}$. We may interchangeably refer to $\mathcal{X}^\zeta = (\mathcal{X} \circ \zeta)$ or $\mathcal{X}$ as controls or decision rules and use the notation $\mathcal{X} = (\mathcal{X}_1, ..., \mathcal{X}_T)^\top$. The aggregate decision rule $\mathcal{X}_t = (\mathcal{V}_t, \mathcal{S}_t,\ \mathcal{Q}_t, \mathcal{R}_t, \mathcal{E}_t^-,\ \mathcal{E}_t^+)^\top$ is simply the stacking of each decision at time t. Refer to table (1) for identification of implementable decisions.

[3]We can see the decision variables of our deterministic model as constant functions of the uncertainty.

5.1 Affine decision rules

In order to obtain an upper bound on the optimal value of our stochastic programming problem, we first consider decision rules that are affine in ζ. If $\mathscr{K}_t = \{n_{t-1}+1, ..., n_{t-1}+n_t\}$ represents the indices associated with decisions at time $t \geq 1$ and $n_0 = 0$, we can write the controls in the form:

$$\mathcal{X}_{kt}(\zeta) = \mathcal{X}^0_{kt} + \sum_{t'=1}^{T} \mathcal{X}^{t'}_{kt}\zeta_{t'}, \quad k \in \mathscr{K}_t$$

where $\mathcal{X}^0_{kt}, \mathcal{X}^{t'}_{kt} \in \mathbb{R}$, $\forall k \in \mathscr{K}_t, t' \in \{1, ..., T\}$. In order to ensure that controls that represent real implementable decisions are non-anticipative, it is sufficient to force $\mathcal{X}^{t'}_{kt} = 0$, $\forall k \in \mathscr{K}^I_t, t = 1, ..., T, t' \geq t$ where $\mathscr{K}^I_t \subset \mathscr{K}_t$ represents the set of indices associated with real implementable decisions.

The volumes, droughts and floods can also be anticipative, since they represent analysis/auxiliary variables that are only required for computation (refer to table (1)). Because of the full-dimensionality of $\mathcal{Z}_\nu$ and the additive uncertainty, flow conservation constraints (2) are satisfied by solving a non-homogeneous linear system of equations in $\mathcal{X}^{t'}_{kt}, \mathcal{X}^0_{kt}$ (see (30) for example).

We reformulate the variation of flow constraints across time by writing :

$$(\mathcal{R}_{it} - \mathcal{R}_{it-1})(\zeta) \leq \bar{\Delta}_i, \qquad \forall i, t = 2, ..., T \tag{16}$$

$$-(\mathcal{R}_{it} - \mathcal{R}_{it-1})(\zeta) \leq \bar{\Delta}_i, \qquad \forall i, t = 2, ..., T \tag{17}$$

where all constraints must hold $\forall \zeta \in \mathcal{Z}_\nu$. As mentioned previously, we consider affine approximations of the real maximum spillage constraints: $\mathcal{C}_i : v \to \mathcal{C}^0_i + \Delta\mathcal{C}_i v$ with $\mathcal{C}^0_i, \Delta\mathcal{C}_i \in \mathbb{R}$ for all plants $i \in I^{evac}$. Hence if $j^-(i)$ indicates the unique reservoir upstream of plant i, the following evacuation curve constraints must hold $\forall \zeta \in \mathcal{Z}_\nu$:

$$S_{it}(\zeta) \leq \mathcal{C}^0_i + \Delta\mathcal{C}_i \mathcal{V}_{j^-(i)t}(\zeta), \qquad \forall i \in I^{evac}, t \tag{18}$$

For $\alpha \in (0, 1)$ and random variable $X \in \mathcal{L}^\infty$, we can write $\text{CVaR}_\alpha(X) = \inf_{t\in\mathbb{R}}\{t + \frac{1}{\alpha}E[\max\{X - t, 0\}]\}$. We use the definition of (31), but replace $1-\alpha$ by α so that the CVaR_α we consider in continuous and non-increasing in α. Hence, we obtain a conservative (upper bound) on $\text{CVaR}_\alpha(\sum_{j\in J}\sum_{t=1}^{T}(\kappa^+_{jt}\mathcal{E}^+_{jt})$

by adding the decision variables $\varphi, t \in \mathbb{R}$ as well as the anticipative decision rule $\mathcal{D}(\zeta) = \mathcal{D}^0 + \sum_{t'=1}^{T} \mathcal{D}^{t'} \zeta_{t'}$ and adding the constraints:

$$t + \frac{1}{\alpha}\mathcal{D}^0 \leq \varphi \tag{19}$$

$$\mathcal{D}(\zeta) \geq 0 \tag{20}$$

$$\mathcal{D}(\zeta) \geq \sum_{j \in J} \sum_{t=1}^{T} (\kappa_{jt}^{+} \mathcal{E}_{jt}^{+}(\zeta)) - t \tag{21}$$

that must hold $\forall \zeta \in \mathcal{Z}_\nu$ and set φ as the objective function [4]. This modelling approach is namely similar to that of (31). This formulation is correct because we consider affine decision rules and $E[\zeta_t] = 0, \forall t$ without loss of generality.

The case $\alpha = 0$ (ess sup) is handled separately by adding the variable φ and the constraint $\varphi \geq \sum_{j \in J} \sum_{t=1}^{T} (\kappa_{jt}^{+} \mathcal{E}_{jt}^{+}(\zeta))$ while the case $\alpha = 1$ (expected value) is treated by adding φ and $\varphi \geq \sum_{j \in J} \sum_{t=1}^{T} (\kappa_{jt}^{+} \mathcal{E}_{jt}^{0+})$. In all cases, we use this epigraph form and set φ as the objective value to minimize.

The rest of the constraints are readily obtained from the deterministic problem where we replace decision variables with their associated affine decision rules. Deriving the deterministic equivalent program is straightforward by using well known methods from robust optimization. For all the inequality constraints that are meant to hold $\forall \zeta \in \mathcal{Z}_\nu$, we namely use strong linear programming duality, which is possible since $\mathcal{Z}_\nu$ is a non-empty compact polyhedron.

5.2 Affine decision rules on lifted probability space

In order to obtain a better upper bound on the optimal solution of our risk averse stochastic program, we consider a special case of the lifted decision rules of (18). The general idea is to apply a non-linear (bijective in our case) transformation $L^{'} : \mathbb{R}^T \to \mathbb{R}^{rT}$ for some $r \in \mathbb{N}$ to $\mathcal{Z}_\nu$ in order to lift each realization of the random vector $\zeta(\omega) \in \mathbb{R}^T$ onto a higher dimensional space $\mathbb{R}^{rT}$. We then define the retraction operator $R^{'} : \mathbb{R}^{rT} \to \mathbb{R}^T$ such that

[4] The approximation bounds the true CVaR from above, because we restrict ourselves to a limited class of decision rules and (20) - (21) must hold $\forall \zeta \in \mathcal{Z}_\nu$.

$(R^{'} \circ L^{'}) = I$. We fix our initial controls to be affine functions of the *new* lifted random variables and impose that the retracted lifted random vector $(R^{'} \circ L^{'})\zeta$ be in the original set $\mathcal{Z}_\nu$. We can then express the robust counterpart directly in terms of these new random variables.

The heart of the problem is then to judiciously define $L^{'}$ and the corresponding $R^{'}$ in order to maintain the tractability of the new lifted program while ensuring an efficient representation of the feasible domain of the lifted random variables and improving the upper bound on the optimal solution of our program.

Following the approach of (18), we choose to fix $L^{'} = L \circ F$ where L is a non-linear bijective mapping and $F : \mathbb{R}^T \to \mathbb{R}^T$ is a simple bijective linear mapping (F^{-1} is a $T \times T$ matrix which exists) preserving the non-anticipativity of decision rules ($F_{ij} = 0$, $\forall i > j$) and which can be interpreted as a change of basis.

For our particular experiments, we consider $F = I$ and the lower triangular mapping $F_{ij} = 1$ for $i \leq j$ and 0 when $i > j$ $\forall j = 1, ..., T$. In the latter case, we can interpret each $F_t\zeta = \zeta_t^{'}$ as the cumulated history from time 1 until t. We present additional details on the linear mapping F in section (6.3).

The non-linear operator $L : \mathbb{R}^T \to \mathbb{R}^{rT}$ is defined through the partial lifting $L_t : \mathbb{R} \to \mathbb{R}^r$ as follows:

$$L_{tk}(\zeta^{'}) = \begin{cases} \min\{\zeta_t^{'}, z_1^t\} & k = 1 \\ \max\{0, \min\{\zeta_t^{'}, z_k^t\} - z_{k-1}^t\} & k = 2, ..., r-1 \\ \max\{\zeta_t^{'} - z_{k-1}^t, 0\} & k = r \end{cases}$$

for all t where $z_1^t < z_2^t < ... < z_{r-1}^t$ are breakpoints dividing the range of the original $\zeta_t^{'}$ in r dimensions. If we consider only one breakpoint ($r = 2$) and fix $z_1^t = 0, \forall t$, then the lifting is equivalent to decomposing every $\zeta_t^{'}$ into its negative and positive parts.

Similarly, we define the partial retraction operator $R_t : \mathbb{R}^r \to \mathbb{R}$, $\forall t$ as follows:

$$R_t(\zeta^{''}) = \sum_{k=1}^{r} \zeta_{tk}^{''} \equiv \zeta_t^{'} = [F\zeta]_t = F_t\zeta$$

Where F_t is the t^{th} row of F. We notice that R_t is linear and we can represent R as a block angular $T \times rT$ matrix:

$$R = \begin{pmatrix} \underbrace{1 \quad \cdots \quad 1}_{r\,terms} & & & & & & \\ & 1 \quad \cdots \quad 1 & & & & & \\ & & & \ddots & & & \\ & & & & & & \\ & & & \ddots & & & \\ & & & & & 1 \quad \cdots \quad 1 \end{pmatrix}$$

To derive the robust equivalent of the lifted problem, we need an efficient representation of the convex hull of $L'(\mathcal{Z}_\nu)$. We know that $F\mathcal{Z}_\nu \subseteq \bigtimes_t [l_t, u_t]$ where $l_t = \min_{\zeta \in \mathcal{Z}_\nu} F_t^T \zeta$ and $u_t = \max_{\zeta \in \mathcal{Z}_\nu} F_t^T \zeta$ and $L \bigtimes_t [l_t, u_t] = \bigtimes_t L_t[l_t, u_t]$.

For every t, $L_t[l_t, u_t]$ is a connected non-convex set in $\mathbb{R}^r$ defined by the breakpoints $z_1^t, ...z_{r-1}^t$ and we have conv $L \bigtimes_t [l_t, u_t] = \bigtimes_t \text{conv} L_t[l_t, u_t]$ since the Cartesian product of convex sets is also convex. The convex hull of $L_t[l_t, u_t]$ is the set of convex combinations of the following $r+1$ affinely independent points in $\mathbb{R}^r$:

$$v_1^t = \begin{pmatrix} l_t \\ 0 \\ ... \\ 0 \end{pmatrix} v_2^t = \begin{pmatrix} z_1^t \\ 0 \\ ... \\ 0 \end{pmatrix} v_3^t = \begin{pmatrix} z_1^t \\ z_2^t - z_1^t \\ ... \\ 0 \end{pmatrix} \quad ... \quad v_{r+1}^t = \begin{pmatrix} z_1^t \\ z_2^t - z_1^t \\ ... \\ u_t - z_{r-1}^t \end{pmatrix}$$

If we regroup these points in the following invertible matrix:

$$V_t = \begin{pmatrix} 1 & ... & 1 \\ v_1^t & ... & v_{r+1}^t \end{pmatrix} \in \mathbb{R}^{(r+1)\times(r+1)}$$

then we have:

$$\begin{aligned} \text{conv} L_t[l_t, u_t] &= \{\zeta_t'' \in \mathbb{R}^r : (1, \zeta_t^{T''})^T = V_t \lambda, \lambda \in \mathbb{R}_+^{r+1}\} \\ &= \{\zeta_t'' \in \mathbb{R}^r : V_t^{-1}(1, \zeta_t^{T''})^T \geq 0\} \\ &= \{\zeta_t'' \in \mathbb{R}^r : b_t \zeta_t'' \geq -a_t\} \end{aligned}$$

where a_t represents the 1^{st} column of V_t^{-1} and $b_t \in \mathbb{R}^{(r+1)\times r}$ is the matrix composed of the last r columns of V_t^{-1}. We notice that $\text{conv}L_t[l_t, u_t]$ is a polyhedron and hence so is $\bigtimes_t \text{conv}L_t[l_t, u_t]$. In general, we have:

$$\begin{aligned}(L \circ F)\mathcal{Z}_\nu &= \{\zeta^{''} \in \mathbb{R}^{rT} : \exists \zeta \in \mathcal{Z}_\nu, \zeta^{''} = (L \circ F)\zeta\} \\ &\subseteq \{\zeta^{''} \in \mathbb{R}^{rT} : \exists \zeta^{'} \in [l, u] \supset F\mathcal{Z}_\nu, \zeta^{''} = L\zeta^{'}, \} \\ &\subseteq \{\zeta^{''} \in \mathbb{R}^{rT} : \exists \zeta^{'} \in [l, u] \supset F\mathcal{Z}_\nu, \zeta^{''} \in \bigtimes_t \text{conv}L_t[l_t, u_t]\} \\ &\equiv \hat{\mathcal{Z}}_\nu^{''}\end{aligned}$$

and so $(F^{-1} \circ R)\hat{\mathcal{Z}}_\nu^{''} = \{\zeta \in \mathbb{R}^T : \exists \zeta^{''} \in \hat{\mathcal{Z}}_\nu^{''} : (F^{-1} \circ R)\zeta^{''} = \zeta\} \supseteq \mathcal{Z}_\nu$. We have effectively constructed an exterior (conservative) convex (polyhedral because of the lifting) of the image of $\mathcal{Z}_\nu$ under the lifting $L^{'} = L \circ F$ and coincidently an exterior approximation of $\mathcal{Z}_\nu$ through $(F^{-1} \circ R)\hat{\mathcal{Z}}_\nu^{''}$.

If $\mathcal{Z}_\nu \subset \mathbb{R}^n$ is a some hyperectangle, then using $F = diag(\lambda_1, ..., \lambda_n)$ for $\lambda_i \neq 0, \forall i$ will guarantee that $F\mathcal{Z}_\nu = \bigtimes_t [l_t, u_t]$. Since $\bigtimes_t R_t\zeta'' \in F\mathcal{Z}_\nu$ when ζ'' is an extreme point of $\bigtimes_t \text{conv}L_t[l_t, u_t]$ for any piecewise linear lifting L, it follows that $(F^{-1} \circ R)\zeta_\nu^{''} \in \mathcal{Z}_\nu$ for all $\zeta_\nu^{''} \in \bigtimes_t \text{conv}L_t[l_t, u_t]$. By restricting our new affine decision rules to a certain class of functions, we obtain the initial (unlifted) problem with affine decision rules and so the lifted problem provides an upper bound that is at least as good as the one with the original affine decisions.

However, except for this rare case, we cannot guarantee that the lifting L' will improve the solution. Indeed, it may even do worse if the uncertainty set considered is larger than the original one. For instance, if we consider the symmetric 2-dimensional uncertainty set given by $\mathcal{Z}_\nu = [-1, 1]^2$ and use the mapping $F = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, then $F\mathcal{Z}_\nu \subset [\min_{\zeta\in\mathcal{Z}} F_1^\top\zeta, \max_{\zeta\in\mathcal{Z}} F_1^\top\zeta] \times [\min_{\zeta\in\mathcal{Z}_\nu} F_2^\top\zeta, \max_{\zeta\in\mathcal{Z}_\nu} F_2^\top\zeta]$ and the inclusion is strict since $(2, -1)^\top \notin F\mathcal{Z}_\nu \Leftrightarrow F^{-1}(2, -1)^\top = (1, -3)^\top \notin \mathcal{Z}_\nu$ and $(2, -1) \notin F\mathcal{Z}_\nu \Leftrightarrow F^{-1}(-2, 1)^\top = (-1, 3)^\top \notin \mathcal{Z}_\nu$. We also have $\mathcal{Z}_\nu \subset (F^{-1} \circ R)\hat{\mathcal{Z}}_\nu^{''}$ regardless of the lifting operator L used (see figure (1).

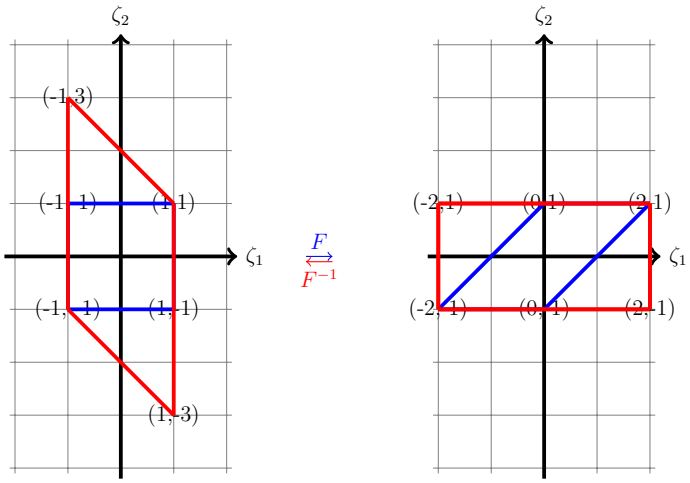

Figure 1: Deficiencies of lifting

6 Case study

6.1 The river system

We apply the risk averse reservoir management problem to the case of the Gatineau river in Québec. This hydro electrical complex is part of the larger Outaouais river basin and is managed by Hydro-Québec, the largest hydro-electricity producer in Canada (32). It is composed of 3 run-of-the river plants and 5 reservoirs, only 2 of which have significant capacity (see figure (2)).

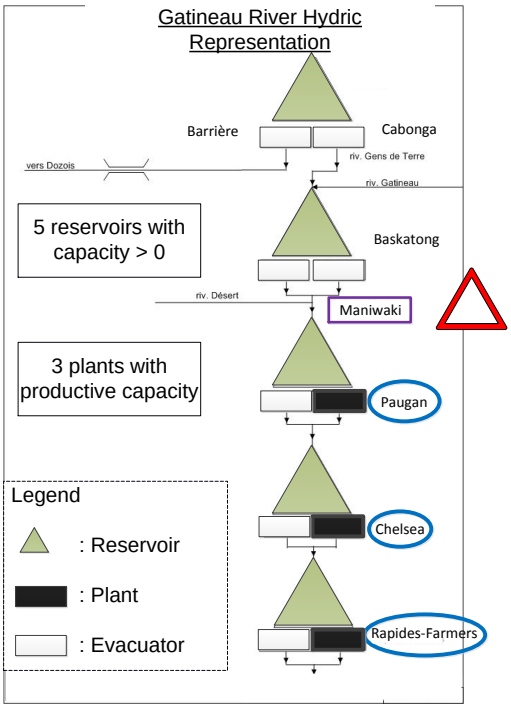

Figure 2: Simplified representation of the Gatineau river system

The Baskatong reservoir is the largest of the larger Outaouais-Gatineau catchment and plays a critical role in the management of the river. It is used to manage risk of floods during the freshet period as well as droughts during the summer months. It has been used to control baseflow at the greater Montréal region several hundreds of kilometres downstream. As such, respect of minimum and maximum water volume threshold is essential for river operators. We consider a time horizon of 30 days with daily time steps to reflect the real decision process for decisions such as setting water target releases for this reservoir during the freshet.

The maximum spillage at facilities Cabonga, Baskatong, Barrière and Rapides-Farmers is bounded by evacuation curves that reflect the particular structure of the dams and the associated reservoirs. As mentioned previously, we use affine functions to approximate the "real" curves used by Hydro-Québec (see figure 3). Our approximations are conservative on most of the feasible domain, since they bound the maximum spillage from below. Being completely

conservative for all feasible volumes would be impossible as this would lead to a very conservative or even infeasible solution.

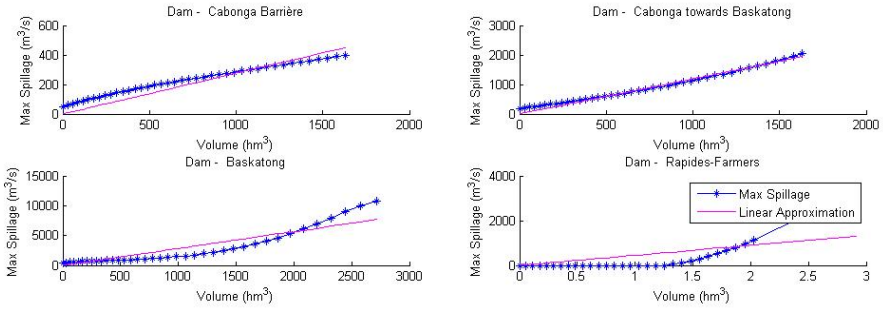

Figure 3: Evacuation curves

Although there are no structures to control water outflow near Maniwaki, there are extremely tight constraints on maximum and minimum flow at that river segment. These bounds represent critical outflow level used to control the risk of drought and flood for the neighbouring town and are therefore essential.

The Gatineau represents an excellent case study as it has relatively small productive capacity compared to the rest of the park. More importantly, the river runs near the small town of Maniwaki which is subject to high risks of flooding, particularly during the spring freshet. Indeed, the city has suffered 4 significant floods in 1929, 1936, 1947 and 1974.

Even today, there exists a need to mitigate the risk of these natural disasters. As figure (4) illustrates, it is possible that we fail to find a solution eliminating all floods even when we have *perfect foresight*. This is namely the case when $\zeta_t = \mu_t + 2\sigma_t$, $\forall t$. Although this synthetic scenario is wet over the entire month, it remains plausible because of the strong serial correlation of inflows as well as evidence from past observations. In this case, even if we spill as much as possible at Baskatong so that the total flow at Maniwaki downstream is constantly at its upper bound, the deterministic program cannot prevent the reservoir from filling up and eventually flooding.

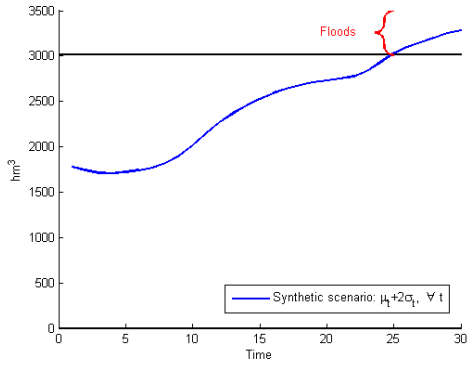

Figure 4: Violations with perfect foresight

6.2 The inflows

Based on our sample observations, the assumption of perfectly spatially correlated inflows is not unreasonable for the Gatineau river. As is clear from the figure (5), water inflows are particularly important during the months of March through April (freshet) as snow melts. There is a second surge during Fall caused by greater precipitations and finally there are very little inflows during the winter months. For this reason, we tested our model on the first 30 days of winter (lowest inflows, lowest variability) and the first 30 days of the spring freshet (highest inflows and highest variability). We estimate the true population mean and standard deviation by computing the sample mean and sample variance with these 6 years of data.

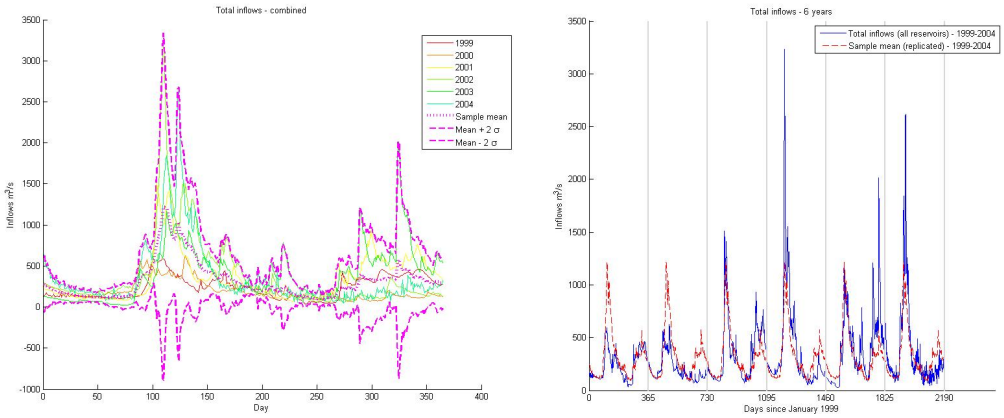

Figure 5: Sample observations (6 years) for inflows

6.3 Upper bounds on the risk of floods

We begin by presenting the optimal solutions obtained for 4 classes of decision rules: 1) the standard affine (dimension 1 i.e. $r = 1$), 2) the lifted decision rules with 1 breakpoint at 0 and the identity mapping ($r = 2$), 3) the lifted decision rules with 2 breakpoints at the 2/6 and 5/6 quantiles of the empirical distribution of $\zeta_t, \forall t$ and the identity mapping ($r = 3$), 4) the lifted decision rules with 1 breakpoint at 0 and the lower triangular mapping $F_{ij} = 1$ for $i \leq j$ and 0 when $i > j \, \forall j = 1, ..., T$ ($r = 2$). We attempted to solve the model with the same mapping F and $r = 3$, but ran into numerical issues.

All solutions where obtained by setting $\nu = 3$, which ensured the support contained all historical inflows from 1999-2004 and provided good quality solutions across all models. We tested our models with varying values of ν and found that increasing its value above a certain threshold lead to infeasibility issues as all bounds cannot be simultaneously respected for very large uncertainty sets. As expected, small ν's lead the model to predict small violations. However, the solutions displayed poor performance when tested on realistic inflows which did not belong to the uncertainty set. Our sensitivity analysis did not seem to reveal any clear link between the size of ν and the

performance of our models when tested with $\zeta \notin \mathcal{Z}_\nu$.

The values are displayed in line "Optimal solution" of table (2) and are good indicators of the CVaR of the *volume* of floods (in hm^3) at Baskatong predicted by the model and is useful to analyse the graphs of section (6.6). Since these solutions are obtained by restricting ourselves to specific family of decision rules, they represent upper bounds on the optimal value of the "true" problem with general decision rules.

	Affine - 1 dim			**Lifted - I - 2 dim**			**Lifted - I - 3 dim**			**Lifted - F - 2 dim**		
	Expected Value ($\alpha = 1$)	$CVaR_{0.5}$	Ess Sup ($\alpha = 0$)	Expected Value ($\alpha = 1$)	$CVaR_{0.5}$	Ess Sup ($\alpha = 0$)	Expected Value ($\alpha = 1$)	$CVaR_{0.5}$	Ess Sup ($\alpha = 0$)	Expected Value ($\alpha = 1$)	$CVaR_{0.5}$	Ess Sup ($\alpha = 0$)
Optimal solution	2.78E+03	5.57E+03	9.77E+03	1.11E+03	2.23E+03	9.77E+03	3.26E+02	6.52E+02	9.77E+03	1.09E+03	2.18E+03	3.74E+03
Optimal solution improvement (%)	-	-	-	60%	60%	0%	88%	88%	0%	61%	61%	62%
Time (seconds)	1.17E+01	6.86E+01	5.37E+02	3.09E+03	8.79E+03	9.95E+03	2.93E+03	4.07E+04	1.03E+05	1.66E+05	2.43E+05	2.30E+05
Time improvement (%)	0%	0%	0%	26367%	12722%	1751%	24966%	59256%	19057%	1422516%	354334%	42611%
Ratio (opt sol/time)	-	-	-	2.27E-03	4.71E-03	-1.00E-08	3.54E-03	1.49E-03	-9.22E-10	4.28E-05	1.72E-04	1.45E-03

Table 2: Optimal values [5]

[5] All results were obtained by using AMPL with solver CPLEX 12.4, on computers with 16.0 GB RAM and i7 CPU's @ 3.4GHz.

The line "Optimal solution improvement (%)" of table (2) shows the improvement with respect to the base affine policies for any fixed risk measure. We observe that the optimal solution of the lifted decision rules strictly improves when considering CVaR_α for $\alpha \in \{0.5, 1\}$ as the number of dimensions rises from 1 to 3 with the identity mapping. However, the optimal solution is constant with respect to the number of breakpoints with the identity mapping for the ess sup. The F mapping with dimension 2 offers the best theoretical bounds for the ess sup (62% improvement compared to the affine policies) while the identity mapping with 2 breakpoints offers the largest improvement (88%) for $\alpha \in \{0, 0.5\}$.

In all cases, the computational burden imposed by the lifted decision rules is much greater than the standard affine decision rules. Problems took up to 67 hours to solve with the F mapping and $r = 2$ while the same problem took only 68.6 *seconds* with the standard affine rules. Although it is surely beneficial to reduce the violations because of the potentially dramatic consequences of floods, computing times of several hours are unacceptable for practical operational purposes.

We explored acceleration techniques based on cutting plane methods as suggested by (30). However, our higher dimension uncertainty set made it difficult to identify and enumerate extreme scenarios explicitly a priori, even with the standard affine decision rules. Dynamic identification of violated inequalities therefore proved longer than full reformulation through duality.

6.4 Evaluating the policies through simulation

Following (33), the optimal solution of the problem can be useful to derive probabilistic bounds on the risk of floods and droughts. However, a small optimal value does not in itself guarantee the quality of the solutions provided by the problem since there are no guarantees when the uncertainty falls outside of the parametrized support considered in the model. It is imperative to simulate the behaviour of the system with the optimal decision rules under various scenarios to see how well it would perform different situations.

For this reason, we perform various tests using different scenario generators. We consider independent non-negative truncated normal random variable as well as independent log normal random variable. These random processes are

serially independent. As illustrated by figure (5), this is not an acceptable hypothesis since there exists strong cyclical effects combined with multi-lag autocorrelation. To remedy some of these problems and use a simulator closer to real observed inflows, we calibrate a SARIMA model on the raw inflow data (34). Once again, we fixed a $\nu = 3$ for all our models.

6.5 Comparison with stochastic model based on a simple scenario tree

In order to compare the value of affine decision rules, we consider a multi-stage stochastic programming model based on scenario trees. We consider a simple binary tree with branching occurring roughly each week over a horizon of 30 days. This implies that an entire week becomes known after each decision is made. At each branching period t, the tree considers two extreme and opposed scenarios with equal probability: $\mu_t + \nu\sigma_t$ or $\mu_t - \min\{\mu_t, \nu\sigma_t\}$. We limit the size of the tree since greater branching factors or more frequent branching quickly leads to severe numerical difficulties. Each problem took on average 1 second to solve with the scenario tree.

Since it is reasonable to suppose that inflows are jointly continuous, any of the scenarios considered in a stochastic tree model have probability 0 of materializing exactly. Using the solutions found by a single solution of the stochastic tree model in simulation is therefore highly likely to lead to poor quality solutions. To mitigate this effect, we simulate the use of this model by solving it in a rolling horizon fashion.

We also tested affine decision rules in a rolling horizon, but with a shorter lookahead period, namely 1 week. These problems took on average 5 seconds to solve and are therefore comparable to the scenario tree policies in terms of computing times.

6.6 Simulation results

We evaluate the empirical conditional value at risk of violations of different constraints at different levels of risk $\alpha' = 100(1-\alpha)$ with $\alpha \in \{0.01, ..., 0.99\}$. We also show the expected value at $\alpha' = 0$ and the ess sup at $\alpha' = 100$.

In all cases, we find that the theoretical upper bound found by solving the stochastic model is loose. Even if we do not restrict the random variables to lie within the support considered by our model, we find that the empirical CVaR_α is usually smaller than the theoretical upper bound.

For sake of space, we present only results for the independent log-normally distributed random variables for which the sample mean and variance are unbiased estimates of the true parameters. Results for the normal independent variables and the SARIMA time series are similar in nature.

Increasing the value of $\nu^{'}$ from its standard value changes the shape of the empirical distribution, but the relative performance of each model stays relatively constant. Moreover, excessively increasing $\nu^{'}$ may lead to scenarios that are more extreme than natural and that may have limited practical interpretation. The treatment of rare extreme events is an open question we do not wish to tackle here.

As figure (6) illustrates, when considering a single reservoir and a fixed decision rule, different risk measures may lead to solutions that are empirically dominated by others. For instance, minimizing the ess sup often leads to higher CVaR_α for total volume violations at Baskatong for all $\alpha \in [0, 1]$ than minimizing $\text{CVaR}_{0.5}$ or the expected value. However, solutions that perform well at a given segment often do worse at other river segments or reservoirs.

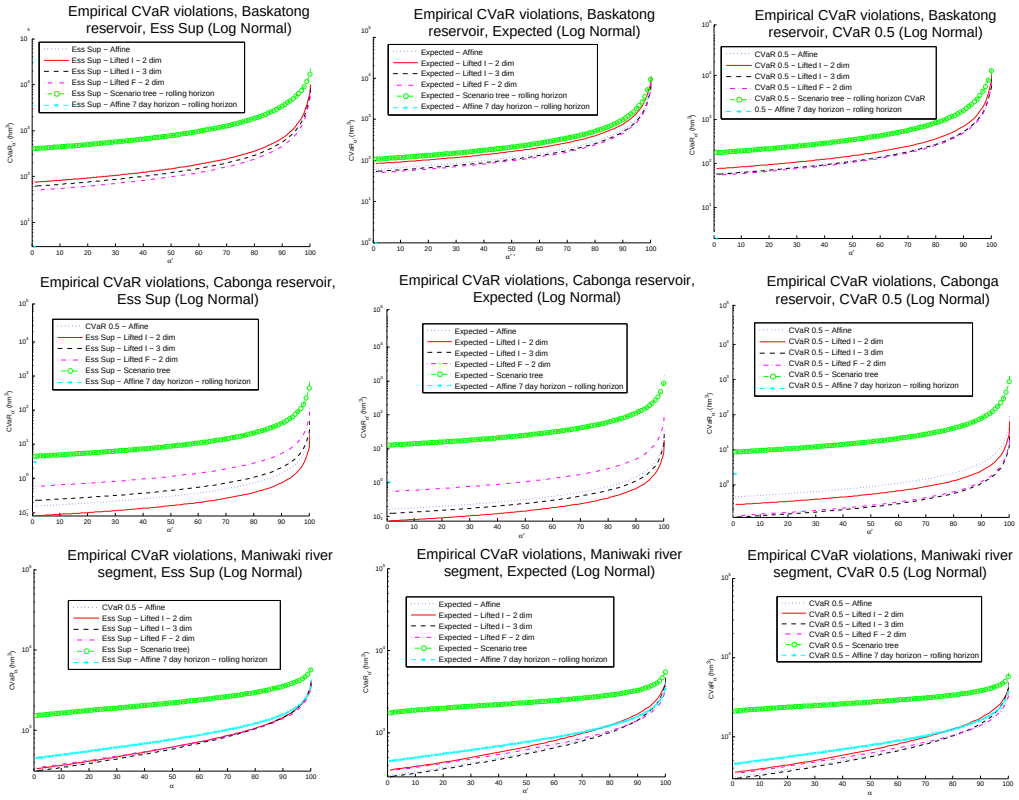

Figure 6: Empirical CVaR_α of violations with log-normally simulated random variables and $\alpha \in \{0, 0.5, 1\}$ [7]

[7] Violations are 0 for Baskatong and Cabonga with the affine policies with 7 day lookahead horizon used in a rolling horizon fashion for all risk measures.

Plotting the results for a given reservoir for all risk measures and decision rules does not clearly indicate superiority of any policy. Indeed, figure (6) make it hard to discriminate between decision rules. In particular, these results do not indicate that lifted decision rules are more performant than the basic affine decision rules. This is surprising as the the theoretical upper bounds computed by lifted decision rules are sizeably smaller than those returned by the affine ones.

This may be due to the fact that the lifted decision rules use the same information and representation of uncertainty as the standard affine controls. Moreover, the empirical CVaR_α results are chiefly influenced by a few (<4) extreme scenarios that fall outside of the support considered by the model. Lifted decision rules do not seem to provide better coverage than the basic affine controls against these extreme events.

However, our results clearly indicate the superiority of the affine and lifted decision rules with regards to the scenario tree model. Indeed, figure (6) highlights the systematically higher violations at Baskatong, Maniwaki and Cabonga for all risk measures compared to any decision rules. Even when solved in a rolling horizon fashion, the lack of adjutibility of the scenario tree provide unsatisfactory results.

Although they do not dominate all decision rules for all risk measures and risk levels, the 7 day lookahead affine policies used in rolling horizon provide excellent results. Limited computing ressources and time make it difficult to simulate policies considering a longer lookahead horizon, but it is likely that they would yield even better results.

6.7 Real scenarios results

We evaluate each policy on real historical observed scenarios. We consider the first 6 years 1999-2004 used to calibrate the model as well as 6 additional out-of-sample years 2008-2013 for validation. Once again, we only present the important reservoirs and river segments, since the other constraints at other reservoirs are always respected. We focus on the first 30 days of the freshet period since other periods of the year are usually free of any violations and we always present results for models using an objective of $\text{CVaR}_{0.5}$ since results are relatively similar for different choice of α.

As figure (7) illustrates, all types of decision rules violate some of the constraints on one of the 12 scenarios. Affine and lifted decision rules only violate constraints on the last 6 out-of-sample years (in red) since they have higher inflows while the scenario tree violate constraints on the 2 sets of data.

For the model based on scenario tree, there are no volume violations for any of the reservoirs. However, there are significant flow violations at Maniwaki. This is not surprising since the average flow computed by the tree model is always very close to the upper flow bound. By emptying the large upstream reservoir Baskatong, this policy reduces the risk of having volume violations upstream, but greatly increases the risk of violating minimum and maximum flow bounds near the city.

On the other hand, affine and lifted decision rules provide a generally adequate solution. Although there are maximum volume violations at the end of the time horizon for very wet years as well as maximum flow violations at periods of high inflow, these violations are usually more acceptable than those of the tree model. It is also rather surprising to see that lifted decision rules do not consistently outperform the simple affine decision rules.

We also include the trajectories for the deterministic model used in a rolling horizon fashion as well as the standard affine decision rules with a limited lookahead horizon of 1 week. The deterministic model does very poorly with regards to flow bound violations at Maniwaki. The limited affine model in rolling horizon takes a similar amount of computing resources as the decision tree, but seems to achieve a better trade-off between violations of volume bounds at Baskatong and flow bounds at Maniwaki.

Although we do not push the volumes towards a desired goals at the end of the time horizon, our solution never completely empties the reservoirs nor fills them up when physically possible. This is an advantage of the model since it yields more realistic water release schedules and does not necessarily require taking into account a future monetary value of water.

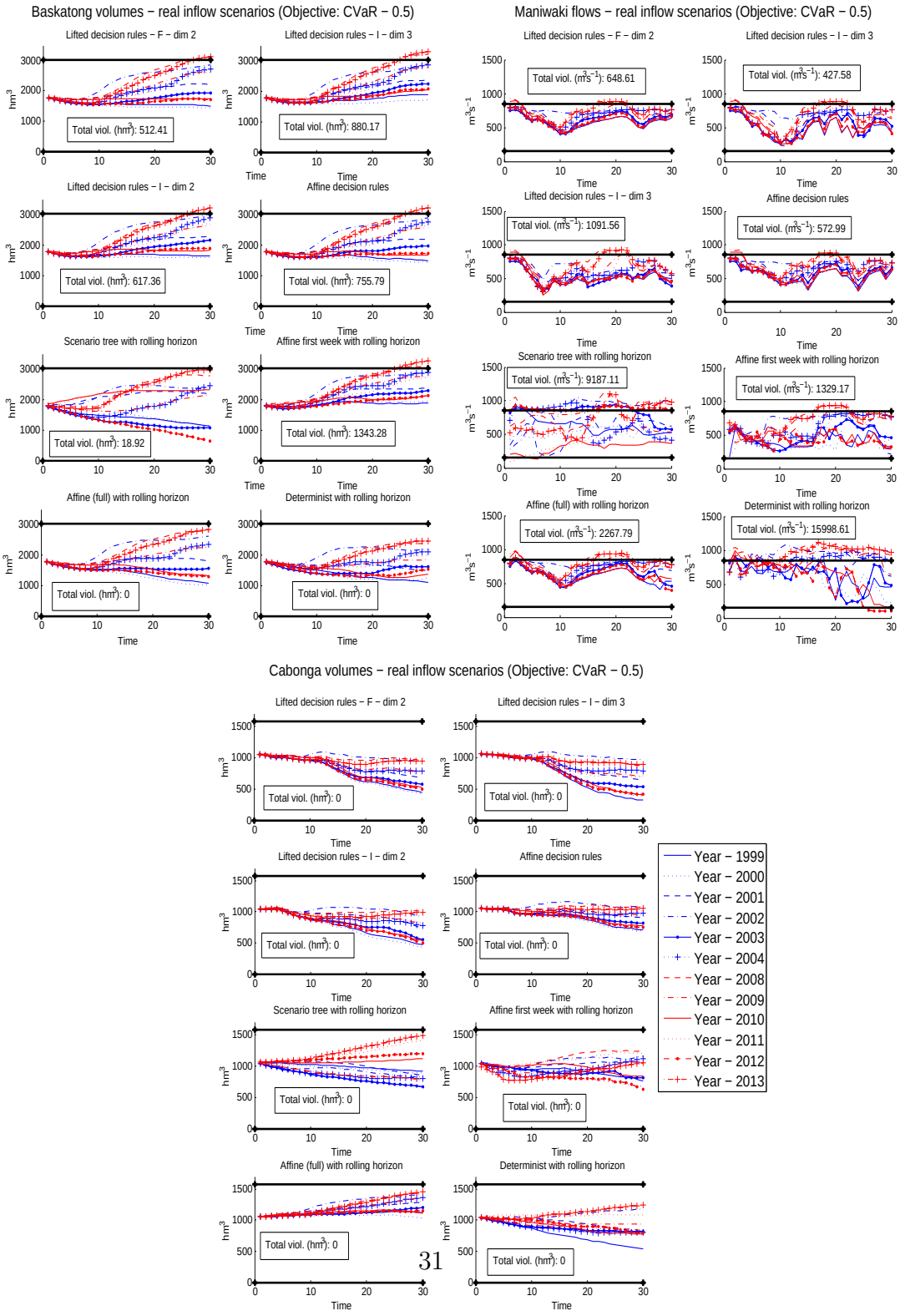

Figure 7: Actual values - 12 historical scenarios

7 Conclusion

In conclusion, we propose a novel distribution-free formulation for the risk averse stochastic reservoir management problem. Our model is extremely flexible, can easily be obtained from an existing deterministic model and is relatively easy to understand and implement.

Contrary to other popular methods such as stochastic dynamic programming, we can easily incorporate constraints involving different time steps, such as flow variation constraints. Our method also allows the consideration of objective functions such as CVaR that do not permit time decomposition.

As opposed to scenario tree models, we do not require discretization of the underlying stochastic process. Consequently, it is easier to represent the uncertainty and we do not have to use interpolation or a rolling horizon optimization to simulate the behaviour of our model, which can yield important computational advantages.

For the first time to the best of our knowledge, we also formalize the notion of multi-objective risk aversion for reservoir operations through *systemic risk measures*, which in our case is simply the composition of the coherent one dimensional risk measure CVaR and a strictly increasing linear aggregation function. This provides a simple parametrization of risk aversion and allows efficient sensitivity analysis.

Our various numerical experiments confirm the value of using affine decision rules to provide good implementable solutions while maintaining the tractability of the model. Our experiments also suggest that these decision rules offer little optimality loss over more complex piece-wise linear decision rules while providing extremely important computational savings. Affine and piece-wise linear decision rules seem to perform consistently better than decision trees, even when considering a rolling horizon framework.

8 Acknowledgments

The authors would like to thank everyone at Hydro-Québec and IREQ for their ongoing support, particularly Grégory Émiel, Louis Delorme and Pierre-

Marc Rondeau. This research was supported by the Natural Sciences and Engineering Reasearch Council of Canada (NSERC) and Hydro-Québec through the industrial research chair on the stochastic optimization of electricity generation.

9 References

References

[1] J. Labadie, Optimal operation of multireservoir systems: state-of-the-art review, Journal of water resources planning and management 130 (2004) 93–111.

[2] S. Hajkowicz, A. Higgins, A comparison of multiple criteria analysis techniques for water resource management, European Journal of operational research 184 (1) (2008) 255–265.

[3] A. Tilmant, R. Kelman, A stochastic approach to analyze trade-offs and risk associated with large-scale water resources systems, Water resources research 43 (2007) W06425.

[4] A. Castelletti, S. Galetti, M. Restelli, R. Soncini-Sessa, Tree-based reinforcement learning for optimal water reservoir operation, Water resources research 46 (2010) W09507.

[5] M. Dyer, L. Stougie, Computational complexity of stochastic programming problems, Mathematical Programming Series A.

[6] D. Bertsimas, D. Brown, C. Caramanis, Theory and applications of robust optimization, SIAM review 53 (2011) 464–501.

[7] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, Robust Optimization, Princeton University Press, 2009.

[8] M. Saad, A. Turgeon, Application of principal component anbalysis to long-term reservoir management, Water ressources research 24 (1988) 907–912.

[9] M. Saad, P. Bigras, A. Turgeon, Fuzzy learning decomposition for the scheduling of hydroelectric power system, Water ressources research 32 (1996) 179–186.

[10] A. Turgeon, R. Charbonneau, An aggregation-disaggregation approach to long-term reservoir management, Water resources research 34 (1998) 3585–3594.

[11] J. R. Stedinger, B. A. Faber, Reservoir optimization using sampling sdp with ensemble streamflow prediction (esp) forecast, Journal of Hydrology 249 (2001) 113–133.

[12] M. Pereira, L. Pinto, Multi-stage stochastic optimization applied to energy planning, Mathematical progamming 52 (1991) 359–375.

[13] R. E. Gonçalves, E. C. Finardi, E. L. da Silva, Applying different decomposition schemes using the progressive hedging algorithm to the operation planning problem of a hydrothermal system, Electric Power Systems Research 83 (1) (2012) 19 – 27.

[14] P.-L. Carpentier, M. Gendreau, F. Bastin, Optimal scenario set partitioning for multistagestochastic programming with the progressive hedging algorithm, Tech. rep., École Polytechnique de Montréal (2013).

[15] A. Ben-Tal, E. Goryashko, A. Guslitzer, A. Nemirovski, Adjustable robust solutions of uncertain linear programs, Mathematical progamming 99 (2004) 351–378.

[16] X. Chen, M. Sim, P. Sun, J. Zhang, A linear decision-based approximation approach to stochastic programming, Operations Research 56 (2008) 344–357.

[17] J. Goh, M. Sim, Distributionally robust optimization and its tractable approximations, Operations Research 58 (2010) 902–917.

[18] A. Georghiou, W. Wiesemann, D. Kuhn, Generalized decision rule approximations for stochastic programming via liftings, Mathematical Programming (2014) 1–38.

[19] R. Apparigliato, Règles de décision pour la gestion du risque: Application à la gestion hebdomadaire de la production électrique, Ph.D. thesis, École Polytechnique (2008).

[20] L. Pang, M. Housh, P. Liu, X. Chen, Robust stochastic optimization for reservoir operation, Water ressources research 51.

[21] C. Chen, G. Iyengar, C. Moallemi, An axiomatic approach to sysystem risk, Management Science.

[22] C. A. Bana e Costa, J.-C. Vansnick, A Theoretical Framework for Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH), Springer, 1997.

[23] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, Coherent measures of risk, Mathematical finance 9 (1999) 203–228.

[24] T. R. Rockafellar, S. Uryasev, Conditional value-at-risk for general loss distributions, Journal of banking and finance 26 (2002) 1443–1471.

[25] C. Acerbi, D. Tasche, On the coherence of expected shortfall, Journal of banking and finance 26 (2002) 1487–1503.

[26] E. Delage, J. Y. Li, Minimizing risk exposure when the choice of a risk measure is ambiguous, Tech. rep., Les cahiers du GERAD (2015).

[27] E. Jouini, M. Meddeb, N. Touzi, Vector-valued coherent risk measures, Finance and stochastics.

[28] B. Acciacio, I. Penner, Advanced Mathematical Methods for Finance, Springer, 2011, Ch. Dynamic Risk Measures, pp. 1–34.

[29] A. Shapiro, On a time consistency measures concept in risk averse multistage stochastic programming, Operations Research Letters.

[30] A. Lorca, A. Sun, E. Litvinov, T. Zheng, Multistage adaptive robust optimization for the unit commitment problem, Tech. rep., Georgia Institue of Technology and ISO New England (2014).

[31] G. Pflug, Some remarks on the value-at-risk and the conditional value-at-risk, in: S. Uryasev (Ed.), Probabilistic Constrained Optimization, Vol. 49 of Nonconvex Optimization and Its Applications, Springer US, 2000, pp. 272–281.

[32] Hydro-Québec, Hydro-québec, rapport annuel 2012 (2012).

[33] W. Chen, M. Sim, J. Sun, C.-P. Teo, From cvar to uncertainty set: implications in joint chance-constrained optimization, Operations Research 58 (2010) 470–485.

[34] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, Time Series Analysis: Forecasting and Control, 4th edition, John Wiley & Sons, Inc., Hoboken, New Jersey., 2008.