A successive linear programming algorithm with non-linear time series for the reservoir management problem

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

July 20, 2017

[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 proposes a multi-stage stochastic programming formulation based on affine decision rules for the reservoir management problem. Our approach seeks to find a release schedule that balances flood control and power generation objectives while considering realistic operating conditions as well as variable water head. To deal with the non-convexity introduced by the variable water head, we implement a simple, yet effective, successive linear programming algorithm. We also introduce a novel non-linear inflow representation that captures serial correlation of arbitrary order. We test our method on a small real river system and discuss policy implications. Our results namely show that our method can decrease flood risk compared to historical decisions, albeit at the cost of reduced final storages.

Keywords: Mathematical programming, Stochastic processes, Forecasting, Risk analysis

1 Nomenclature

1.1 Sets and parameters

T	Total number of periods
L	Length of look-ahead period ($L \leqslant T$)
$\mathbb{T}_{t,L}$	Look-ahead period of L periods starting at time t
$\mathbb{L}$	Lead times
$\mathbb{K}^{prod.\,ref.}$	Piecewise linear segments approximating the reference production curve
$\mathbb{K}^{flood\,dmg.}$	Piecewise linear segments approximating the flood damage curve
$\mathbb{K}^{f.val.}$	Piecewise linear segments approximating the storage value function curve
I	Set of plants
I^{evac}	Set of plants with evacuation curves
$I^+(j), I^-(j)$	Set of plants upstream and downstream of reservoir j, respectively
J	Set of reservoirs
$j^+(i)$, $j^-(i)$	Unique reservoir upstream or downstream of plant i, respectively
$\mathcal{H}_i^{ref}$	Reference water head at plant i
$\mathcal{N}_{jt}^{ref}$	Reference water level at reservoir j
λ_{il}	Fraction of the releases from plant i reaching the unique downstream reservoir after l periods
δ_i^{min}, δ_i^{max}	Minimum and maximum water delays for plant i
γ_j	Relative importance of flood damages at reservoir j
ϕ	Relative importance of power generation objectives compared to flood control
α_{jt}	Average proportion of the total inflows entering reservoir j at time t
$\underline{l}_i$, $\overline{l}_i$	Upper and lower bounds on total spillage at plant i
$\underline{r}_i$, $\overline{r}_i$	Upper and lower bounds on releases at plant i
$\underline{f}_i$, $\overline{f}_i$	Upper and lower bounds on total releases/total flow at plant i
Δ_i	Maximum flow variation between periods at plant i
$\underline{s}_j$, $\overline{s}_j$	Upper and lower bounds on storage at reservoir j
l_i^{Δ}, l_i^0	Parameters for the linear approximation of the evacuation curve at plant i
p_{ik}^{Δ}, p_{ik}^0	Parameters for the k^{th} segment of piecewise linear approximation of the reference production function for plant i
n_j^{Δ}, n_j^0	Parameters for the linear approximation of the water level curve for reservoir j
e_{jk}^{Δ}, e_{jk}^0	Parameters for the k^{th} segment of piecewise linear approximation of the flood penalty at reservoir j
s_{jk}^{Δ}, s_{jk}^0	Parameters for the k^{th} segment of piecewise linear approximation of the final water value function at reservoir j
ν_t	Price of 1 MWh during time t

1.2 Decision variables

$\mathcal{R}_{it}$	Releases at plant i during time t
$\mathcal{L}_{it}$	Spills at plant i during time t
$\mathcal{F}_{it}$	Total flow at plant i during time t
$\mathcal{S}_{jt}$, $\bar{\mathcal{S}}_{jt}$	Storage and value of storage at reservoir j at the end of time t
$\mathcal{E}_{jt}$, $\bar{\mathcal{E}}_{jt}$	Floods and flood damages at reservoir j at the end of time t
$\mathcal{H}_{it}$	Relative water head at plant i and time t
$\mathcal{P}_{it}$, $\mathcal{P}_{it}^{total}$	Reference and real production at plant i and time t
$\mathcal{N}_{jt}$	Water level at reservoir j and time t
$\mathcal{Q}_t^{floods}$	Cost associated with floods at time t
$\mathcal{Q}_t^{prod}$	Cost associated with electricity generation and final value of storage at time t

1.3 Random variables

ξ_t	Total inflows over all reservoirs during time t
ξ_{jt}	Inflows at reservoir j during time t
ζ_t	Logged (total) inflows during time t
$\hat{\zeta}_t(l)$	Conditional expectation of ζ_{t+l} at time t for lead time l
$\rho_t(l)$	Forecast error at time t for lead time l
ϱ_t	Residuals at time t
$\mathcal{G}_t$	Information known at time t
Ξ_t	Dynamic uncertainty set given $\mathcal{G}_t$
$\hat{\Xi}_t$	Exterior polyhedral approximation of Ξ_t

2 Introduction

This paper considers the optimal operation of a multi-reservoir hydro-electrical system over a given time horizon subject to uncertainty on inflows. We seek to balance hydro-electrical production and flood control while respecting a host of operating constraints on storages, spills, releases and water balance. Our model explicitly considers water delays as well as variable water head. In addition, we consider inflow persistence through ARIMA time series of *arbitrary* order [1].

As mentioned in [2], this stochastic multi-stage non-convex problem is extremely challenging to solve. The variable water head and the associated non-convexity make it impractical to find a globally optimal solution while the multi-dimensional nature of the problem often requires various simplifications that may lead to policies of limited practical use [3].

To solve this problem, we propose a stochastic program based on affine decision rules (ADR), non-linear ARIMA time series and a successive linear programming (SLP) algorithm. We incorporate our lightweight model in a rolling horizon framework and test it with historical inflows. Our approach builds on previous work on ADR [4, 5, 6, 7, 8], but makes important improvements in terms of modeling, namely by introducing a non-linear time series model together with variable water head. Our model also considerably improves our past empirical results regarding flood reduction [4, 5].

ADR were initially proposed in the water resources literature under a restrictive form with chance-constraints [9]. However, it is only after the work of [10, 11], that they have gained considerable attention for various energy-related problems [4, 6]. Although they do not guarantee optimality in general, these simple parametric policies often lead to good-quality solutions that can be obtained very efficiently.

Unlike competing methods such as traditional stochastic dynamic programming (SDP) [12] and various other extensions such as sampling stochastic dynamic programming (SSDP) [13], ADR do not require discretizing the state-space, which results in important computational gains.

As illustrated in [5], ADR also make it possible to consider long water delays as well as highly persistent inflows without increasing the computational burden as would be the case of algorithms belonging to the family of SDP. Considering these long inflow delays can be essential to avoid overly optimistic solutions in the face of floods [5, 12, 14].

This paper also extends [5] and [6] by introducing a new inflow representation based on a simple non-linear time series model. Although this leads to a non-convex support, the resulting model has nice statistical interpretation, good forecasting ability and can provide lower bounds on key metrics such as the risk of floods.

This is a key advantage of our approach compared with stochastic dual dynamic programming (SDDP), which to the best of our knowledge can only handle linear (P)ARMA models with serially independent residuals [15, 16]. Although the authors of [17] consider a non-linear PAR inflow model similar to the one used in this paper for their SDDP algorithm, this method is only a heuristic that does not provide valid upper and lower statistical bound, as does the original SDDP algorithm. Furthermore, the linear approximation used in [18] retains valid theoretical properties, but does not represent the stochastic process exactly.

By exploiting our non-linear representation along with affine decision rules, we are able to obtain a compact bilinear expression for the objective function. In order to tackle the non-convexity, we implement a simple SLP algorithm. These algorithms have been successfully utilized in hydro-electrical reservoir management problems in the context of stochastic programs based on scenario-trees [19], distributionally robust optimization [20] as well as at the operational level for deterministic and stochastic optimization, namely at Hydro-Québec. These experimental results suggest that this simple algorithm can provide good quality solutions since the objective is bilinear and therefore well approximated by a linear first-order Taylor approximation in a restricted neighborhood.

Considering variable water head is important since it captures the fact that a better production schedule can produce more electricity by using less water than the corresponding optimal plan with fixed head [21]. This is also an important consideration from an operational perspective in various regions such as Québec where cascaded river systems are widespread.

Although recent SDDP extensions can approximate non-convexities through McCormick envelopes and Lagrangian relaxation [22, 23, 24, 25, 26, 27], as well as mixed integer approximations of the cost-to-go [28], these methods are considerably more involved than our simple SLP algorithm. These methods also require longer computing times than our method, which only needs to solve a few small linear programs.

The paper is structured as follows. We present the deterministic reservoir management problem in Section (3). We then discuss our non-linear inflow representation and the associated non-convex uncertainty set in Section (4). Section (5) presents the multi-stage stochastic program based on ADR as well as the SLP algorithm used to solve it approximately. We then present the river system used for our numerical experiments in Section (6). Section (7) discusses results and Section (8) draws concluding comments.

3 Deterministic formulation

3.1 Daily river operations

River operators must determine the average water release that will flow through the turbines of the powerhouse i during day $\tau = t + l$ where $\tau \in \mathbb{T}_{t,L} = \{t, \cdots, t + L - 1\}$. For the level of granularity considered by our model (daily decisions), the different turbines are aggregated and it is assumed that no maintenance occurs so that all turbines are available for power generation. Because of plant specificities and other physical limitations, the total releases are required to reside within some closed interval:

$$\underline{r}_i \leqslant \mathcal{R}_{i\tau} \leqslant \bar{r}_i \qquad \forall i \in I, \tau \in \mathbb{T}_{t,L}. \tag{1}$$

Based on physical characteristics of the sites, a given amount of water can be unproductively spilled:

$$\underline{l}_i \leqslant \mathcal{L}_{i\tau} \leqslant \bar{l}_i \qquad \forall i \in I, \tau \in \mathbb{T}_{t,L}. \tag{2}$$

Some sites $i \in I^{evac} \subset I$ can also evacuate water through a separate spillway. Due to the physical structure of these evacuators, the maximal amount of water that can be physically spilled is bounded by a function of $\mathcal{S}_{j\tau}$, the storage (hm^3) in the unique upstream reservoir $j^-(i)$ at the end of time τ. We approximate these functions with affine functions parametrized by $l_i^\Delta, l_i^0 \in \mathbb{R}$ to obtain:

$$\mathcal{L}_{i\tau} \leqslant l_i^\Delta \mathcal{S}_{j^-(i)\tau} + l_i^0, \qquad \forall i \in I^{evac}, \tau \in \mathbb{T}_{t,L}. \tag{3}$$

We define $\mathcal{F}_{i\tau} = \mathcal{R}_{i\tau} + \mathcal{L}_{i\tau}$ as the total flow (m^3/day) $\forall i \in I, \tau \in \mathbb{T}_{t,L}$, which is simply the sum of unproductive spillage and releases. These quantities are bounded for navigation and recreation purposes, environmental needs as well as to ensure smooth ice formation during the winter months:

$$\underline{f}_i \leqslant \mathcal{F}_{i\tau} \leqslant \bar{f}_i \qquad \forall i \in I, \tau \in \mathbb{T}_{t,L}. \tag{4}$$

To avoid plant damages, we also ensure that the total flow does not change excessively from one day to the next:

$$|\mathcal{F}_{i\tau} - \mathcal{F}_{i\tau-1}| \leqslant \bar{\Delta}_i \qquad \forall i \in I, \tau \in \mathbb{T}_{t,L}, \tag{5}$$

where $\mathcal{F}_{i,t+l}$ have already occurred at time t for any $l \leqslant 0$. We require that the final water storage at the end of time τ at reservoir j lie within fixed bounds:

$$\underline{s}_j \leqslant \mathcal{S}_{j\tau} \leqslant \bar{s}_j \qquad \forall j \in J, \tau \in \mathbb{T}_{t,L}. \tag{6}$$

Given very wet conditions, it is possible that the upper bound constraint (6) are violated. We therefore introduce the non-negative variables $\mathcal{E}_{j\tau} \geqslant 0$ representing floods and modify $\mathcal{S}_{j\tau} \leqslant \bar{s}_j$ to: $\mathcal{S}_{j\tau} - \mathcal{E}_{j\tau} \leqslant \bar{s}_j, \forall j \in J, \tau \in \mathbb{T}_{t,L}$.

We model the monetary value of flood damage ($) as a piecewise linear increasing convex function of floods, which reflects the assumption that greater floods have increasing marginal consequences. We specifically introduce the decision variable $\bar{\mathcal{E}}_{j\tau} \geqslant 0$ and impose:

$$\bar{\mathcal{E}}_{j\tau} \geqslant e_{jk}^\Delta \mathcal{E}_{j\tau} + e_{jk}^0, \forall k \in \mathbb{K}^{flood\,dmg.}, \tag{7}$$

for all j, τ where $\mathbb{K}^{flood\,dmg.} = \{1, \cdots, K^{flood\,dmg.}\}$ and e_{jk}^Δ, e_{jk}^0 define the hyperplanes.

Water balance constraints ensure that the current storage is given as the previous storage plus the net inflow. The net inflow is itself defined as the sum, over all upstream reservoirs, of past releases (m^3/day) weighted by the flow delay coefficients plus the daily inflows $\xi_{j,\tau}$ (m^3/day) during time τ less the releases from all plants during the day (m^3/day):

$$\mathcal{S}_{j,\tau} = \mathcal{S}_{j,\tau-1} + \sum_{i^- \in I^-(j)} \sum_{\bar{l}=\delta_{i^-}^{min}}^{\delta_{i^-}^{max}} \lambda_{i^-\bar{l}} \mathcal{F}_{i^-,\tau-\bar{l}} - \sum_{i^+ \in I^+(j)} \mathcal{F}_{i^+,\tau} + \xi_{j,\tau}, \tag{8}$$

for all $j \in J, \tau \in \mathbb{T}_{t,L}$ where $\delta_{i^-}^{min}$ and $\delta_{i^-}^{max}$ represent the minimum and maximum time taken for a drop of water released from the plant $i^-(j) \equiv i^-$ to reach reservoir j and $\mathcal{S}_{j,t+l}, \mathcal{F}_{i,t+l}$ are already known at time t for any $l \leqslant 0, \forall i, j$. $I^-(j), I^+(j)$ represent the sets of plants upstream and downstream of reservoir j, respectively (see [29, 4, 5] for a similar formulation).

Although easy to model, preliminary numerical experiments suggests that the flow delay constraints can lead to important numerical difficulties for SDDP (and *a fortiori* for SDP). Indeed, the addition of such constraints linking time stages leads to increases in the state-space which make it hard even for SDDP to converge within a limited amount of time.

We make the assumptions that $\xi_{j,\tau} = \alpha_{j,\tau}\xi_\tau$ where $\alpha_{j,\tau}$ is the average proportion of inflows entering reservoir j at time τ and ξ_τ are the inflows over the entire catchment at time τ. This implicitly assumes perfect spatial correlation across sites. A multidimensional ARMA time series model [30, 31] or other more sophisticated models [15] could capture this spatial component. It would also be possible to consider $|J|$ independent time series and then use the Cholesky decomposition of the covariance matrix of the residuals [32, 33]. However, as we detail in the case study in Section 6 our approach appears reasonable in practice.

As is customarily done, we model the final water value function with a piecewise linear increasing concave function [34]. More precisely, we introduce the decision variable $\bar{\mathcal{S}}_{j\tau} \geqslant 0$ and impose :

$$\bar{\mathcal{S}}_{j\tau} \leqslant s^{\Delta}_{jk}\mathcal{S}_{j\tau} + s^{0}_{jk}, \forall k \in \mathbb{K}^{f.val.}, \tag{9}$$

for all j, τ where $\mathbb{K}^{f.val.} = \{1, \cdots, K^{f.val.}\}$ and $s^{\Delta}_{jk}, s^{0}_{jk}$ define the hyperplanes, which are based on the output of other long-term scheduling tools.

The total power produced (MW) during time τ for plant i is taken as the product of the reference production $\mathcal{P}_{i\tau}$ at some fixed reference water head times the relative water head $\mathcal{H}_{i\tau}$ [35, 36, 19]:

$$\mathcal{P}^{total}_{i\tau} = \mathcal{P}_{i\tau}\mathcal{H}_{i\tau}. \qquad \forall i \in I, \tau \in \mathbb{T}_{t,L}. \tag{10}$$

Note that since we consider daily decisions over a monthly horizon, we can realistically assume that generators are used at a fixed level of efficiency. It is therefore not necessary to consider generator level granularity and our production function need not be as detailed as the one in [37, 38].

The relative water head is given by the net water head divided by the reference water head $\mathcal{H}^{ref}_i$:

$$\mathcal{H}_{i\tau} = (\mathcal{N}_{j^-\tau} - \mathcal{N}_{j^+\tau})/\mathcal{H}^{ref}_i \quad \forall i \in I, \tau \in \mathbb{T}_{t,L}. \tag{11}$$

The net water head is simply the difference between forebay $\mathcal{N}_{j^+\tau}$ and tailrace $\mathcal{N}_{j^-\tau}$ elevations during time τ, where $j^+ \equiv j^+(i)$ and $j^- \equiv j^-(i)$ respectively represent the reservoir upstream and downstream of plant i.

The forebay and tailrace elevation (m) at a given plant i are concave increasing functions of the average storage in the upstream $j^-(i) \equiv j^-$ and downstream reservoirs $j^+(i) \equiv j^+$ during time τ respectively. These functions are sufficiently well approximated by affine functions of the water storage in the associated reservoir:

$$\mathcal{N}_{j^-\tau} = n^0_{j^-} + n^{\Delta}_{j^-}\mathcal{S}^{avg}_{j^-\tau} \qquad \forall j \in J, \tau \in \mathbb{T}_{t,L} \tag{12a}$$

$$\mathcal{N}_{j^+\tau} = n^0_{j^+} + n^{\Delta}_{j^+}\mathcal{S}^{avg}_{j^+\tau} \qquad \forall j \in J, \tau \in \mathbb{T}_{t,L}, \tag{12b}$$

where $n^0_j, n^{\Delta}_j \in \mathbb{R}, \forall j$ and $\mathcal{S}^{avg}_{j\tau} = (\mathcal{S}_{j\tau} + \mathcal{S}_{j\tau+1})/2$ is the average water storage in j during time τ. For plants where the downstream reservoir is sufficiently far, the tailrace water level is taken as a constant $\mathcal{N}_{j^-} = \mathcal{N}^{ref}_{j^+}$.

Our approach shares similarities with the one of [37], who assume that the difference in reservoir levels from on time period to the next is a linear function of the net inflow into that reservoir and where the tailrace elevation is in part influenced by the forebay elevation of the downstream reservoir.

The reference production itself depends on the total flows (m^3/day). For every i and τ, we use the following approximation for the reference production function by using $K^{prod.\,ref.} \in \mathbb{N}$ hyperplanes:

$$\mathcal{P}_{i\tau} \leqslant p^{\Delta}_{ik}\mathcal{F}_{i\tau} + p^{0}_{ik}, \forall k \in \mathbb{K}^{prod.\,ref.}, \tag{13}$$

where $\mathbb{K}^{prod.\,ref.} = \{1, \cdots, K^{prod.\,ref.}\}$ and $p^{\Delta}_{ik}, p^{0}_{ik} \in \mathbb{R}$. Each piecewise function is increasing only on a subset of its domain to reflect negative tailrace effects caused by excessive flows. Indeed, for each plant, we have $p^{\Delta}_{ik} > 0, k \in \mathbb{K}\backslash K^{prod.\,ref.}$ and $p^{\Delta}_{iK^{prod.\,ref.}} < 0$. Section (6) provides concrete examples and additional details.

3.2 Biobjective problem

We wish to find a production plan that balances some systemic measure of flood occurrence while maximizing electricity production and avoiding completely emptying the reservoirs at the end of the horizon.

Based on previous work [4], we let $\gamma_j \geqslant 0$ reflect the relative importance of floods at reservoir j where $\sum_{j\in J}\gamma_j = 1$. For our problem instance, these weights are defined naturally since some reservoirs are much more important than others in terms of capacity and flood control. In general, decision makers might use structured weight elicitation schemes such as [39]. We then consider $\mathcal{Q}_\tau^{floods}$, the weighted flood damages over the entire catchment at time $\tau \in \mathbb{T}_{t,L}$ ($), which is given by the following simple relationship:

$$\mathcal{Q}_\tau^{floods} = \sum_{j\in J} \gamma_j \bar{\mathcal{E}}_{j\tau}. \tag{14}$$

We evaluate the hydro-electric productive benefits associated to a reservoir release schedule at time $\tau \in \mathbb{T}_{t,L}$ by considering the (negative) total hydroelectric production value and the value of the final storage of reservoirs ($):

$$\mathcal{Q}_\tau^{prod} = \begin{cases} -\sum\limits_{i\in I^{prod}} \mathcal{P}_{i,\tau}^{total,}\nu_\tau - \sum\limits_{j\in J}\bar{\mathcal{S}}_{j\tau} & \tau = t+L-1 \\ -\sum\limits_{i\in I^{prod}} \mathcal{P}_{i,\tau}^{total}\nu_\tau & \text{else,} \end{cases} \tag{15}$$

where ν_τ is the average price of 1 MWh per day and $\bar{\mathcal{S}}_{j\tau}$ indicates the final value of this water storage. Electricity generation is assumed constant throughout each period τ. We also assume that electricity prices are known constants, which makes sense in the regulated context of Hydro-Québec. Finally, at the beginning of time $t \in \mathbb{T}$, we seek to minimize

$$\sum_{\tau=t}^{t+L-1} \left((1-\varphi)\mathcal{Q}_\tau^{floods} + \varphi\mathcal{Q}_\tau^{prod}\right), \tag{16}$$

for some $\varphi \in [0,1]$ reflecting the relative importance of the aggregate production measure $\mathcal{Q}_\tau^{prod}$ compared to the aggregated flood measure $\mathcal{Q}_\tau^{floods}$. Decision makers may exploit the parametric nature of the objective in order to explore the solution space. While this does not formally allow us to explore the Pareto-front, due to the non-convexity of the problem, varying the φ parameter can help gain some insights on the model. It would also be possible to consider an alternative hierarchical scheme where floods are first minimized and energy is then maximized subject to additional constraints on floods.

4 Incorporating uncertainty

We want to model $\{\xi_t\}_{t\in\mathbb{Z}}$, the discrete time stochastic process representing total inflows. At each time $t \in \mathbb{T} = \{1,\cdots,T\}$ for the lead times $\mathbb{L} = \{0,\cdots,L-1\}$ (future uncertain horizon) , we assume that the $\xi \equiv \xi_{[t,t+L-1]}$ lie within the following set with probability 1:

$$\Xi_t = \left\{ \xi \in \mathbb{R}^L \;\middle|\; \begin{array}{ll} \exists \zeta, \varrho \in \mathbb{R}^L & \text{(17a)} \\ \xi_{t+l} = e^{\zeta_{t+l}}, \forall l \in \mathbb{L} & \text{(17b)} \\ \zeta = V_t\varrho + v_t & \text{(17c)} \\ W\varrho \leqslant w, & \text{(17d)} \end{array} \right\}$$

where $\zeta \equiv \zeta_{[t,t+L-1]} = (\zeta_t,\cdots,\zeta_{t+L-1})^\top$ and $\varrho \equiv \varrho_{[t,t+L-1]}$ represent random variables for the next future L days. We fix $V_t \in \mathbb{R}^{L\times L}$, $W \in \mathbb{R}^{c\times L}$, $v_t \in \mathbb{R}^L$ and $w \in \mathbb{R}^c$ for some $c \in \mathbb{N}$. For any $l \in \mathbb{L}$, $v_{t,l}$ is the forecast of ζ_{t+l} given information up to time t and $\zeta_{t+l} - v_{t,l} = V_{t,l}^\top\varrho$ is the forecast error.

Representation (17) essentially states that the logarithms of inflows follow some ARIMA process, whose structure is codified by the matrix V_t and the vector v_t, and where the residuals ϱ are random variables taking values in the polyhedral set $P = \{\varrho \in \mathbb{R}^L : W\varrho \leqslant w\}$ (see the appendix for more details).

This representation captures commonly used inflow representations as a particular case. For instance, considering large P and assuming $\varrho_{[t,t+L-1]}$ follow a multivariate normal distribution implies that the inflows are correlated and approximately follow the popular log-normal distribution, often used in reservoir management problems [40, 41, 42, 43, 44].

The exponential function used in equation (17b) also possesses various intuitive and statistically appealing properties. Indeed, the function ensures non-negativeness of inflows and belongs to the family of Box-Cox transforms commonly used to obtain a more adequate fit with a theoretical model [45].

An important advantage of ARMA inflow representations compared with periodic PARMA models commonly used in mid-term scheduling (see namely [44, 26, 43, 15]) is their parsimony. For models with monthly decisions, this might be less of a concern. However, for *daily* decisions as is the case in this paper, a PAR(p) model could require calibrating up to $365 \times p$ parameters. This is not merely a theoretical consideration, as considering such a large number of parameters for a limited data set like the one we use will likely lead to over-fitting issues and poor out-of-sample prediction performance. Nonetheless, we could directly use a PARMA(p) model calibrated on the logarithm of inflows within the framework we propose here.

Alternative inflow representations previously used with ADR include [8] that considers the convex hull of a set of scenarios or [4] that uses simple box sets, which disregard the serial correlation. Contrary to these approaches, our model exploits the structure of the time series model to express inflows as a function of uncorrelated random variables. Using Markov's inequality can therefore provide joint probabilistic bounds (see [5] as well as the appendix). Moreover, considering the convex hull of a set of scenarios requires sampling S points $(\check{\xi}^1, \cdots, \check{\xi}^S)^\top$ from Ξ_t and then adding the constraints $\mathcal{R}^0_{it} + \sum_{t'=t}^{t+L-1} \mathcal{R}^{t'}_{it} \check{\xi}^s_{t'} \leqslant \bar{r}, \forall s = 1, \cdots, S$. We experimented with this approach and found that while it yielded results similar to those of our exterior approximation, the model was much longer to solve when S was large.

5 Stochastic multi-stage formulation

We consider a dynamic setting where the inflows are progressively revealed as time unfolds over the horizon of T days. At each time $t \in \mathbb{T}$, the decision maker fixes a sequence of policies to be implemented at future times $\tau = t, \cdots, t+L-1$ under the assumptions that the future inflows $\xi_{[t,t+L-1]}$ belong to Ξ_t. Decisions must be non-anticipative in the sense that each $\mathcal{X}_\tau$ must only be a function of the past random variables $\xi_{[\tau-1]}$.

5.1 Affine decision rules

Solving the true stochastic multistage problem to optimality poses important computational difficulties. We therefore limit ourselves to affine policies, which are suboptimal but provide an upper bound and might still adequately capture the general structure of optimal future decisions. These functions take the following form at time $\tau = t + l$ for $t \in \mathbb{T}$ and $l \in \mathbb{L}$:

$$\mathcal{X}_{k\tau}(\xi) = \mathcal{X}^0_{k\tau} + \sum_{\tau'=t}^{t+L-1} \mathcal{X}^{\tau'}_{k\tau} \xi_{\tau'}, \tag{18}$$

where $k \in \{n_{\tau-1}+1, \cdots, n_{\tau-1}+n_\tau\}$ and $n_\tau \in \mathbb{N}$ represents the number of decisions at time τ. $\mathcal{X}^0_{k\tau}, \mathcal{X}^{\tau'}_{k\tau} \in \mathbb{R}, \forall k, \tau$ become the decision variables in our deterministic equivalent formulation. Non-anticipativity is accounted for by imposing $\mathcal{X}^{\tau'}_{k\tau} = 0, \forall \tau' \geqslant \tau$.

As mentioned in [4, 46], solving the stochastic problem with ADR requires replacing each decision variable in the deterministic formulation by its associated affine decision in the stochastic model. To ensure that the inequalities hold with probability 1, we solve an optimization problem. For instance, constraint (1) is respected with probability 1 if $\max_{\xi \in \Xi_t} \mathcal{R}^0_{it} + \sum_{t'=t}^{t+L-1} \mathcal{R}^{t'}_{it} \xi_{t'} \leqslant \bar{r}$, where $\mathcal{R}^0_{it}$ and $\mathcal{R}^{t'}_{it}$ correspond to $\mathcal{X}^0_{k\tau}, \mathcal{X}^{\tau'}_{k\tau}$ for some specific indices i, k. However, this optimization problem is intractable given our choice of Ξ_t, since this set is not generally convex for an arbitrary $t \in \mathbb{T}$ (see the appendix). Fortunately, it is possible to obtain the following exterior polyhedral conservative approximation :

$$\hat{\Xi}_t = \bigtimes_{l=1}^{L} \left[e^{\min_{W\varrho \leqslant w} V_l^\top \varrho + v_{t,l}}, e^{\max_{W\varrho \leqslant w} V_l^\top \varrho + v_{t,l}} \right], \tag{19}$$

where the interval bounds can be conveniently computed by minimizing or maximizing $\{V_l^\top \varrho : W\varrho \leqslant w\}$ once in a static fashion. The parameter $v_{t,l}$ depends on past observed inflows and W, w determine the joint probabilistic guarantees. We can check that $\Xi_t \subseteq \hat{\Xi}_t$, where $\hat{\Xi}_t$ is simply the Cartesian product of intervals that can be represented with only $2L$ hyperplanes. We then use strong linear programming duality to write the deterministic equivalent as a large linear program (see [4] for more details). Since we consider the conservative exterior approximation $\hat{\Xi}_t$, solving our deterministic equivalent yields an upper bound on the true problem.

More specifically, the value of the two following LP's is the same:

$$\begin{array}{lll|lll} \max\limits_{\xi} & \sum\limits_{l=0}^{L} \mathcal{R}_{it}^{t+l} \xi_{t+l} & & \min\limits_{\pi^-,\pi^+} & \sum\limits_{l=1}^{L} \left(\pi_l^+ \overline{\vartheta}_{t,l} - \pi_l^- \underline{\vartheta}_{t,l}\right) & \\ \text{s. t.} & \xi_{t+l} \leqslant \overline{\vartheta_{t,l}}, \forall l & & \text{s. t.} & \pi_l^+ - \pi_l^- = \mathcal{R}_{it}^{t+l}, \forall l & \\ & -\xi_{t'} \leqslant -\underline{\vartheta}_{t,l} \forall l & & & \pi_l^-, \pi_l^+ \geqslant 0, \forall l & \end{array}$$

where $\underline{\vartheta}_{t,l} = \exp(\min_{W\varrho \leqslant w} V_l^\top \varrho + v_{t,l})$ and $\overline{\vartheta}_{t,l} = \exp(\max_{W\varrho \leqslant w} V_l^\top \varrho + v_{t,l})$. We therefore only need to add the π^+, π^- variables as well as the linear constraints $\mathcal{R}_{it}^0 + \sum_{l=1}^{L} \left(\pi_l^+ \overline{\vartheta}_{t,l} - \pi_l^- \underline{\vartheta}_{t,l}\right) \leqslant \bar{r}$ together with $\pi_l^+ - \pi_l^- = \mathcal{R}_{it}^{t+l'}, \forall l'$ and $\pi_l^-, \pi_l^+ \geqslant 0, \forall l$ to ensure that the original constraint holds for all $\xi \in \hat{\Xi}_t \supset \Xi_t$. We then optimize over $\mathcal{R}_{it}^0, \mathcal{R}_{it}^{t'}, \pi^+$and π^-.

We point out that while this box uncertainty set may seem overly conservative and simple, it still provides joint probabilistic bounds that would be impossible to obtain with Markov's inequality if we simply considered a box of the form, say, $\xi_t \in [\mu_t - \sigma_t \nu, \mu_t + \sigma_t \nu], \forall t$ for some $\nu > 0$, where μ_t, σ_t are the mean and standard deviation of inflows at time t, respectively, as is used in [4]. Even if we lose some precision with our proposed conservative convexification, we still ensure that the uncertainty only takes plausible values given the observed history, since the inflows are expressed as a function of the independent ϱ (see [5] as well as the appendix for more details on these joint guarantees).

5.2 Objective function and composite risk

We define the (composite) risk of the production plan at the beginning of time t over the next T days as:

$$\mathrm{E}\left[\sum_{l=0}^{T-1} (1-\varphi) \mathcal{Q}_{t+l}^{floods}(\xi) + \varphi \mathcal{Q}_{t+l}^{prod}(\xi) | \mathcal{G}_{t-1} \right], \tag{20}$$

where $\mathcal{G}_t = \sigma(\varrho_\tau : \tau \leqslant t)$ is the σ-algebra generated by the past ϱ_τ and that we can simply interpret as the information known at time t. By linearity of the conditional expectation, it follows that this composite risk is simply a convex combination of the flood risk and the production risk. The appendix details the analytical expression of these terms.

5.3 Using successive linear programming to approximate the true problem

Due to (10), (20) contains a bilinear term (see the appendix for details). In order to deal with this non-convexity, we consider a commonly used first order Taylor approximation, which is detailed in the appendix of the technical report, as well as a SLP algorithm based on the general trust-region algorithm of [47].

Our algorithm begins by considering the linear problem of minimizing flood damages, *i.e.*, we set $\varphi = 0$ in (16), without any trust region constraints. After obtaining the initial solution $\hat{\mathcal{X}}_\nu$ at iteration $\nu = 0$, we add norm infinity (box) constraints $\left\| \hat{\mathcal{X}} - \mathcal{X}_\nu \right\|_\infty \leqslant \epsilon_\nu$ for some $\epsilon_\nu > 0$ since the linearized approximation of the bilinear term is only valid within a limited region around the incumbent solution.

We then reset $\varphi > 0$ to its prescribed value and resolve the modified linearized problem iteratively. At each iteration ν, we update the value of ϵ_ν according to the basic trust region algorithm of [47] and consider a greedy version where we only move to a new feasible point if the objective value increases. We stop the algorithm if the value of ϵ_ν falls below a fixed threshold, if the incumbent objective shows no improvement with respect to the previous iteration and the norm constraints are not binding, or if we reach a predetermined number of iterations.

The SLP algorithm displayed very rapid convergence and after only 4 iterations, the flood and production risk remained roughly constant. We obtained similar results with different stopping criteria, period of the year and weights φ in (20).

Rather than using this SLP algorithm, it is possible to consider McCormick envelopes to convexity the bilinear term and progressively refine the representation with integer programming as done in [23, 27]. This would allow us to obtain a lower bound on the minimization problem. However, this procedure will likely negatively affect the tractability and very low computing times of our approach, which are paramount to its practical usage for Hydro-Québec.

5.4 Rolling horizon simulation

We embed the SLP algorithm within a broader rolling horizon simulation that reflects the real behavior of river operators (see Figure (1)). At each time $t \in \{1, \cdots, 30\}$, we compute the forecasts $\mathrm{E}\left[\xi_{t+l}|\mathcal{G}_t\right]$ for the next 30 days. We then optimize the deterministic equivalent, fix the first period's decisions, simulate daily inflows for this period, compute final storages and resulting violations, update the initial conditions of our model and finally move on to the next time step. We repeat the preceding steps for a set of $S \in \mathbb{N}$ simulated or historical inflow trajectories. Hence, for each of the S scenarios, we may find different optimal $\mathcal{X}^0_{\tau,k}, \mathcal{X}^{\tau'}_{\tau,k}$ for the affine functions (18). In the next sections, we only report results for the first 30 periods of the horizon, which is the period of interest.

Although each model considers a horizon of 30 days, we only "robustify" the constraints for the next $L = 7$ periods because this speeds up computations while still providing good solutions. For instance, we require that the following constraints on releases hold: $\mathcal{R}_{i,t+l}(\xi) \leqslant \bar{r}, \forall \xi \in \hat{\Xi}_t$ for $l = 0, \cdots, L-1$, but only require that the "nominal" constraints $\mathcal{R}^0_{i,t+l} \leqslant \bar{r}$, hold for $l = L+1, \cdots, T$.

6 The Gatineau river system

We apply our framework to the Gatineau river in Western Québec. This hydro-electrical complex is part of the larger Outaouais catchment and is managed by Hydro-Québec. It is composed of 5 reservoirs and 3 run-of-the river plants.

The first 2 head reservoirs, Baskatong and Cabonga have storage capacity of 1563 hm^3 and 3049 hm^3, respectively, while the remaining downstream reservoirs Paugan, Chelsea and Rapides-Farmers have virtually none.

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 flood risk during the freshet as well as baseflow during summer. Furthermore, it is located upstream of the town of Maniwaki, which has witnessed 5 significant floods in the past, including the recent and important flood of spring 2017, which has affected a large portion of Southern Québec. We therefore impose extremely tight operating constraints on water flows at this segment.

Like other flood-prone river systems [48, 49, 50], reservoirs on the Gatineau were created primarily to mitigate the short term negative effects of extreme hydrological conditions. In spring 2017 for instance, operators were able to limit floods downstream at Maniwaki and the greater Montréal region by maintaining Baskatong at high levels. In this context, hydro-electrical objectives are secondary. From this operational perspective, it also makes sense to consider daily discharge and spills as this is the level of granularity considered by river operators during the freshet and other periods that need the most active management.

The three run-of-the-river plants Paugan, Chelsea and Rapides-Farmers have relatively small installed capacity. As illustrated in Figure (2), up to some critical threshold, increasing flow will increase the reference production. On the other hand, excessive releases and spills above this value will have negative effect on

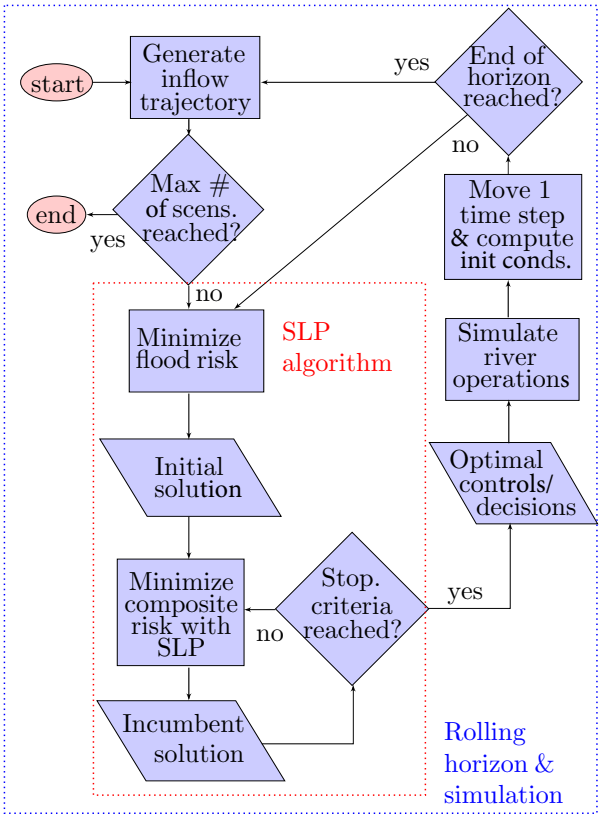

Figure 1: Flow chart for the rolling horizon simulation algorithm.

water head, which is captured with an additional linear segment of negative slope. This representation was proposed by Hydro-Québec, but it could be interesting to evaluate alternative models that capture tailrace effects such as [37, 51].

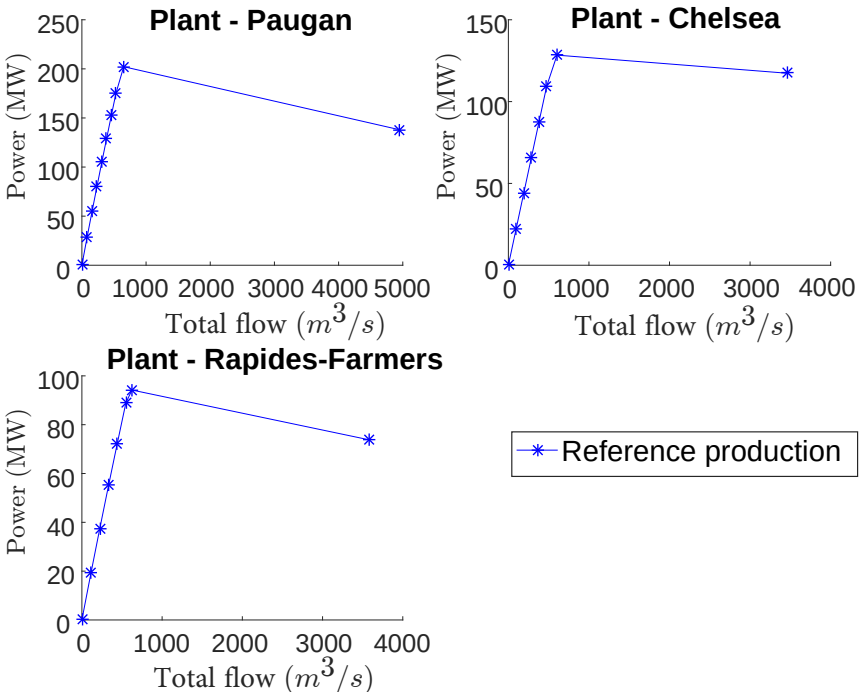

Figure 2: Production functions at reference water head

7 Numerical experiments

7.1 Daily inflows

We consider 12 years of historical daily inflows provided by Hydro-Québec. We calibrated our model on 6 years of in-sample data 1999-2004, but we also tested our framework on 6 additional out-of-sample years 2008-2013, which are on average wetter than the in-sample years.

As mentioned previously, we assume that each reservoir receives a fixed proportion of the total inflows at each time. This approach is reasonable, because the most important flood dynamics are directly related to the inflows at Baskatong, the largest reservoir accounting for the majority of inflows, and the town of Maniwaki. Since the inflows at these two sites are highly correlated with total inflows, our approximation is sufficient to obtain good performances. It might be beneficial to consider multivariate ARMA models [31], but this will significantly complicate the identification and calibration of the time series while also increasing the size of the resulting deterministic equivalent. We therefore leave this to future work.

7.2 Comparative study of flood reduction with high initial storage

We highlight the performance of the non-linear ARMA representation (17) with ADR by comparing it with 2 alternative ADR models based on simpler inflow representations as well as a convex SDDP benchmark. For sake of simplicity, we only consider the linear flood minimization problem, which reflects the real behavior of operators in given periods of the year like the freshet.

7.2.1 ADR models

The three ADR models discussed in this section all consider functions of the form (18), but exploit different uncertainty representations.

The first naive representation is a static uncertainty set that consists of intervals around the historical sample mean and does not explicitly consider serial correlation (see [4]). Contrary to the non-linear model proposed in this paper, this naive representation considers static forecasts with fixed uncertainty sets. How-

ever, it shares the same box structure as (19) and hence the dual and resulting deterministic equivalent are structurally similar to the one in this paper.

The second one considers the linear ARMA(1,1) model described in [5], which although conceptually similar to (17), does not take into account the non-linear log transformation and leads to a deterministic equivalent with relatively different structure.

The third model is based on the non-linear ARMA model (17) where $\xi_t = e^{\zeta_t + \hat{\mu}_t}$, $\hat{\mu}_t$ is the sample log average and the AIC criterion suggests that the ζ_t follow an AR(4) model.

7.2.2 SDDP benchmark

We also implement a simple version of the SDDP algorithm described in [18]. We avoid aggregating reservoirs or past flows. Although this may slow down convergence (see [25]), aggregation leads to extremely bad solutions when simulated on real inflows. We also implemented a procedure to identify and remove redundant cuts since CPLEX seemed to stall when a large number of similar cuts were incorporated (see [25]). Since to the best of our knowledge SDDP can only exactly consider serially independent linear time series, we consider the same affine ARMA(1,1) inflow model as in the second model based on ADR [5].

We consider the sample average approximation to the true problem and independently sample $\{\varrho'_{t,n}\}_{n=1}^{N}$ with $N = 10$ from a normal distribution with moments matching the empirical values found from our calibrated time series model. Hence our recombining tree implicitly consider 10^{30} possible scenarios. We always use $M = 2$ of these scenarios without replacement in the forward phase. We run the SDDP algorithm for 200 forward and backward iterations before performing the simulation/re-optimization on the real scenarios. We only consider 200 iterations, because this already takes more than an hour to compute and additional iterations does not necessarily translate into better simulated performance.

7.2.3 Experimental method

We focus on hard initial conditions with high storage to provide a better basis for comparison since average levels lead to very limited violations over all models. This also allows us to determine if Hydro-Québec can maintain higher storages at the beginning of the freshet period.

Results were obtained by considering a 30-day period starting at the beginning of the spring freshet for each of the 12 historical inflow scenarios under the same hard initial storages. We set $\varphi = 0$ in the objective (16). All problems were solved using AMPL with solver CPLEX 12.5 on computers with 16.0 GB RAM and i7 CPU's @ 3.4 GHz.

All models based on ADR are run in a rolling horizon framework, where each model solved considers the next 30 days, but with uncertainty only on the next $L = 7$ days. Only the constant current decisions are implemented.

The SDDP algorithm is re-optimized based on the initially computed cost-to-go piecewise convex approximation since recomputing the cost-to-go functions in a rolling horizon would be much too heavy in terms of computing times.

	Bask. storage viol. (hm^3)	Mani. flow viol. (m^3/s)	Agg. viol. (m^3/s)	Energy (GWh)	Final Volume ($10^3\,hm^3$)	Total Time (s)
Non-lin. ARMA	0.000	0.000	0.000	140.4	22.22	1288
Linear ARMA	1605	27.63	166.3	147.3	40.33	1888
Naive forecast	1768	28.08	180.9	136.4	43.56	1200
Linear SDDP	0.000	118.1	118.1	143.9	38.41	3704

Table 1: Impact of different uncertainty representations.

7.2.4 Results analysis

Results presented in Table (1) represent the sum over the 12 historical scenarios. The aggregated violations represents the sum of storage and flow violations, when considering the conversion of hm^3/day to m^3/s. The total time reported considers the entire rolling horizon procedure as well as the initial cost-to-go approximation for SDDP. We report the energy and final volumes for information purposes only since the optimization model is indifferent to these factors.

The absence of violations for the non-linear AR(4) model for all 12 scenarios in the case of these extremely challenging stress tests is very encouraging. Moreover, these flood reduction objectives are achieved without sacrificing much energy generation compared to the best model. These results suggest that Hydro-Québec can maintain higher storages without violating operating constraints nor social obligations. However, Table (1) reveals that maintaining higher storage is of questionable value as a significant amount of water needs to be evacuated out of the system to avoid floods.

It is likely that the superiority of the non-linear inflow representation comes from the fact that the linear ARMA model considered by the second ADR model and SDDP is a distortion of the true inflow model. For instance, it implicitly assumes that inflows can be negative. Using truncation to avoid this issue will bias and skew the empirical distribution. In addition, as can be seen from the higher variance residuals, the model is less skillful than the non-linear AR(4) model at predicting future values. Finally, the residuals from the ARMA(1,1) model display problematic signs of heteroscedasticity, which is not the case for the AR(4) model. As for the naive forecast, it always considers the same nominal static forecasts (the unconditional mean).

Given the same ARMA(1,1) inflow representation, the superiority of SDDP compared with the ADR model can be attributed to the fact that SDDP converges to a general flood minimizing policy while we only obtain an optimal *affine* policy with ADR.

Nonetheless, the results of this paper and past works [4, 5] support the idea that it is preferable to consider more accurate inflow representations than policy structures. Hence, the ability to consider non-linear inflows with better forecasting skill proves an important advantage for our non-linear ADR model compared with other ADR and SDDP.

It is possible that using linear schemes [17, 18] to approximate the true non-linear inflow effects could also provide results of similar quality with SDDP. We suggest performing a more detailed comparison between the 2 methods in future work. However, we believe that ADR offer important computational advantages over SDDP. ADR do not require sampling and will obtain an optimal solution in polynomial time since they exploit linear programming. On the other hand, SDDP is only guaranteed to converge within a finite, but possibly very large number of steps [16]. This consideration is not merely theoretical as we found that stopping the algorithm prematurely led to myopic solutions of questionable value. Furthermore, our models based on ADR do not display the numerical issues experienced by SDDP such as stalling due to numerous cuts. Since our ADR models can be solved extremely rapidly, they can be re-optimized efficiently given various initial conditions, which are tantamount for short term flood dynamics. While SDDP can be partially re-optimized based on the previously computed cost-to-to approximations, these cuts may have limited value if they were only computed based on average value of states, which have a greater likelihood of being visited than the extreme flood-generating scenarios. From a short-term operational perspective, we believe these computational aspects to be crucial.

7.3 Model performance under realistic historical conditions compared with historical decisions

7.3.1 Flood reduction

Next, we consider the very wet historical scenario of fall 2003. Figure (3) illustrates that these meteorological conditions led to very high storages at the Baskatong and Cabonga reservoirs around early November and in turn to upper flow bounds violations of 55 m^3/s at the downstream town of Maniwaki on 3 occasions during the period November 14 to December 13. We consider November 14 as the starting period, which is the day the Baskatong storage reached a level very close to maximal normal operating limits. We therefore naturally fix $\varphi = 0$ and only focus on minimizing flood risk for the catchment. We fixed the initial storages at their historical levels and used the realized inflows over this 30-day period. The time series 'Average' represents

the daily average storage at Cabonga, daily average storage at Baskatong and the daily average flow at Maniwaki, respectively. These averages are computed over the 5 years 1999-2003 using the historical data. The time series 'Average', 'Historical' and 'Model' correspond to the left axis (storage for Baskatong and Cabonga, flow for Maniwaki), while the right axis corresponds to the 'Inflows' time series. Upper and lower bounds are indicated by solid black lines. All models are run in a rolling horizon framework considering a look-ahead period of $L = 7$ days each time.

Figure (3) reveals that our model out-performed the historical decisions since the production plan leads to solutions that respect all operating constraints. We are able to respect flow bounds at Maniwaki by emptying the Baskatong reservoir faster and better anticipating the very large inflows at times 5 and 6. Our release schedule maintains approximately the same final storages in Baskatong as the historical decisions, which illustrates the importance of properly timing the decisions. However, as in Section (7.2), a significant amount of water is released from Cabonga out of the system. This result suggests that under extreme scenarios, it is very hard to simultaneously reduce floods and maintain the same level of storage.

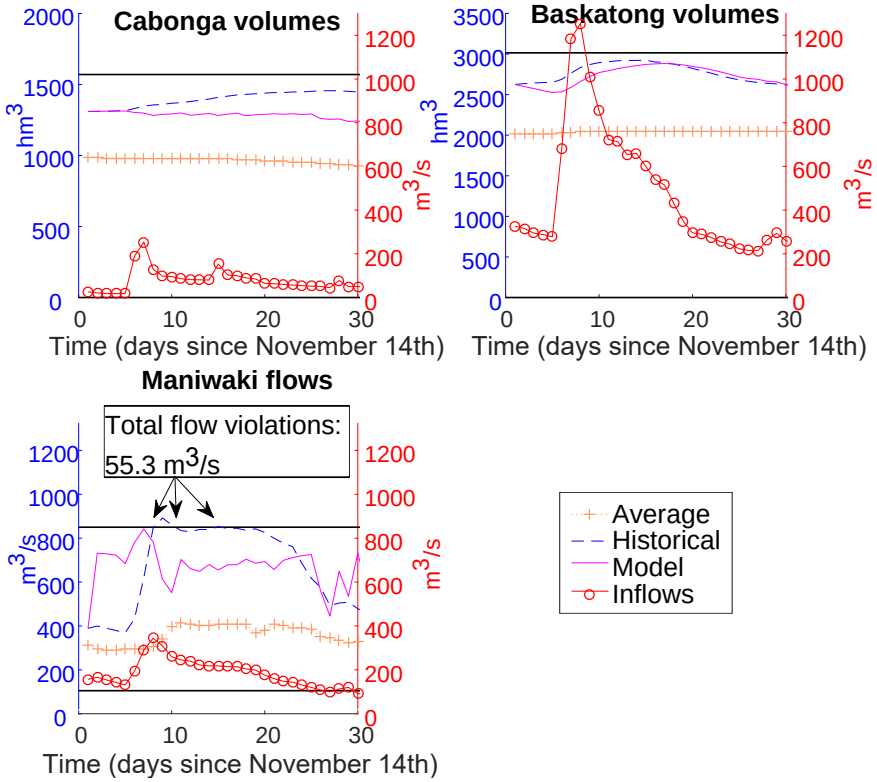

Figure 3: Model results on hard November 2003 scenario with historical inflows and initial storage. Only the two head reservoirs and the town of Maniwaki are presented.

7.3.2 Energy generation

To validate the utility of our approach for electricity generation purposes, we tested our model on the years 1999-2003 for the summer period, which corresponds to the days 200-229. We only considered these years, because the data for the historical production plan for the years 2008-2013 was not available. We considered this period, since it represents the moment where the reservoirs are relatively full, there are no pressing flood risks and it is possible to exploit the variable water head to derive operating efficiencies. Table (2) reveals that our model yields an average energy increase of 0.3 % over a 30-day horizon relative to the historical production plan. These figures are obtained by using the same amount of water as the historical plan and avoiding floods. The values in this table where obtained by setting $\varphi = 1$ and focusing on energy production.

Table 3 suggests that this increase in production is the result of exploiting the variable water head

actively. Indeed, it exploits a larger range of head values than the historical production plan at all plants. We also notice that on average, our model proposes compensating the lower head at Chelsea with a higher head at Rapides-Farmers.

Year	Paugan	Chelsea	Rapides-Farmers	Total
1999	0.5	-6.5	10	0.2
2000	0.4	-5.4	8.6	0.4
2001	0.5	-5.2	8.0	0.4
2002	0.4	-3.5	5.4	0.2
2003	0.1	-3.5	5.1	0.0
Average	0.4	-4.8	7.4	0.3

Table 2: Energy generation increase (%) with model water head.

	Paugan		Chelsea		Rapides-Farmers	
	Non-lin. ARMA	Hist.	Non-lin. ARMA	Hist.	Non-lin. ARMA	Hist.
Min.	0.99	0.99	0.94	0.97	0.92	0.92
Average	1.00	1.00	0.98	1.03	1.02	0.95
Max.	1.00	1.00	1.06	1.05	1.08	1.03
Range	0.01	0.01	0.12	0.08	0.16	0.11

Table 3: Comparison of proposed and historical relative head using the same values as Table (2).

8 Conclusion

This paper presents encouraging evidence that stochastic programming models based on ADR can improve short-term river management operations. Our framework based on non-linear time series and a SLP algorithm can successfully consider important phenomena such as water delays, variable water head and inflow persistence of arbitrary order.

Our model yields sizable improvements when compared with historical decisions on both flood reduction and energy generation. We show that it is possible to find feasible release schedules with storage considerably larger than average historical levels, even with high inflows. This suggests it is possible to revise the drawdown-refill cycle currently used in practice to maintain higher storage at the beginning of the freshet.

In the context of ADR, the non-linear inflow representation appreciably improves flood management compared with past works [5, 4]. We also show that our method can provide results competitive with the popular SDDP algorithm in less computing time.

Finally, the simple convex combination of flood control and production objectives we consider can allow river operators to easily explore the solution space. Although we could not give additional results due to lack of space, it is interesting to assess possible trade-offs using different inflow simulations. Since our model takes only a few (<5) seconds to completely re-optimize, it is highly feasible to perform such sensitivity analysis in a realistic setting.

Acknowledgment

The authors would like to thank Grégory Émiel, Louis Delorme, Pierre-Marc Rondeau, Sara Séguin, Jasson Pina and Pierre-Luc Carpentier. This research was supported by NSERC/Hydro-Québec through the Industrial Research Chair on the Stochastic Optimization of Electricity Generation and grant 386416-2010. We thank seven anonymous reviewers for their valuable comments.

9 Appendix

10 Details on Ξ_t

10.1 Time series model linking ζ and ϱ

At each time $t \in \mathbb{Z}$, we assume there exists some finite constant $\bar{v}_t$ such that $\zeta_t = \ln \xi_t - \bar{v}_t$ are zero mean random variables. Given S years of data where $\xi_{t,i}$ represents the inflows on the t^{th} period of the i^{th} series, we can namely consider $\bar{v}_t = \sum_{i=1}^{S} \frac{1}{S} \ln \xi_{t,i}, \forall t$, the sample daily logged mean, which is also used in [41]. The ζ_t are also assumed to satisfy the equation:

$$\phi(B)\zeta_t = \theta(B)\varrho_t, \tag{21}$$

where B represents the backshift operator acting on time indices such that $B^p\zeta_t = \zeta_{t-p}$ for all $t, p \in \mathbb{Z}$ [1, 52], $\phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\theta(B) = 1 + \sum_{i=1}^{q} \theta_i B^i$ with $\phi_i, \theta_i \in \mathbb{R}$ for $p, q \in \mathbb{N}$.

We also suppose $\{\varrho_t\}$ are second order stationary zero-mean i.i.d. random variables [1] and that we can find $\psi(B) = \sum_{i=0}^{\infty} \psi_i B^i$ where $\phi(B)\psi(B) = \theta(B)$ such that $\sum_{i=0}^{\infty} |\psi_i| < \infty$. Under suitable conditions on $\phi(B)$ and $\theta(B)$, the representation $\zeta_t = \psi(B)\varrho_t$ holds and is essentially unique [52].

We then exploit the following linear decomposition:

$$\hat{\zeta}_t(l) \equiv \mathrm{E}\left[\zeta_{t+l}|\mathcal{G}_t\right] = \sum_{i=0}^{l-1} \psi_i \varrho_{t+l-i} \tag{22}$$

$$\rho_t(l) \equiv \zeta_{t+l} - \hat{\zeta}_t(l) = \sum_{i=0}^{l-1} \psi_i \varrho_{t+l-i}, \tag{23}$$

for any $t \in \mathbb{T}$ and $l \in \mathbb{L}$. $\hat{\zeta}_t(l)$ can be naturally interpreted as a forecast of ζ_{t+l} given information up to time t and $\rho_t(l)$ as the forecast error.

If we set $\rho_{t-1,L} \equiv (\rho_{t-1}(1), \cdots, \rho_{t-1}(L))^\top$ for any $t \in \mathbb{Z}$, we can then express the forecast error vector $\rho_{t-1,L}$ as a linear function of the independent $\varrho_{[t,t+L-1]}$. More specifically, $\rho_{t-1,L} = V_t \varrho_{[t,t+L-1]}$ holds for all $L \in \{1, \cdots, T-t+1\}$, where $V_t \equiv V \in \mathbb{R}^{L \times L}$ is the following invertible and lower triangular square matrix, which is constant across all $t \in \mathbb{Z}$:

$$V = \begin{pmatrix} 1 & & \cdots & 0 \\ \psi_1 & 1 & & \vdots \\ \vdots & & \ddots & \\ \psi_{L-1} & \cdots & \psi_1 & 1 \end{pmatrix} \tag{24}$$

We then have the system of equalities:

$$\zeta_{[t,t+L-1]} = \hat{\zeta}_{t-1,L} + V\varrho_{[t,t+L-1]} \tag{25}$$

and it follows that $\xi_{t+l} = e^{\bar{v}_{t+l} + \hat{\zeta}_{t-1}(l+1) + \rho_{t-1}(l+1)}, \forall l \in \mathbb{L}$ and $\hat{\zeta}_{t-1,L} + \bar{v}_{[t,t+L-1]} \equiv v_t \in \mathbb{R}^L$ in representation (17) where $\hat{\zeta}_{t-1,L} \equiv (\hat{\zeta}_{t-1}(1), \cdots, \hat{\zeta}_{t-1}(L))^\top$ and $\bar{v}_{[t,t+L-1]} = (\bar{v}_t, \cdots, \bar{v}_{t+L-1})^\top$. For more details, see [5].

10.2 Polyhedral support of the ϱ_t

We fix some $\epsilon > 0$ and consider the following polyhedron:

$$P = \{\varrho \in \mathbb{R}^L : \sum_{i=1}^{L} |\varrho_i| \leqslant L\sqrt{\epsilon\sigma_\varrho^2}; |\varrho_i| \leqslant \sqrt{L\epsilon\sigma_\varrho^2}, \forall i\}. \tag{26}$$

This set is motivated by Markov's inequality. Indeed, for the independent random variables $\tilde{\varrho} \equiv \varrho_{[t,t+L-1]}$, we know that $\mathbb{P}(\tilde{\varrho} \in P) \geqslant 1 - \epsilon^{-1}$. For more details, see [5].

10.3 Counterexample showing that Ξ_t is generally non-convex

We show that in general, Ξ_t is not convex for an arbitrary $t \in \mathbb{T}$. Consider the following instance:

$$
\begin{aligned}
&V = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \\
&v_t = (0,0)^\top \\
&P = \{\varrho \in \mathbb{R}^L : -1 \leqslant \varrho_l \leqslant 1, \forall l = 1, \cdots, L; \sum_{i=1}^{L} |\varrho_i| \leqslant \sqrt{L}\} \\
&L = 2.
\end{aligned}
$$

Figure (4) displays $\Xi_t = \{\xi \in \mathbb{R}^L : \exists \varrho \in P, \xi_l = \exp(V_l^\top \varrho), \forall l = 1, \cdots, L\}$ and illustrates that Ξ_t is in general not convex.

For a slightly more formal demonstration, it is possible to show that given the two points $\hat{\varrho}_1 = (1 - \sqrt{2}, 1)^\top \in P$ and $\hat{\varrho}_2 = (\sqrt{2} - 1, 1)^\top \in P$ illustrated in Figure (4) as well as $\lambda = \frac{1}{2}$, there exists no $\varrho \in P$ such that:

$$
\lambda \begin{pmatrix} e^{V_1^\top \hat{\varrho}_1} \\ e^{V_2^\top \hat{\varrho}_1} \end{pmatrix} + (1-\lambda) \begin{pmatrix} e^{V_1^\top \hat{\varrho}_2} \\ e^{V_2^\top \hat{\varrho}_2} \end{pmatrix} = \begin{pmatrix} e^{V_1^\top \varrho} \\ e^{V_2^\top \varrho} \end{pmatrix}. \tag{27}
$$

Equivalently, we can show that $\forall \varrho \in P$,

$$
\left\| \begin{pmatrix} e^{V_1^\top \varrho} \\ e^{V_2^\top \varrho} \end{pmatrix} - \lambda \begin{pmatrix} e^{V_1^\top \hat{\varrho}_1} \\ e^{V_2^\top \hat{\varrho}_1} \end{pmatrix} + (1-\lambda) \begin{pmatrix} e^{V_1^\top \hat{\varrho}_2} \\ e^{V_2^\top \hat{\varrho}_2} \end{pmatrix} \right\|_\infty > 0.
$$

This can be shown by solving the following linear program and observing that its optimal value is strictly larger than 0:

$$
\begin{aligned}
\min_{\varrho^+ \geqslant 0, \varrho^- \geqslant 0, t \geqslant 0} \quad & t \\
\text{s. t.} \quad & \sum_{i=1}^{2} (\varrho_i^+ + \varrho_i^-) \leqslant \sqrt{2} \\
& \varrho_i^+ + \varrho_i^- \leqslant 1, \quad \forall i = 1, 2 \\
& V_i^\top (\varrho^+ - \varrho^-) - l_i^\lambda \leqslant t, \quad i = 1, 2 \\
& l_i^\lambda - V_i^\top (\varrho^+ - \varrho^-) \leqslant t, \quad i = 1, 2,
\end{aligned}
$$

where $l_i^\lambda = ln(\lambda e^{V_i^\top \hat{\varrho}_1} + (1-\lambda) e^{V_i^\top \hat{\varrho}_2})$ is a known constant.

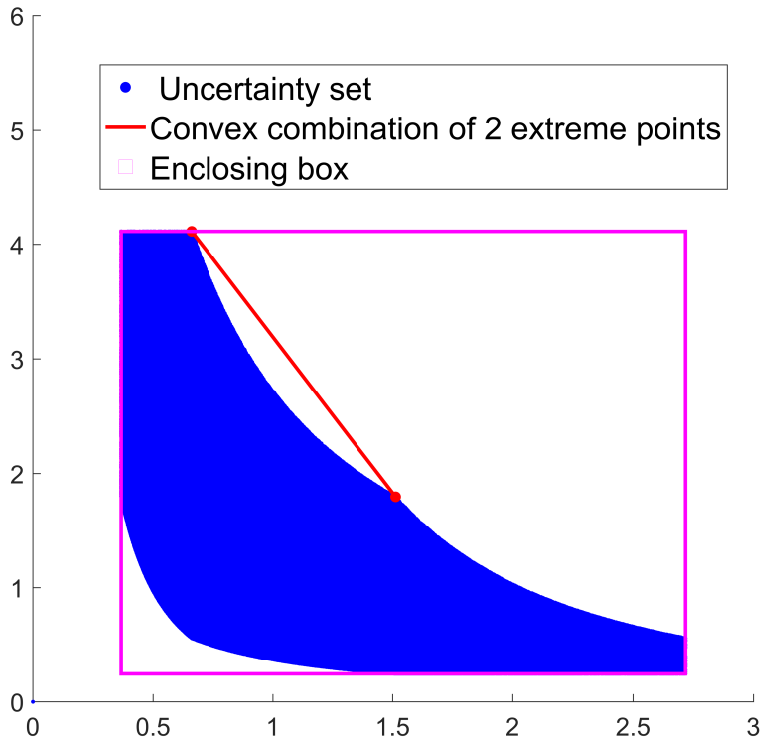

Figure 4: Non convex uncertainty set

11 Analytical expression for the composite risk

11.1 Conditional expected value of $\sum_{l=0}^{L-1} \mathcal{Q}_{t+l}^{floods}(\xi)$ and $\sum_{l=0}^{L-1} \mathcal{Q}_{t+l}^{prod}(\xi)$

By considering affine decision rules, deriving an analytical expression for the flood and production risk becomes straightforward. Indeed,

$$\mathrm{E}\left[\sum_{l=0}^{L-1} \mathcal{Q}_{t+l}^{floods}(\xi)|\mathcal{G}_{t-1}\right] = \sum_{l=0}^{L-1}\sum_{j\in J} \gamma_j \mathrm{E}\left[\bar{\mathcal{E}}_{j,t+l}(\xi)|\mathcal{G}_{t-1}\right], \tag{28}$$

where

$$\mathrm{E}\left[\bar{\mathcal{E}}_{j,t+l}(\xi)|\mathcal{G}_t\right] = \bar{\mathcal{E}}_{j,t+l}^0 + \sum_{l'=0}^{L-1} \bar{\mathcal{E}}_{j,t+l}^{t+l'} \mathrm{E}\left[\xi_{t+l'}|\mathcal{G}_t\right]. \tag{29}$$

The production risk is defined as:

$$\begin{aligned}
&\mathrm{E}\left[\sum_{l=0}^{L-1} \mathcal{Q}_{t+l}^{prod}(\xi)|\mathcal{G}_{t-1}\right] \\
&\quad = -\sum_{l=0}^{L-1}\sum_{i\in I^{prod}} \mathrm{E}\left[\mathcal{P}_{i,t+l}(\xi)\mathcal{H}_{i,t+l}(\xi)|\mathcal{G}_{t-1}\right] \\
&\qquad -\sum_{j\in J} \mathrm{E}\left[\bar{\mathcal{S}}_{j,t+L-1}(\xi)|\mathcal{G}_{t-1}\right].
\end{aligned} \tag{30}$$

The conditional expected value of the final value of storages $\mathrm{E}\left[\bar{\mathcal{S}}_{j,t+L-1}(\xi)|\mathcal{G}_{t-1}\right]$ is defined similarly to (29) while $\mathrm{E}\left[\mathcal{P}_{i,t+l}(\xi)\mathcal{H}_{i,t+l}(\xi)|\mathcal{G}_{t-1}\right]$, which involves the product of two affine functions, is defined as:

$$\begin{aligned}\mathrm{E}\left[\mathcal{P}_{i,t+l}(\xi)\mathcal{H}_{i,t+l}(\xi)|\mathcal{G}_{t-1}\right] = \quad & \mathcal{P}^0_{i,t+l}\mathcal{H}^0_{i,t+l} \\ & + \sum_{k=0}^{L-1}(\mathcal{P}^0_{i,t+l}\mathcal{H}^k_{i,t+l} + \mathcal{H}^0_{i,t+l}\mathcal{P}^k_{i,t+l})\mathrm{E}\left[\xi_{t+k}|\mathcal{G}_{t-1}\right] \\ & + \sum_{m=0}^{L-1}\sum_{k=0}^{L-1}\mathcal{P}^m_{i,t+l}\mathcal{H}^k_{i,t+l}\mathrm{E}\left[\xi_{t+m}\xi_{t+k}|\mathcal{G}_{t-1}\right]. \end{aligned} \tag{31}$$

The conditional expectations $\mathrm{E}\left[\xi_{t+k}|\mathcal{G}_{t-1}\right]$ and $\mathrm{E}\left[\xi_{t+k}\xi_{t+m}|\mathcal{G}_{t-1}\right]$ in (29) and (31) are detailed in the following Section (11.2).

11.2 Conditional expected value of ξ_{t+l}

Theorem 11.1. *For any $t \in \mathbb{Z}$ and $l \in \mathbb{N}$:*

$$E[\xi_{t+l}|\mathcal{G}_t] = e^{\hat{\zeta}_t(l)} \cdot E\left[e^{\rho_t(l)}\right] \tag{32}$$

and for $k \geqslant l \geqslant 1$:

$$\begin{aligned}E[\xi_{t+l}\xi_{t+k}|\mathcal{G}_t] = e^{\hat{\zeta}_t(l)+\hat{\zeta}_t(k)} \cdot E\left[e^{\sum_{i=0}^{k-l-1}\psi_i\varrho_{t+k-i}}\right] \\ \cdot E\left[e^{\sum_{i=k-l}^{k-1}(\psi_i+\psi_{l-k+i})\varrho_{t+k-i}}\right]. \end{aligned} \tag{33}$$

Proof. This is a direct application of Theorem 34.3 of [53] together with the observation that the $\{\varrho_t\}$ are independent and hence $e^{\rho_t(l)} = e^{\sum_{i=0}^{l-1}\psi_i\varrho_{t+l-i}}$ is independent of $\{\varrho_\tau\}_{\tau=-\infty}^{t}$, which in turn implies that $\mathrm{E}\left[e^{\rho_t(l)}|\mathcal{G}_t\right] = \mathrm{E}\left[e^{\rho_t(l)}\right]$. □ □

At time t, $e^{\hat{\zeta}_t(l)}$ is a known deterministic value for any $l \in \mathbb{N}$ that depends on the past observed $\{\varrho_\tau\}_{\tau=-\infty}^{t}$. To compute $\mathrm{E}\left[e^{\sum_{i=0}^{k}\chi_i\varrho_{t+i}}\right]$ for any $k \in \{0,\cdots,L\}$ and $\chi_i \in \mathbb{R}, \forall i \in \{0,\cdots,L\}$, we perform Monte Carlo simulation.

12 First order Taylor approximation of the composite risk

We first fix $\mathrm{E}\left[\mathcal{P}_{i,t+l}(\xi)\mathcal{H}_{i,t+l}(\xi)|\mathcal{G}_{t-1}\right] \equiv F_{i,t+l}(\mathcal{H}_{i,t+l}, \mathcal{P}_{i,t+l})$ where $\mathcal{H}_{i,t+l} = (\mathcal{H}^0_{i,t+l}, \mathcal{H}^{t+1}_{i,t+l}, \cdots, \mathcal{H}^{t+L-1}_{i,t+l})^\top \in \mathbb{R}^L$ and $\mathcal{P}_{i,t+l} = (\mathcal{P}^0_{i,t+l}, \mathcal{P}^{t+1}_{i,t+l}, \cdots, \mathcal{P}^{t+L-1}_{i,t+l})^\top \in \mathbb{R}^L$. Given the point $(\hat{\mathcal{H}}^\top_{i,t+l}, \hat{\mathcal{P}}^\top_{i,t+l})^\top \in \mathbb{R}^{2L}$, we then obtain:

$$
\begin{aligned}
&F_{i,t+l}(\mathcal{H}_{i,t+l}, \mathcal{P}_{i,t+l}) \\
&\quad \approx F_{i,t+l}(\hat{\mathcal{H}}_{i,t+l}, \hat{\mathcal{P}}_{i,t+l}) \\
&\qquad + \hat{\mathcal{H}}^0_{i,t+l}(\mathcal{P}^0_{i,t+l} - \hat{\mathcal{P}}^0_{i,t+l}) \\
&\qquad + \sum_{k=0}^{L-1} \hat{\mathcal{H}}^k_{i,t+l}(\mathcal{P}^0_{i,t+l} - \hat{\mathcal{P}}^0_{i,t+l})\mathrm{E}\left[\xi_{t+k}|\mathcal{G}_{t-1}\right] \\
&\qquad + \hat{\mathcal{P}}^0_{i,t+l}(\mathcal{H}^0_{i,t+l} - \hat{\mathcal{H}}^0_{i,t+l}) \\
&\qquad + \sum_{k=0}^{L-1} \hat{\mathcal{P}}^k_{i,t+l}(\mathcal{H}^0_{i,t+l} - \hat{\mathcal{H}}^0_{i,t+l})\mathrm{E}\left[\xi_{t+k}|\mathcal{G}_{t-1}\right] \\
&\qquad + \sum_{k=0}^{L-1} \hat{\mathcal{P}}^0_{i,t+l}(\mathcal{H}^{t+k}_{i,t+l} - \hat{\mathcal{H}}^{t+k}_{i,t+l})\mathrm{E}\left[\xi_{t+k}|\mathcal{G}_{t-1}\right] \\
&\qquad + \sum_{m=0}^{L-1}\sum_{k=0}^{L-1} \hat{\mathcal{P}}^{t+m}_{i,t+l}(\mathcal{H}^{t+k}_{i,t+l} - \hat{\mathcal{H}}^{t+k}_{i,t+l})\mathrm{E}\left[\xi_{t+m}\xi_{t+k}|\mathcal{G}_{t-1}\right] \\
&\qquad + \sum_{k=0}^{L-1} \hat{\mathcal{H}}^0_{i,t+l}(\mathcal{P}^{t+k}_{i,t+l} - \hat{\mathcal{P}}^{t+k}_{i,t+l})\mathrm{E}\left[\xi_{t+k}|\mathcal{G}_{t-1}\right] \\
&\qquad + \sum_{m=0}^{L-1}\sum_{k=0}^{L-1} \hat{\mathcal{H}}^{t+m}_{i,t+l}(\mathcal{P}^{t+k}_{i,t+l} - \hat{\mathcal{P}}^{t+k}_{i,t+l})\mathrm{E}\left[\xi_{t+m}\xi_{t+k}|\mathcal{G}_{t-1}\right]
\end{aligned} \tag{34}
$$

References

[1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: forecasting and control, 4th edition*. John Wiley & Sons, Inc., Hoboken, New Jersey., 2008.

[2] J. Labadie, "Optimal operation of multireservoir systems: state-of-the-art review," *Journal of Water Resources Planning and Management*, vol. 130, pp. 93–111, 2004.

[3] A. Castelletti, F. Pianosi, and R. Soncini-Sessa, "Water reservoir control under economic, social and environmental constraints," *Automatica*, vol. 44, no. 6, pp. 1595–1607, 2008.

[4] C. Gauvin, E. Delage, and M. Gendreau, "Decision rule approximations for the risk averse reservoir management problem," *European Journal of Operational Research*, vol. 261, pp. 317–336, January 2017.

[5] ——, "A stochastic program with tractable time series and affine decision rules for the reservoir management problem," Les cahiers du GERAD, Tech. Rep. G-2016-24, 2016.

[6] A. Lorca and X. A. Sun, "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Transactions on Power Systems*, 2015.

[7] Álvaro Lorca, X. A. Sun, E. Litvinov, and T. Zheng, "Multistage adaptive robust optimization for the unit commitment problem," *Operations Research*, vol. 64, no. 1, pp. 32–51, 2016.

[8] S. V. Braaten, O. Gjonnes, K. Hjertvik, and S.-E. Fleten, "Linear decision rules for seasonal hydropower planning: Modelling considerations," *Energy Procedia*, vol. 87, pp. 28 – 35, 2016, 5th International Workshop on Hydro Scheduling in Competitive Electricity Markets.

[9] C. S. ReVelle and W. Kirby, "Linear decision rule in reservoir management and design, 2, performance optimization," *Water Resources Research*, vol. 6, pp. 1033–1044, 1970.

[10] A. Ben-Tal, E. Goryashko, A. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Mathematical Programming*, vol. 99, pp. 351–378, 2004.

[11] D. Kuhn, W. Wiesemann, and A. Georghiou, "Primal and dual linear decision rules in stochastic and robust optimization," *Mathematical Programming Series A*, vol. 130, pp. 177–209, 2011.

[12] A. Turgeon, "Solving a stochastic reservoir management problem with multilag autocorrelated inflows," *Water Resources Research*, vol. 41, p. W12414, 2005.

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

[14] F. Pianosi and R. Soncini-Sessa, "Real-time management of a multipurpose water reservoir with a heteroscedastic inflow model," *Water Resources Research*, vol. 45, no. 10, pp. 1–12, 2009, w10430.

[15] T. Lohmann, A. S. Hering, and S. Rebennack, "Spatio-temporal hydro forecasting of multireservoir inflows for hydro-thermal scheduling," *European Journal of Operational Research*, vol. 255, no. 1, pp. 243 – 258, 2016.

[16] A. Shapiro, "Analysis of stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 209, pp. 63–72, 2011.

[17] H. Poorsepahy-Samian, V. Espanmanesh, and B. Zahraie, "Improved inflow modeling in stochastic dual dynamic programming," *Journal of Water Resources Planning and Management*, vol. 142, no. 12, December 2016.

[18] A. Shapiro, W. Tekaya, J. P. da Costa, and M. P. Soares, "Final report for technical cooperation between georgia institute of technology and ons - operador nacional do sistema eletrico," Georgia Institute of Technology and Operador Nacional do Sistema Eletrico, Tech. Rep., 2012.

[19] D. De Ladurantaye, M. Gendreau, and J.-Y. Potvin, "Optimizing profits from hydroelectricity production," *Computers and Operations Research*, vol. 36, pp. 499–529, 2009.

[20] L. Pan, M. Housh, P. Liu, X. Cai, and X. Chen, "Robust stochastic optimization for reservoir operation," *Water Resources Research*, vol. 51, no. 1, pp. 409–429, 2015.

[21] S. Séguin, P. Côté, and C. Audet, "Short-term unit commitment and loading problem," Les cahiers du GERAD, Tech. Rep., 2014.

[22] G. Steeger and S. Rebennack, "Dynamic convexification within nested benders decomposition using lagrangian relaxation: An application to the strategic bidding problem," *European Journal of Operational Research*, vol. 257, no. 2, pp. 669 – 686, 2017.

[23] S. Cerisola, J. M. Latorre, and A. Ramos, "Stochastic dual dynamic programming applied to nonconvex hydrothermal models," *European Journal of Operational Research*, vol. 218, no. 3, pp. 687 – 697, 2012.

[24] T. N. Santos and A. L. Diniz, "A new multiperiod stage definition for the multistage benders decomposition approach applied to hydrothermal scheduling," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1383–1392, August 2009.

[25] A. L. Diniz and M. E. P. Maceira, "A four-dimensional model of hydro generation for the shortterm hydrothermal dispatch problem considering head and spillage effects," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1298–1308, August 2000.

[26] T. N. dos Santos and A. L. Diniz, "A new multiperiod stage definition for the multistage benders decomposition approach applied to hydrothermal scheduling," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1383–1392, Aug 2009.

[27] F. Thome, M. Pereira, S. Granville, and M. Fampa, "Non-convexities representation on hydrothermal operation planning using sddp."

[28] A. Phillpott, F. Wahid, and F. Bonnans, "Midas: A mixed integer dynamic approximation scheme," Tech. Rep., 2016. [Online]. Available: http://www.optimization-online.org/DB_FILE/2016/05/5431.pdf

[29] A. L. Diniz and T. M. Souza, "Short-term hydrothermal dispatch with river-level and routing constraints," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2427–2435, Sept 2014.

[30] R. S. Tsay, *Analysis of Financial Time Series, Second Edition.* John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.

[31] B. Klöckl and G. Papaefthymiou, "Multivariate time series models for studies on stochastic generators in power systems," *Electric Power Systems Research*, vol. 80, no. 2010, 2009.

[32] A. Tilmant and R. Kelman, "A stochastic approach to analyze trade-offs and risk associated with large-scale water resources systems," *Water Resources Research*, vol. 43, p. W06425, 2007.

[33] M. Maceira, V. Duarte, D. Penna, L. Moraes, and A. Melo, "Ten years of application of stochastic dual dynamic programming in official and agent studies in brazil - description of the newave program," in *16th PSCC, Glasgow, Scotland*, 2008.

[34] P.-L. Carpentier, M. Gendreau, and F. Bastin, "Long-term management of a hydroelectric multireservoir system under uncertainty using the progressive hedging algorithm," *Water Resources Research*, vol. 49, pp. 2812–2827, 2013.

[35] D. De Ladurantaye, M. Gendreau, and J.-Y. Potvin, "Strategic bidding for price-taker hydroelectricity producers," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2187–2203, 2007.

[36] A. Gjelsvik, B. Mo, and A. Haugstad, *Long- and Medium-term Operations Planning and Stochastic Modelling in Hydro-dominated Power Systems Based on Stochastic Dual Dynamic Programming.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 33–55.

[37] A. Hamann, G. Hug, and S. Rosinski, "Real-time optimization of the mid-columbia hydropower system," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 157–165, January 2017.

[38] S. Séguin, P. Côté, and C. Audet, "Self-scheduling short-term unit commitment and loading problem," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 133–142, 2016.

[39] C. A. Bana e Costa and J.-C. Vansnick, *A theoretical framework for measuring attractiveness by a categorical based evaluation technique (MACBETH)*, J. Climaco, Ed. Springer, 1997.

[40] A. Castelletti, S. Galetti, M. Restelli, and R. Soncini-Sessa, "Tree-based reinforcement learning for optimal water reservoir operation," *Water Resources Research*, vol. 46, p. W09507, 2010.

[41] A. Shapiro, W. Tekaya, J. P. da Costa, and M. P. Soares, "Risk neutral and risk averse stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 224, no. 2, pp. 375 – 391, 2013.

[42] A. Turgeon and R. Charbonneau, "An aggregation-disaggregation approach to long-term reservoir management," *Water Resources Research*, vol. 34, pp. 3585–3594, 1998.

[43] B. Bezerra, . Veiga, L. A. Barroso, and M. Pereira, "Stochastic long-term hydrothermal scheduling with parameter uncertainty in autoregressive streamflow models," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 999–1006, March 2017.

[44] M. Maceira and J. Damázio, "The use of par(p) model in the stochastic dual dynamic programming optimization scheme used in the operation planning of the brazilian hydropower system," ser. 8th International Conference on Probabilistic Methods Applied to Power Systems. Iowa State University, September 12-16 2004, pp. 397–402.

[45] G. E. P. Box and D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society*, vol. 2, pp. 211–252, 1964.

[46] R. Egging, S. E. Fleten, I. Grønvik, A. Hadziomerovic, and N. Ingvoldstad, "Linear decision rules for hydropower scheduling under uncertainty," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 103–113, 2017.

[47] J. Nocedal and S. J. Wright, *Numerical Optimization, 2nd Edition*, ser. Springer Series in Operations Research. Springer, 2006.

[48] J. T. Needham, D. W. Watkins, J. R. Lund, and S. K. Nanda, "Linear programming for flood control in the iowa and des moines rivers," *Journal of water resources planning and management*, vol. 126, no. 3, May 2000.

[49] X. Li, S. Guo, P. Liu, and G. Chen, "Dynamic control of flood limited water level for reservoir operation by considering inflow uncertainty," *Journal of Hydrology*, vol. 391, no. 1, pp. 124 – 132, 2010.

[50] C.-C. Wei and N.-S. Hsu, "Optimal tree-based release rules for real-time flood control operations on a multipurpose multireservoir system," *Journal of Hydrology*, vol. 365, no. 3, pp. 213 – 224, 2009.

[51] A. Borghetti, C. D'Ambrosio, A. Lodi, and S. Martello, "An milp approach for short-term hydro scheduling and unit commitment with head-dependent reservoir," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1115–1124, Aug 2008.

[52] P. J. Brockwell and R. A. Davis, *Time series: theory and methods.* Springer-Verlag New York, 1987.

[53] P. Billingsley, *Probability and measure, third edition.* John Wiley & Sons Inc, 1995.