Sustainable Management of a Water Reservoir System Facing Climate Uncertainty

Felipe Caro
UCLA Anderson School of Management
felipe.caro@anderson.ucla.edu

Martin Glanzer
University of Mannheim Business School
martin.glanzer@uni-mannheim.de

Kumar Rajaram
UCLA Anderson School of Management
kumar.rajaram@anderson.ucla.edu

Abstract. Water scarcity is a growing global issue. It affects even highly developed regions like California, the world's fifth largest economy, where years of severe droughts interspersed with unusually wet seasons have caused societal debates over the management of the water resources stored in the state's extensive reservoir system. This paper presents a strategic optimization model for the long-term, sustainable management of such a system. Introducing a new modeling paradigm based on cycles of stochastic length, defined by the event when all reservoirs are simultaneously full, we avoid the limitations of traditional models and obtain sustainable management policies. To address the unpredictable effect of climate change on future water supply, we adopt a distributionally robust framework where nature chooses adverse inflow distributions. This leads to a stochastic shortest path problem under distributional ambiguity. Using tools from stochastic dynamic programming and aggregation methods, we obtain policy insights and overcome the curse of dimensionality typically associated with systems models. In a case study for California's Sacramento River Basin, we report suboptimality gaps between 3% and 15%, with our sustainable management policy reducing average cycle shortage costs by 40% compared to the current policy, demonstrating the significant cost saving potential.

Key words: Water resource management, distributionally robust optimization, stochastic shortest path

1. Introduction

Water reservoirs are man-made structures designed to store available water resources, while dams control their outflow. Typically, reservoirs are filled during periods with high precipitation to ensure stable water supplies during dry seasons. This concept has been benefiting humanity for millennia (see Fahlbusch 2009), from the famous Jawa Dam in ancient Mesopotamia to the sophisticated reservoir systems of the Roman Empire, and the over 91,000 dams constructed across the United States since the nineteenth century (see American Society of Civil Engineers, ASCE 2021). Today, the competing needs of urban populations, agriculture, industry, and the environment make the storage and distribution of water resources a contentious issue, both socially and economically.

In addition, climate change has significantly increased the unpredictability of precipitation patterns. These challenges call for sustainable water resource management policies, a need that is also reflected in several United Nations Sustainable Development Goals (SDGs), especially SDG 6 (“Clean Water and Sanitation”) and SDG 12 (“Responsible Consumption and Production”).

Day-to-day reservoir management typically involves short-term decision support models based on planning horizons ranging from a few days to the end of a season. The report of Georgakakos et al. (2018) comprehensively details a decision support model used by the California Department of Water Resources. Depending on the prevailing regime, the focus of short-term operations models is either on flood protection (during wet periods) or on satisfying immediate demand subject to detailed operational constraints (during dry periods). Focusing on a single wet year can lead to extremely low reservoir levels if there is an extended drought. For instance, in July 2022, after a years-long dry spell, Shasta Lake in California was at 38% capacity, the lowest point on record for that time of the year (Chun 2022).

In contrast, sustainable management requires taking future uncertainties into account that go beyond immediate needs and currently available resources. This requires strategic planning on a coarser time scale, delegating intraseasonal operations management to complementary models.

Aiming at a strategic planning model raises the question of an appropriate planning horizon. Traditional approaches face limitations: finite-horizon models require artificial terminal conditions to avoid depleting reservoirs at the end of the planning period, and infinite-horizon models yield solutions that depend heavily on the discount rate used for future water shortages. To address these issues, we propose a novel modeling paradigm based on a stochastic time horizon. We define a cycle as a period of time in which all reservoirs in the system are simultaneously full at the beginning and at the end, but not at any intermediate time. The length of such a cycle is influenced by releases and inflows (and, to a minor extent, evaporation losses). Inflows are uncertain, while releases can be controlled. The objective of our model is to manage releases in such a way as to minimize the total expected shortage costs over a cycle. This creates a non-trivial trade-off between meeting immediate demand and shortening the expected cycle, i.e., getting the system back to full. Further trade-offs can be incorporated by assigning higher shortage costs to prioritized demand streams, such as drinking water. Then, immediate non-essential demand competes with future high-priority demand.

Historical yearly inflow data exhibit significant variability, making seasonal inflow a key source of exogenous uncertainty in our model. However, reliable probabilistic forecasts are difficult, not only due to relatively few available data points but especially because of climate change, see Le et al. (2023). We address this challenge by incorporating ambiguity into our model. We introduce an adversarial player ("nature") to choose the worst-case inflow distribution for any given release decision. We adopt a penalty-based distributionally robust optimization approach, which uses relative entropy to measure the deviations of the worst-case distribution from our best-guess nominal uncertainty model. This method ensures that the policies induced by our model are robust against climate uncertainty.

We consider a system of reservoirs that collectively satisfy demand, exemplified by California, where most major reservoirs are located in the Sacramento River Basin in the northern part. This system contributes in a controlled way to the Sacramento River, from which urban and ecological demand in the Bay Area as well as agricultural demand in the southern Central Valley are supplied. The case of California, for which we carefully assessed all details in consultation with the operating agency (White 2022), serves as a running example across the paper. However, our results are readily applicable to any other reservoir system sharing similar characteristics.

For any such system, efficient distribution of available water resources requires a systemic perspective, as reflected in our model. From a computational perspective, however, stochastic dynamic optimization models typically suffer from a curse of dimensionality. Considering all possible states of a dynamic system of even a few reservoirs can thus be computationally intractable. To overcome this issue, we suggest an efficiently computable heuristic based on two steps. First, we use an aggregation approach to determine the total release amount from the system. Then, we exploit our analytical policy insights to optimally split the total release among the individual reservoirs. The expected costs for the aggregated system are shown to represent a lower bound for the optimal (worst-case) expected costs. While the worst-case inflow distributions cannot be extracted in closed form, we derive upper bounds on the worst-case expected shortage costs associated with our policy, allowing for performance evaluation.

Water resources management has a tradition of research at the interface of several scientific communities, generating a large body of literature. The survey of Elleuch et al. (2022) reports nearly seven hundred publications between the year 1985 and 2021 on water management problems, classifying the related literature along several dimensions. Giuliani et al. (2021) categorize over 300 papers on prescriptive reservoir management models with respect to the applied policy design, including a detailed analysis of 114 papers using (approximate) stochastic dynamic programming techniques. Finite-horizon models with end-of-horizon penalties and infinite-horizon models with discounting (or considering average immediate costs) are identified in their review as the only two approaches used across the existing literature. Hence, our proposed approach of using a stochastic time horizon introduces a new modeling paradigm, enriching the existing literature.

Extreme weather events have generally led to increased awareness of climate change and underscored the need for sustainable management of natural resources. Yet, the recent literature remains sparse in presenting innovative models for strategically optimizing the distribution of water resources facing climate uncertainty. An exception is the work of Park and Bayraksan (2023), who employ a scenario tree based multistage distributionally robust optimization approach to model water distribution in the area of Tucson, Arizona. While Park and Bayraksan (2023) also propose integrating climate uncertainty into strategic water resource management models through distributionally robust optimization, our paper differs significantly both from a modeling and methodological perspective. Unlike their focus on algorithmic solution techniques for scenario tree models, our stochastic dynamic programming approach allows for finer time granularity of yearly decisions and uncertain parameters, emphasizing policy insights and implications. Moreover, their model considers a finite horizon (the middle of the century) with discounting, whereas we introduce a new modeling approach. Gauvin et al. (2017, 2018a,b) also apply robust optimization techniques. However, in contrast to our strategic planning model, their models for a river system in Western Québec address flood risk management over a natural finite horizon defined by the freshet period.

From an inventory management perspective, the studied optimization problem can be seen as a problem of sequentially matching random supply (under model ambiguity) into a network of storage facilities, with known central demand. To our knowledge, such models have not been extensively studied in the classical inventory management literature. While random supply issues have been explored, for instance, in the context of distributing donated products (see, e.g., Zhang et al. 2020), water reservoir management remains a distinct dynamic resource allocation problem with unique challenges.

The main contributions of this paper can be summarized as follows:

1. **A New Strategic Model** (Section 2): We present a strategic model for sustainable water resource management across a network of reservoirs facing climate uncertainty. The model prescribes the aggregate release from each reservoir in the system over an entire upcoming (dry) season, given all storage levels at the beginning of the season. The strategic nature of the model stems from its objective: we propose to optimize the expected sum of shortage costs

over cycles of stochastic length, thereby overcoming the limitations of standard modeling approaches. Acknowledging epistemic inflow uncertainty due to climate change, we incorporate ambiguity, optimizing expected cycle costs with respect to worst-case inflow distributions.

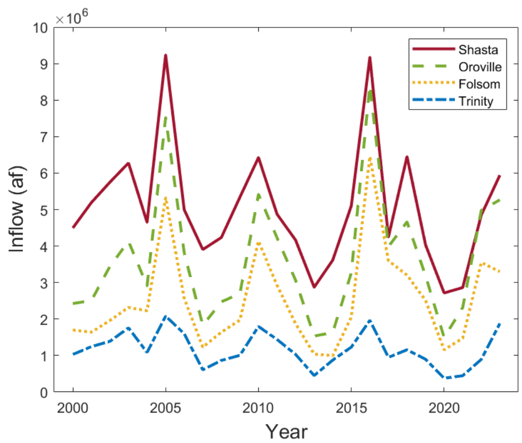
2. **Theoretical Analysis and Policy Insights** (Section 3): The proposed model is thoroughly analyzed, deriving structural properties and policy insights. The analytical results underscore the intuitive nature of our model. Moreover, we establish a link to the literature on stochastic shortest path models in the sense of Bertsekas and Tsitsiklis (1991), presenting a distributionally robust version for the first time (to the best of our knowledge).
3. **Heuristics and Bounds** (Section 4): From a computational perspective, we overcome the curse of dimensionality using aggregation methods, suggesting a scalable heuristic together with performance bounds.
4. **Case Study** (Section 5): Using real data, we apply our model to the case of the Sacramento River Basin, the largest water resource system in California. In a comprehensive computational study, we demonstrate the reliability and robustness of our policy. Suboptimality gaps ranging from 3% to 15% indicate that our heuristic is close to optimal. Applying historical inflow scenarios, our policy significantly outperforms several benchmark policies. In particular, the average cycle cost is reduced by 40% compared to the policy used in practice.

In Section 6, we present our conclusions. All proofs and additional supplemental material are provided in the Electronic Companion (EC).

2. Model Formulation

We consider a discrete-time dynamic system of $N \in \mathbb{N}$ reservoirs located in a river basin. The capacity of reservoir i is C_i and its state, $S_{t,i} \in [S_i^{\min}, C_i]$, describes the volume of water that is stored in reservoir i at time $t \geq 0$. Minimum levels $S_i^{\min} > 0$ are typically imposed for environmental reasons. Let $\mathcal{S} = [S_1^{\min}, C_1] \times \ldots \times [S_N^{\min}, C_N]$ denote the state space of the system.

California's official water year is a 12 months period starting on October 1 of the calendar year that precedes the reference year. For instance, the water year 2000 started on Oct. 1, 1999 and ended on Sep. 30, 2000. Each water year is assumed to consist of two seasons, namely a rainy/wet season (October to April) and a dry season (May to September). On average, about 70% of the yearly inflow occurs through direct precipitation during the wet season, while the inflow during the dry season is mostly due to snow melt. Each time point $t \geq 0$ in our discrete-time system marks the beginning of a dry season, i.e., May 1 of year t. The inflow into reservoir i between time t and $t+1$, which corresponds to the aggregate inflow over a dry season and the subsequent rainy season, is modeled by $\alpha_i \cdot \Gamma_t$, where Γ_t is an independent realization of a univariate random variable Γ representing the reference annual inflow, and $\alpha_i \geq 0$ is a constant factor. Let $\mathbb{P}$ denote the stationary distribution of Γ. Using a univariate source of exogenous uncertainty is motivated by the fact that we consider a system of reservoirs belonging to a single river basin. The assumption can be validated by observing high correlations in historical inflow data. Figure 1 illustrates such correlations for the major reservoirs in the Sacramento River basin.



Correlation	Trinity	Folsom	Oroville
Shasta	88%	85%	91%
Trinity		74%	82%
Folsom			98%

Figure 1 Left: Yearly inflow data (May to April of the following year) from the (starting) year 2000 to 2023. Right: Corresponding correlation matrix.

Reservoir management in the two seasons follows two completely different objectives. During the wet season, the focus is on flood protection and dam failure must be avoided by all means. Such flood protection policies are typically implemented in practice by dynamically adjusting the so-called top of conservation storage level. Upon reaching that level, any additional inflow is immediately released. The top of conservation storage level is set high, if either the current storage level is low, or the end of the wet season is approaching, or little precipitation is predicted for the remainder of the season. By the end of the wet season, the top of conservation storage level typically reaches one hundred percent of storage capacity. As our model considers the entire wet season as a whole without actively controlling its releases, we simplify the exogenous flood protection policy by multiplying the (yearly) inflow $\alpha_i\Gamma$ into reservoir i by a constant factor $(1-\beta_i)$, such that $\beta_i\alpha_i\Gamma$ represents the outflow from reservoir i over a wet season. This simplifying assumption reflects the fact that large/small release amounts are typically observed in wet seasons with large/small precipitation amounts (for example, the correlation between wet season outflows and yearly inflows is more than 90% in our data for Shasta Lake).

To account for evaporation losses, a constant amount of stored water e_i is subtracted from reservoir i during each time period. For ease of notation, let $\xi_{t,i} := (1-\beta_i)\alpha_i\Gamma_t - e_i$ be the uncontrolled net inflow into reservoir i between time t and $t+1$, and let $\xi_t = (\xi_{t,1}, \dots, \xi_{t,N})$. Let also $S_t = (S_{t,1}, \dots, S_{t,N})$ and $C = (C_1, \dots, C_N)$ denote the state and capacity vectors respectively.

The outflow during dry seasons is the focus of our management model. Given the state $S_t \in \mathcal{S}$, our model controls the total outflow $x(S_t) = (x_1(S_t), \dots, x_N(S_t))$ over an upcoming dry season. Hence, the state update from period t to $t+1$ for each reservoir i is described by the equation

$$S_{t+1,i} = \min\left\{S_{t,i} + \xi_{t,i} - x_i(S_t), C_i\right\}. \tag{1}$$

During each dry season, we assume a known demand D, which occurs downstream of all the reservoirs in the system and therefore can be satisfied with water coming from any of the reservoirs. Such a situation is common. For the Sacramento River basin, the demand consists of the following three main sources: (i) demand in the San Joaquin-Sacramento River Delta, including urban consumption in the San Francisco Bay Area; (ii) agricultural demand in the Central Valley, for which water is exported south via a system of pipes; and (iii) environmental demand, satisfying the need from a sensitive ecosystem (e.g., salmon populations) to provide enough fresh water to keep salinity levels low. All of these demand sources occur downstream of the reservoirs in the system. While hydroelectric energy is produced as a consequence of releases, none of the reservoirs considered are primarily for hydroelectric power generation. Hence, we do not explicitly model the corresponding energy market in this paper. However, this is not a limitation of our model or the results. If part of the demand D cannot be satisfied, shortage costs are incurred according to a convex, decreasing cost function $c : \mathbb{R}_+^N \to \mathbb{R}_+$ on outflows $x(S_t)$. The assumption that all demand occurs downstream implies that $c(x) = c(x')$ holds for any release vectors x and x' with $\sum_{i=1}^N x_i = \sum_{i=1}^N x_i'$. However, the total demand D can be split into different streams $D_1, \dots, D_M$ (such that $\sum_{i=1}^M D_i = D$), with associated shortage cost parameters reflecting their prioritization (cf. Section 5.3.3).

All quantities $S_{t,i}, x_i, C_i, e_i, \Gamma_t$ and D, are measured in acre-feet (af) or million acre-feet (maf).

2.1. Objective function

Release decisions boil down to a tradeoff between satisfying immediate demand versus storing water for future needs in case of little inflow. When addressing this tradeoff in a stochastic model, a key modeling choice is the time horizon.

There does not seem to be a natural choice for a deterministic, finite planning horizon. Indeed, any finite-horizon model will incentivize greedy decisions in the last periods, unless artificial terminal conditions are imposed, in which case the decisions will be driven by those conditions. Similarly, the decisions in an infinite-horizon model are driven by the choice of the discount factor. The steeper the discount factor, the smaller is the incentive to conserve water resources for future needs. Moreover, the concept of a time value of water does not exist in a similar manner as for money. Thus, it remains unclear which discount factor to apply.

To avoid having to choose artificial terminal conditions or an arbitrary discount factor, we adopt an alternative perspective based on stochastic cycles delimited by the event of all reservoirs being full. A cycle of length n is defined as a (sample) path from time t_0 through time $t_n = t_0 + n$, such that $S_{t_0,i} = C_i$ and $S_{t_n-1,i} + \xi_{t_n-1,i} - x_i(S_{t_n-1}) > C_i$ hold for all $i = 1,\dots,N$, and for each $t_0 < t < t_n$ there exists some reservoir i such that $S_{t,i} < C_i$. Starting from all storage levels equaling their capacities, a cycle thus ends as soon as all reservoirs $i = 1,\dots,N$ in the system are simultaneously full again. While, for technical reasons, our definition of a cycle is based on sufficient inflow to *exceed* all capacities, the cap C_i included in the state update, c.f. Equation (1), prevents the model from reaching physically impossible storage levels strictly above capacity. The resulting cycles are of stochastic length and the end of a cycle is denoted by the stopping time $\tau := \min\{t > 0 : S_{t,i} \geq C_i \ \forall i = 1,\dots,N\}$. Assuming a continuous inflow distribution ensures that τ is well defined, as the event $S_{t_n-1,i} + \xi_{t_n-1,i} - x_i(S_{t_n-1}) = C_i$ has probability zero. We introduce a virtual absorbing state $\overline{S}$ to channel all potential realizations of inflows that cause the reservoirs to simultaneously exceed their capacities. We refer to $\overline{S}$ as the *target state*, which incurs no cost and in which the system remains once it is reached. However, note that enforcing the reservoir levels equal to the capacity C after the target state $\overline{S}$ is reached, allows resetting the model so it can be interpreted as a renewal process, starting a new cycle at time τ .

Finally, our objective is to minimize the expected sum of shortage costs per cycle over all feasible release policies. For a given distribution $\mathbb{P}$ associated with the inflow variable Γ, the resulting optimization problem can be summarized as follows:

$$v(C) := \min_{x \geq 0} \ \mathbb{E}^{\mathbb{P}} \left[\sum_{t=0}^{\tau} c(x(S_t)) \middle| S_0 = C \right] \tag{2}$$

$$\text{s.t. } S_{t+1,i} = \min\{S_{t,i} + \xi_{t,i} - x_i(S_t), C_i\} \geq S_i^{\min}\ a.s.\ \forall i \leq N, \forall t \geq 0$$

2.2. Formulation as a stochastic shortest path problem

Using a stochastic time horizon model, in general, does not allow drawing on the rich toolkit that is available for both standard finite and infinite horizon models. However, it can be shown that, under mild assumptions, the model given in (2) belongs to the class of stochastic shortest path problems introduced by Bertsekas and Tsitsiklis (1989). In particular, we use the following assumption on the stationary inflow distribution:

Assumption 1. The distribution $\mathbb{P}$ of the inflow variable Γ has a continuous density function with support $(a,b) \subseteq \mathbb{R}_+$, such that $a > \max_{i=1,\ldots,N} \frac{1}{(1-\beta_i)\alpha_i} e_i$ and $b > \max_{i=1,\ldots,N} \frac{1}{(1-\beta_i)\alpha_i}(C_i - S_i^{\min})$.

The condition on a implies that there is a positive inflow (net of evaporation) during each year. The condition on b implies that, for any state of the system, there is a positive chance of enough inflow during one year, such that all reservoirs reach their capacity. Given the historical data, these conditions are both very realistic assumptions: First, evaporation losses are of a negligible magnitude compared to inflows. Second, the inflows required to hit capacity when starting at the minimum level for Shasta, Trinity, Folsom and Oroville, are given by 3.55, 1.97, 0.84, and 2.75 (maf), respectively. These inflows have been exceeded multiple times. For instance, from May 1, 2005 to April 30, 2006, the corresponding total inflows were about 9.24, 2.08, 5.32, and 7.52 (maf).

Assumption 1 guarantees that all policies are *proper* in the sense of Bertsekas and Tsitsiklis (1989):

LEMMA 1. *For any policy x that is feasible for* (2)*, it holds that* $\lim_{t\to\infty} \mathbb{P}[S_t = \overline{S}] = 1$.

Assumption 1 also ensures that state transition probabilities are continuous functions of the release decisions. Given the minimum reservoir levels $S^{\min}$ that must be respected, the action space in a given state is a compact subset of $\mathbb{R}^N$. Moreover, since the cost function $c(\cdot)$ is continuous, all conditions are satisfied for the following statement to hold:

PROPOSITION 1. *Under Assumption 1, the model given in* (2) *is a stochastic shortest path problem in the sense of Bertsekas and Tsitsiklis (1989), for which the optimal cost is the unique solution to the dynamic programming equation*

$$v(S) = \min\left\{c(x) + \mathbb{E}^{\mathbb{P}}\Big[v\big(\min\{S + \xi - x, C\}\big)\Big] \,:\, 0 \leq x \leq S - S^{\min}\right\}, \tag{3}$$

with boundary condition $v(C) = 0$, where the inner min *operator is applied component-wise.*

2.3. Incorporating climate uncertainty via robustification

As discussed in the Introduction, the distribution $\mathbb{P}$ of the annual inflow Γ cannot be estimated reliably, due to climate change. Therefore, we introduce an adversarial player ("nature"), who selects a particularly unfavorable distribution for the state transitions, in response to any release decision taken. We start from a nominal distribution $\mathbb{P}$ (a "best guess" probabilistic estimate based on the available data), but allow nature to modify the density function $p(\cdot)$ in return for paying a penalty. Let $\mathbb{Q}_t = \mathbb{Q}(S_t, x(S_t))$ denote the distribution chosen by nature at time t, after observing the state S_t and the release decision $x(S_t)$, with an absolutely continuous density $q(\cdot)$ with respect to $p(\cdot)$, i.e., $\mathbb{Q}_t \ll \mathbb{P}$. The more the distribution $\mathbb{Q}_t$ deviates from $\mathbb{P}$, the more ambiguity in the distribution of Γ, and the larger the penalty to be paid. To measure the deviation of $\mathbb{Q}_t$ from $\mathbb{P}$, we use relative entropy, also known as Kullback-Leibler (KL) divergence, defined as $D_{\mathrm{KL}}(\mathbb{Q}_t \| \mathbb{P}) := \int q(y) \log(q(y)/p(y))\, dy$. While a variety of alternative distance concepts between probability distributions are available (including moment-based distances, the Hellinger distance, Wasserstein distance, etc.), KL-divergence has several favorable properties (see for instance the discussion in Hu and Hong 2013), and is therefore widely applied in the distributionally robust control literature (Esfahani and Kuhn 2018). Our ambiguity approach is similar to Kim and Lim (2016) who used it to study a robust multi-armed bandit problem.

Starting from the nominal model given in (2), we consider the following robustified version:

$$v_\theta(C) = \min_{x \geq 0} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=0}^{\tau} c(x(S_t)) - \theta \cdot D_{\mathrm{KL}}(\mathbb{Q}_t \| \mathbb{P}) \middle| S_0 = C \right] \tag{4}$$
$$\text{s.t. } S_{t+1,i} = \min\{S_{t,i} + \xi_{t,i} - x_i(S_t), C_i\} \geq S_i^{\min} \; a.s. \; \forall i \leq N, \forall t \geq 0,$$

where $\theta > 0$ penalizes the ambiguity measured by $D_{\mathrm{KL}}(\mathbb{Q}_t \| \mathbb{P})$, and $\mathcal{Q} = \{\mathbb{Q}_t = \mathbb{Q}(S_t, x(S_t)) : S_t \in \mathcal{S}, \mathbb{Q}_t \ll \mathbb{P}, t \geq 0\}$. The smaller the value of θ, the less confidence there is in the nominal model and the more freedom is thus given to nature. On the other hand, a larger θ tells nature not to deviate too much from a well-trusted nominal model. Hence, we refer to θ as the confidence parameter.[1]

Similar to Kim and Lim (2016), the robust model in Equation (4) can be seen as a zero-sum two-player sequential stochastic game between nature and the decision maker, with nature having perfect information about the decision maker's action when determining its adverse response. This strongly adversarial version of nature considers the worst-case distribution for each candidate

[1] The literature refers to θ as the ambiguity parameter, but calling it the confidence parameter is more intuitive in our context.

release decision and thereby induces particularly robust policies. The formulation in (4) can be simplified taking advantage of the Donsker and Varadhan variational formula (see Donsker and Varadhan 1983), as shown in the following proposition.

PROPOSITION 2. *The robust model given in* (4) *describes a stochastic shortest path problem. The optimal cost is the unique solution of the stochastic dynamic programming equation*

$$v_\theta(S) = \min_{0\le x\le S-S^{\min}} c(x) + \theta \log\left(\mathbb{E}^{\mathbb{P}}\left[e^{\frac{v_\theta(\min\{S+\xi-x,C\})}{\theta}}\right]\right). \tag{5}$$

The optimal value function can thus be computed by a standard value iteration algorithm, which is guaranteed to converge to a unique, finite value. Notice that the formulation (5) is based on the nominal distribution $\mathbb{P}$, but involves a nonlinear transformation of the cost-to-go function.

We note that from a decision science perspective, setting $\theta = -1/\eta$ allows interpreting Equation (5) as a risk-averse decision problem of a decision maker who maximizes expected exponential utility. In the finance literature, the quantity $\rho(X) := \frac{1}{\eta}\log(\mathbb{E}[e^{-\eta X}])$ is referred to as the entropic risk measure (for a risky position X). While mathematically equivalent, we emphasize the fundamental conceptual difference between considering (5) as a risk-averse decision problem under uncertainty, which assumes a known probabilistic model of inherent randomness, and considering (5) as a distributionally robust decision problem, which acknowledges the unavailability of a reliable probabilistic uncertainty model and seeks robustness across a range of possible distributions. This distinction is especially relevant when it comes to estimating the parameter θ. For our robust approach, θ is not chosen to reflect any subjective risk preferences within a given uncertainty model, but rather will be tuned to optimize out-of-sample performance. This allows to infer robust decisions from our model, when facing ambiguous inflow distributions.

3. Model Analysis

In this section, we analyze the model presented above. We derive structural results for the special case of a single reservoir and obtain policy insights for the general case of a system of multiple reservoirs. Unless otherwise stated, all assumptions made in Section 2 are presumed. The shorthand notation $\sum_i$ is sometimes used for the sum over all reservoirs $i = 1, \ldots, N$.

3.1. Managing a single reservoir

While the general case of a system of $N \ge 1$ reservoirs is the focus of this paper, the case $N = 1$ of a single reservoir is an important special case: First, it has been well-studied in the management literature that the implementation of centralized versus non-centralized decision-making processes in organizations is a multi-faceted question (see, e.g., the classical book of Mintzberg 1989, Chapter 11). For various organizational reasons, even reservoirs that are located in the same river basin might be managed in an uncoordinated way. Jeuland et al. (2014) analyze the cost of such non-cooperation in water resources systems. Second, the single reservoir problem can easily be solved computationally by straightforward value iteration, as it does not suffer from the curse of dimensionality. Exploiting this fact, in Section 4 we will apply aggregation methods that rely on the single reservoir model to solve the multiple reservoirs case. Given its relevance both from a managerial and a computational perspective, we begin the analysis of our model by examining basic properties in the single reservoir case. To simplify the notation, we drop the i index. Our first result characterizes the optimal value function.

PROPOSITION 3. *The optimal value function $v_\theta(S)$ is convex and decreasing in $S \in [S^{\min}, C]$.*

The fact that the optimal value function is decreasing in our model reflects the intuitive property that a larger amount of stored water is associated with a smaller sum of expected shortage costs until the end of a cycle. The majority of our numerical experiments in Section 5 will be based on a single demand stream D together with a convex (piece-wise linear) cost function $c(x) := \kappa \cdot (D - \sum_{i=1}^{N} x_i)^+$, where $(\cdot)^+$ denotes the positive-part function. For a single reservoir, the convexity property from Proposition 3 allows a full characterization of the optimal policy:

PROPOSITION 4. *Let $c(x) := \kappa \cdot (D - x)^+$. There is a state $S^* \in [S^{\min}, C]$ such that for the optimal policy x^* it holds $x^*(S) = 0$ for all $S \le S^*$, and $x^*(S) = \min\{S - S^*, D\}$ for all $S > S^*$.*

Given a piece-wise linear cost function, Proposition 4 shows that the optimal policy in our model can be interpreted as a "release-down-to" policy with a target level S^*, where the release amount is capped by the demand. This interpretation reflects a mirror image of "order-up-to" (base stock) policies, which represent the most famous policy type in the inventory management literature.

The following proposition shows that the optimal value function in our model also exhibits a natural behavior with respect to the confidence parameter θ: the more confident we are in the nominal uncertainty model (i.e., the larger θ), the smaller the expected shortage costs.

PROPOSITION 5. *Given the state S of the reservoir, the optimal value function $v_\theta(S)$ is a decreasing function with respect to the confidence parameter θ.*

3.2. Managing a system of identical reservoirs

California's reservoirs vary considerably in terms of capacities, water shed sizes, etc. We will study such systems in Section 3.3 below. Before doing so, we build intuition—omitting separate proofs; all results follow from Section 3.3—by considering the much simpler special case in which all reservoirs in the system are identical copies of each other.

Suppose $C_i = C_j, \alpha_i = \alpha_j, \beta_i = \beta_j, e_i = e_j$ and $S_i^{\min} = S_j^{\min}$ for all $i, j = 1, \ldots, N$. We call a (state of the) system *balanced*, if all reservoir levels coincide. If the system is balanced, then it is (only) optimal to release an equal amount from each reservoir. This will keep the system balanced until the terminal state is reached. If the system is not balanced, then it is optimal to first release water only from the reservoir with the highest level, until it matches the reservoir that had the second highest level. Then, an equal amount is released from these two reservoirs, until both match the reservoir that had the third highest level, and so on. When all reservoir levels eventually match the one that had the lowest initial level, a balanced state is reached. Therefore, every optimally chosen release decision brings the system closer to a balanced state. The total release amount from the system is equivalent to the optimal amount released from a single (aggregate) reservoir with capacity $N \cdot C_i$ and inflow $N \cdot \xi_{t,i}$, for any i. Hence, from a managerial perspective, a system of identical reservoirs essentially collapses to a single reservoir, which we studied in Section 3.1.

3.3. Managing a system of non-identical reservoirs

We now turn to the general case of a system of $N > 1$ non-identical reservoirs that collectively satisfy downstream demand. The additional managerial complexity then stems from deciding not only how much demand to satisfy (as in Section 3.1), but also how much water to release from each reservoir.

We first need the following definitions. Let $\delta_i(S_i) := \frac{1}{(1-\beta_i)\alpha_i}(C_i - (S_i - e_i))$ be the inflow that would bring reservoir i to its capacity C_i from the current state S_i. Hence, when $\delta_i(S_i)$ is lower it means that reservoir i is closer to full, i.e., it has more water. Next, we introduce a state classification that generalizes the notion of a balanced state.

DEFINITION 1. A state S of a system of N reservoirs is called

- balanced, if $\delta_i(S_i) = \delta_j(S_j)$ holds for all $i, j = 1, \ldots, N$;
- weakly-balanced, if for a subset $\mathcal{I} \subseteq \{1, \ldots, N\}$ the subsystem $(S_i : i \in \mathcal{I})$ is balanced ($\delta^{\mathcal{I}} := \delta_i(S_i) = \delta_j(S_j)$ for all $i, j \in \mathcal{I}$), and $S_k = S_k^{\min}$ with $\delta_k(S_k) < \delta^{\mathcal{I}}$ holds for all $k \notin \mathcal{I}$.
- unbalanced, if it is neither balanced nor weakly-balanced.

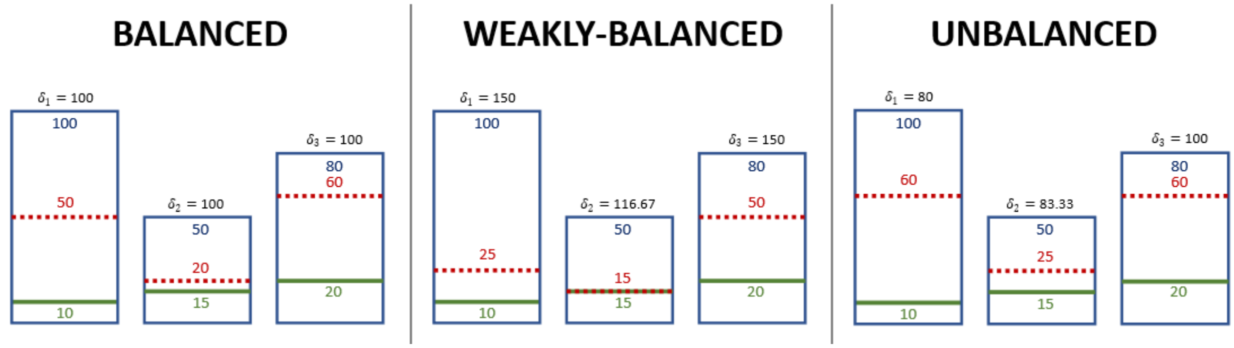


Figure 2 Illustration of Definition 1 for a system of three reservoirs with (absolute) capacities $C_1 = 100, C_2 = 50, C_3 = 80$, minimum levels $S_1^{\min} = 10, S_2^{\min} = 15, S_3^{\min} = 20$, and inflow constants $(1-\beta_1)\alpha_1 = 0.5, (1-\beta_2)\alpha_2 = 0.3, (1-\beta_3)\alpha_3 = 0.2$. For simplicity, evaporation losses are neglected here by setting $e_1 = e_2 = e_3 = 0$. States are indicated by the dotted red lines, the exact levels are written above.

Figure 2 illustrates this definition, which is crucial for the remainder of the paper. The state $S^{(1)} = (50, 20, 60)$ shown in the left plot is balanced, because for a total inflow of $\Gamma = \delta_1(S_1^{(1)}) = \delta_2(S_2^{(1)}) = \delta_3(S_3^{(1)}) = 100$ all reservoirs simultaneously hit capacity. The state $S^{(2)} = (25, 15, 50)$ in the middle is weakly-balanced, because the subsystem $(S_1^{(2)}, S_3^{(2)})$ is balanced $(\delta_1(S_1^{(2)}) = \delta_3(S_3^{(2)}) = 150 > 116.67 = \delta_2(S_2^{(2)}))$ and $S_2^{(2)} = S_2^{\min}$. The state $S^{(3)} = (60, 25, 60)$ in the right plot is unbalanced, because $\delta_1(S_1^{(3)}) \neq \delta_2(S_2^{(3)}) \neq \delta_3(S_3^{(3)})$. The state $S^{(3)}$ would be balanced for alternative inflow constants given by $(1-\beta_1)\alpha_1 = 0.5, (1-\beta_2)\alpha_2 = 0.25, (1-\beta_3)\alpha_3 = 5/16$.

Definition 1 also induces a partial ordering of states in the following sense:

DEFINITION 2. Consider two states $S^{(1)}$ and $S^{(2)}$ of a system of N reservoirs. Then $S^{(1)}$ is called *better balanced* than $S^{(2)}$, if $\max_i \delta_i(S_i^{(1)}) - \min_i \delta_i(S_i^{(1)}) < \max_i \delta_i(S_i^{(2)}) - \min_i \delta_i(S_i^{(2)})$.

The fact that this definition is exclusively based on the states of the two most extreme reservoir levels (taking all specifics into account) is related to the fact that a system is no closer to reaching the target state whether or not "intermediate" reservoir levels are balanced. Notice that for a balanced state $S^{(1)}$, the left hand side in Definition 2 equals zero. Hence, any balanced state is better balanced than any weakly-balanced or unbalanced state (regardless of the total amount of water in the system). Any weakly-balanced state $S^{(2)}$ is better balanced than any unbalanced state $S^{(3)}$ with $\sum_i S_i^{(3)} = \sum_i S_i^{(2)}$. If $\sum_i S_i^{(3)} > \sum_i S_i^{(2)}$, this relation cannot be guaranteed, e.g., consider $S^{(3)} = (49, 20, 61)$ vs. $S^{(2)}$ in the middle plot of Figure 2 as a counterexample.

Our goal now is to characterize the optimal release policy. The following proposition helps build some intuition.

PROPOSITION 6. *If for two state vectors $S^{(1)}$ and $S^{(2)}$ it holds almost everywhere that*

$$\sum_{i=1}^{N} \min\left\{C_i, S_i^{(1)} + \xi_i\right\} \geq \sum_{i=1}^{N} \min\left\{C_i, S_i^{(2)} + \xi_i\right\}, \tag{6}$$

then $v_\theta(S^{(1)}) \leq v_\theta(S^{(2)})$. If strict inequality holds in (6) with positive probability, then $v_\theta(S^{(1)}) < v_\theta(S^{(2)})$.

The key insight of Proposition 6 is that the value of a given (pre- or post-decision) state $S^{(1)}$, i.e., its associated expected costs, is smaller than the value of another state $S^{(2)}$, if the system in state $S^{(1)}$ can hold more water than the system in state $S^{(2)}$, irrespective of the realized inflow over the next year. A direct corollary of Proposition 6 shows the desirability of reaching a balanced state, or if that is not possible, then a weakly balanced state.

COROLLARY 1. *Let the state S of a system be given, and let $v_\theta(x,S) := c(x) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_\theta(\min\{C, S+\xi-x\})}{\theta}}\right]\right)$. Consider two feasible release decision vectors x and x^* with $\sum_{i=1}^N x_i = \sum_{i=1}^N x_i^*$. If (i) $S - x^*$ is balanced, or (ii) $S - x^*$ is weakly balanced and $S - x$ is unbalanced, then $v_\theta(x^*, S) \leq v_\theta(x, S)$.*

Corollary 1 implies that for any system of reservoirs, a release decision that turns them (weakly) balanced is better than any other decision releasing the same total amount. Section EC.1.1 in the appendix provides a few additional preliminary results. In particular, Lemma EC.2 shows that if a system is balanced at the beginning of a dry season, then a release decision vector x^* preserving this property is conditionally optimal, given the total amount of water to be released. The following proposition is our main result and it characterizes an optimal decision vector as a direct consequence of these insights.

PROPOSITION 7. *Consider a system of N reservoirs in state S and let the optimal total release X^* for S be given. Suppose x^* is feasible and satisfies the following conditions:*

(i) $\sum_{i=1}^N x_i^* = X^*$;

(ii) there is an index set $\mathcal{I} \subseteq \{1,\dots,N\}$ with $\delta_i(S_i) < \delta_j(S_j)$ for all $i \in \mathcal{I}, j \notin \mathcal{I}$, such that

- $x_j^* = 0$ *for all* $j \notin \mathcal{I}$
- $\{S_i - x_i^* : i \in \mathcal{I}\}$ *is balanced, or* $\{S_i - x_i^* : i \in \mathcal{I}\}$ *is weakly balanced and there is no* $x \geq 0$ *with* $\sum_{i=1}^N x_i = X^*$ *and* $x_j = 0$ *for all* $j \notin \mathcal{I}$ *such that* $\{S_i - x_i : i \in \mathcal{I}\}$ *is balanced;*

(iii) there is no other vector x' satisfying (i) and (ii) for an index set $\mathcal{I} \subsetneq \mathcal{I}'$;

Then x^* is an optimal release vector.

Intuitively, Proposition 7 shows that an optimal decision is guided by the principle of splitting the total release among the individual reservoirs in such a way that as little water as possible will have to be spilled (i.e., wasted) over all inflow scenarios. This is ultimately achieved by reaching a balanced post-decision state. Part (ii) of Proposition 7 also indicates that it is always optimal to start releasing water from those reservoirs that are closer to full, i.e., with a lower $\delta_i(S_i)$.

To illustrate the proposition, consider the pre-decision state $S^{(3)} = (60, 25, 60)$ given in the right plot of Figure 2, from which a total amount of $X^* = 30$ shall be released. Then, by Proposition 7, the split $x^* = (17.5, 9.5, 3)$ is optimal, as the post-decision state $S^{(3)} - x^*$ is balanced ($\delta_1(60 - 17.5) = \delta_2(25 - 9.5) = \delta_3(60 - 3) = 115$). However, in case of little total release, an unbalanced pre-decision state can also result in a better balanced but still unbalanced post-decision state. For instance, if $X^* = 10$ in the above example, then $x^* = (6.875, 3.125, 0)$ is optimal according to Proposition 7 ($\mathcal{I} = \{1, 2\}$, $\{S_i^{(3)} - x_i^* : i \in \mathcal{I}\}$ is balanced), even though the post-decision state $S^{(3)} - x^*$ is unbalanced. On the other hand, even for large inflow Γ, reaching only a weakly-balanced state can be optimal if a balanced post-decision state cannot be achieved due to the different minimum levels. Continuing with the same example, splitting $X^* = 31.67$ as $x^* = (55/3, 10, 10/3)$ is optimal and results in a balanced state with $S_2^{(3)} - x_2^* = S_2^{\min}$. Hence, for any $X^* > 31.67$ a feasible balanced post-decision state cannot be reached. If, for example, $X^* = 40$, then the split $x^* = (170/7, 10, 40/7)$ is optimal according to Proposition 7 (resulting in a weakly-balanced post-decision state).

After releases have been made, a system in a balanced post-decision state will remain balanced (up to evaporation losses) as inflow occurs. If in a weakly-balanced post-decision state, inflow either results in reaching the terminal state, or in reaching an unbalanced state with a subset of reservoirs remaining balanced (up to evaporation losses) while others are no longer at their minimum. For instance, let the post-decision state be $S^{(2)}$ given in the middle plot of Figure 2 and let $\Gamma = 60$. Then, the resulting state $S^{(2)} + \xi = (55, 33, 62)$ is unbalanced, with the first and third reservoir remaining balanced, but the second one no longer at its minimum. Finally, if at least one reservoir reaches its capacity, inflow into an unbalanced system can make it better balanced. For example, consider $S^{(3)}$ as the post-decision state and let $\Gamma = 90$. Then, the first and second reservoir reach their capacities, but the third one does not. Nevertheless, the system becomes better balanced than before, as $\max_i \delta_i - \min_i \delta_i$ is reduced from 20 to 10. These concepts will be key in developing our bounds and heuristic for problem (5), which is described next.

4. Computational Solution

In principle, the dynamic programming formulation (5) of our model allows applying value iteration to determine the optimal value function. However, even for a system of only a few reservoirs this is computationally infeasible due to the number of states growing exponentially with respect to the number of reservoirs in the system.

We now deal with this curse of dimensionality in three steps. First, we use an aggregation approach to establish a lower bound for the optimal costs. Combining the optimal releases for the resulting aggregate model with the policy insights from Section 3.3, we then propose a heuristic release policy for the multiple reservoirs case. In a third step, we discuss the policy evaluation problem, which requires the development of an upper bound on the (worst-case) expected costs associated with the heuristic.

4.1. A lower bound for the optimal expected costs

Consider a system of two reservoirs, a larger one and a smaller one. Suppose the smaller one gets a large amount of inflow, which exceeds its storage capacity, but the larger one does not get a sufficient amount to reach its capacity. Then, the system does not reach its target state and a substantial amount of valuable inflow cannot be stored for the future, but must be spilled from the smaller reservoir. If, instead, the system consisted of a single reservoir, which has the aggregate capacity and receives the aggregate inflow of the two individual reservoirs, then no inflow would need to be wasted. The aggregate reservoir is closer to its terminal state than the corresponding two reservoirs system and has more water to satisfy future demand. Due to this pooling effect, the expected cycle costs associated with the aggregate reservoir would be smaller. The following result formalizes this intuition for a general system:

PROPOSITION 8. *Consider a system of N reservoirs, as modeled in Section 2. Construct a (virtual) system consisting of a single reservoir, with aggregate capacity $C^{(1)} = \sum_i C_i$, minimum level $S^{(1)}_{\min} = \sum_i S_i^{\min}$, and uncontrolled net inflow $\xi^{(1)} = \sum_i \xi_i$. The same (downstream shortage) cost function $c(X)$, where X denotes the total upstream release, applies to both systems. Denote by $v_\theta^{(1)}$ and $v_\theta^{(N)}$ the optimal value functions corresponding to the aggregate reservoir and the N-reservoirs system, respectively. Then, it holds for any state $S^{(N)}$ of the system of N reservoirs, that $v_\theta^{(1)}\left(\sum_i S_i^{(N)}\right) \leq v_\theta^{(N)}\left(S^{(N)}\right)$. If $S^{(N)}$ is unbalanced, strict inequality holds.*

Proposition 8 shows that properly aggregating a system of reservoirs to a single, virtual reservoir allows to determine a lower bound for the optimal expected cycle costs of the system. The aggregated problem does not suffer from the curse of dimensionality because it is a single reservoir, c.f. Section 3.1, so it is computationally feasible to compute the lower bound by standard value iteration.

When, relative to their capacities, some reservoirs in the system receive substantially more inflow than others, then the amount of total usable inflow into the system (i.e., inflow that does not need to be spilled immediately) can be much larger for the virtual aggregated reservoir. This can lead to the bound of Proposition 8 being loose. However, the constructed overly generous inflow into the aggregate reservoir can be restricted, while still entailing a lower bound. See EC.1.2 in the appendix for details on the construction of a tighter lower bound following this idea.

4.2. Release Heuristic

Our proposed policy $x(\cdot)$ involves two steps. First, the total amount to be released from the system, $X(S^{(N)})$, is determined. Second, the amount $x_i(S^{(N)}) \geq 0$ released from each reservoir i is decided such that $\sum_i x_i(S^{(N)}) = X(S^{(N)})$.

To determine $X(S^{(N)})$, we use the aggregation approach of Section 4.1, incorporating the refinement of EC.1.2, and set $X(S^{(N)}) = X^{(1)}(\sum_i S_i^{(N)})$. In general, $X^{(1)}(\sum_i S_i^{(N)})$ does not coincide with the optimal total release $X^*(S^{(N)})$, which makes this policy a heuristic. Once $X^{(1)}(\sum_i S_i^{(N)})$ has been determined, it can be split among the individual reservoirs in an *optimal* way, based on Proposition 7. Specifically, the split $x(\cdot)$ aims for a balanced post-decision state, which requires sufficient total release $X^{(1)}(\sum_i S_i^{(N)})$ and feasibility with respect to the minimum levels. If a balanced state is not feasible, a weakly-balanced post-decision state is constructed. In case of little total release, at least a subsystem of reservoirs ends up in a (weakly-)balanced post-decision state.

As an example, consider the system illustrated in the left plot of Figure 2. Suppose a total amount of $X = 40$ shall be released from the system in state $S^{(1)} = (50, 20, 60)$. To keep the system balanced, X would need to be split into $x_1^* = 20, x_2^* = 12$, and $x_3^* = 8$. However, this is infeasible due to $S_2^{\min} = 15$. Therefore, the optimal split can be constructed by setting $x_2^* = 5$ first (such that $S_2^{(1)} - x_2^* = S_2^{\min}$) and then splitting the remaining release amount of 35 in such a way that the subsystem $(S_1^{(1)} - x_1^*, S_3^{(1)} - x_3^*)$ is balanced. This is achieved by setting $x_1^* = 25$ and $x_3^* = 10$. The resulting state is the weakly-balanced state $S^{(2)}$ illustrated in the middle plot of Figure 2.

4.3. Policy Evaluation

For standard stochastic models, straightforward Monte-Carlo simulation is typically used to estimate the expected costs associated with a given policy. For our distributionally robust model, however, this would require to simulate from the worst-case distributions $\mathbb{Q}_t^*$, which are not available in closed form. The policy evaluation problem, given by

$$\overline{v}_\theta(x, C) := \max_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\left[\sum_{t=0}^{\tau} c(x(S_t)) - \theta \cdot D_{\mathrm{KL}}(\mathbb{Q}_t \| \mathbb{P}) \middle| S_0 = C\right], \tag{7}$$

therefore requires the construction of a computable upper bound on the worst-case expected costs, where the worst-case refers to the inflow distribution.

Constructing such an upper bound for a maximization problem over probability distributions is challenging. Our approach exploits the structure of the heuristic policy. The idea in developing this upper bound is similar to the one used in EC.1.2. However, conversely to the *most favorable* state in the case of a lower bound, our upper bound is based on the *least favorable* state for a given total amount of water $S^{(1)}$ stored in the system. Constructing the least favorable state requires constructing the most unbalanced state that can be reached. Considering the structure of our release policy, when starting at the beginning of a cycle, the system typically remains in a balanced state (up to evaporation losses) for some time. An unbalanced state can then be reached (for the first time) only as a consequence of inflow, after releases have led to a weakly-balanced state right before. Subsequent releases and inflows will then make the system better balanced. Hence, given $S^{(1)}$, the corresponding most unbalanced state S with $\sum_i S_i = S^{(1)}$ would be given by a state whose preceding post-decision state was all reservoirs being at their minimum levels. If such a state S is not feasible (due to the different capacities), it can be adjusted accordingly. Section EC.1.3 details the construction.

PROPOSITION 9. *Consider the setting of Proposition 8, but let the state update for the aggregated problem be given by* $S_{t+1}^{(1)} = \min\{C^{(1)}, S_t^{(1)} + \sum_i \min\{\xi_{t,i}, C_i - \check{S}_i^{(N)}(S_t^{(1)})\} - X(S_t^{(1)})\}$, *where* $\check{S}_i^{(N)}(S_t^{(1)})$ *denotes the* i*-th component of the least favorable state* $\check{S}^{(N)}$ *with* $\sum_i \check{S}_i^{(N)} = S^{(1)}$. *Fix the policy* $x(\cdot)$ *for the* N*-reservoirs setting as the one suggested in Section 4.2, and the corresponding policy* $X(\cdot)$ *for the aggregated problem. Then,* $\overline{v}_\theta(x, C) \leq \check{v}_\theta^{(1)}(X, C)$, *where* $\check{v}_\theta^{(1)}(X, C)$ *is part of the unique solution of the fix point equations (for all* $S^{(1)} = s$*)*

$$\check{v}_\theta^{(1)}(X, s) = c(X(s)) + \theta \log\left(\mathbb{E}\left[e^{\frac{\check{v}_\theta^{(1)}(X, \min\{C^{(1)}, s + \sum_i \min\{\xi_{t,i}, C_i - \check{S}_i^{(N)}(s)\} - X(s)\})}{\theta}}\right]\right). \tag{8}$$

We also developed an alternative upper bound based on information relaxation which is not limited to the policy suggested in Section 4.2, but can be applied to any feasible policy. The details are provided in EC.1.4. However, in the numerical study discussed next, the tailored upper bound in Proposition 9 turned out to be tighter.

5. Case Study: Sacramento River Basin

We now apply the above model to the case of the Sacramento River Basin, the largest water shed in California. This basin consists of a complex system of natural rivers and man-made or regulated reservoirs. The largest reservoirs, in order of size, are Shasta Lake, Lake Oroville, Trinity Lake, and Folsom Lake. Representing the Northern California System of major reservoirs by these four facilities is in line with Georgakakos and Georgakakos (2007). The reservoir outflows controlled by dams eventually contribute to the Sacramento River, and they are ultimately responsible for the river flow. Shasta Dam, Trinity Dam and Folsom Dam are part of the federal Central Valley Project and are operated by the U.S. Bureau of Reclamation (USBR), while Oroville Dam is part of the California State Water Project and operated by the California Department of Water Resources. Hence, we restrict the system that we study to Shasta, Trinity, and Folsom. The details of this study are based on conversations with a USBR representative (White 2022).

All data that we use are publicly available through the website of the California Data Exchange Center, managed by the California Department of Water Resources. We have queried daily data for the period from May 1, 2000 to April 30, 2024.

The capacities of Shasta, Trinity, and Folsom Lake are $C_1 = 4{,}552{,}000$, $C_2 = 2{,}447{,}650$, and $C_3 = 976{,}000$ acre-feet (af), respectively. The minimum levels imposed have varied significantly over time, according to White (2022). We use values of $S_1^{\min} = 1{,}002{,}000$, $S_2^{\min} = 478{,}423$, and $S_3^{\min} = 136{,}980$ (af), corresponding to the lowest levels observed in the data. These values also seem roughly compatible with the estimated total deadpool (i.e., volume of inaccessible water) in California's reservoirs of 7 million acre-feet (maf), as used in Stanton and Fitzgerald (2011).

To model the nominal inflow uncertainty, we let Γ denote the inflow into Shasta Lake (from May 1 to April 30 of the following year). We estimate from the data the inflow constants $\alpha_1 = 1.0, \alpha_2 = 0.2259$, and $\alpha_3 = 0.4867$ for Shasta, Trinity, and Folsom, respectively. Section EC.1.6 details the estimation of α_i. Shasta Lake serves as the reference reservoir due to its size and the inflow correlations with Trinity Lake and Folsom Lake, which are both larger than the correlation between Trinity and Folsom (see Figure 1). Notice that Trinity Lake has a relatively large capacity compared to its catchment area, receiving only about 22% of Shasta's inflow and less than 50% of the inflow into the much smaller Folsom Lake. Our estimate of the distribution $\mathbb{P}$ of Γ is based on the gamma distribution, which is widely used in the hydrology literature as a probabilistic precipitation model (see, e.g., the discussion in Martinez-Villalobos and Neelin 2019 or Murata et al. 2020). In particular, we fit a 2-gamma-mixture model to avoid oversmoothing and account for the chance of very wet years in a realistic manner. Figure 3 illustrates the fitted density function for our nominal uncertainty model $\mathbb{P}$ alongside historical inflow data for Shasta Lake.

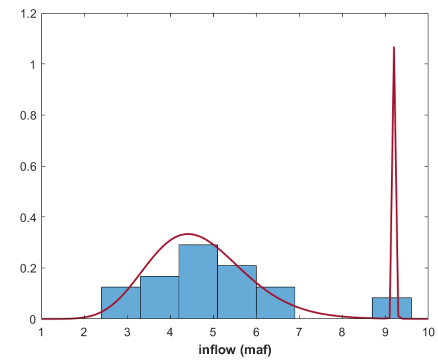


Figure 3 Density function of a 2-gamma-mixture model (red line), fitted to historical yearly (May to April in the following year) inflow data into Shasta Lake, from the (starting) year 2000 to 2023. The inflow data are illustrated by the blue histogram (with normalized counts to reflect probabilities).

To estimate the evaporation losses e_i, we use averages over yearly-aggregated data, resulting in $e_1 = 91{,}335$, $e_2 = 38{,}583$, and $e_3 = 33{,}049$ (af). The outflow constants β_i, which represent the percentage of inflow released over the wet season, are estimated as $\beta_1 = 0.4438, \beta_2 = 0.3489$, and $\beta_3 = 0.5105$, following the procedure outlined in EC.1.7.

For the unit shortage cost parameter, we choose a value of $\kappa = 800$ ($/af), based on Park and Bayraksan (2023), where a liquid market supplying water at a constant exogenous price of 800 $/af is assumed. This value is also reasonable for the present application, considering that the average Nasdaq Veles California Water Index (NQH2O) price over the last three years in the data (05/01/2021 – 04/30/2024) was slightly above 700 ($/af). The NQH20 index tracks the spot water price in California, priced at the source excluding any conveyance costs and losses. A surcharge is appropriate when estimating the unit shortage cost at the location where demand occurs.

Finally, we specify the total demand during the dry season. We consider a single demand stream and use the average outflow over all dry seasons in the data as a proxy for the demand, resulting in a baseline value of $D = 4{,}481{,}756$ (af) with $c(x) := \kappa \cdot (D - \sum_{i=1}^{N} x_i)^+$. In Section 5.3 we analyze the sensitivity with respect to D and we consider the case with two demand streams.

Across all our experiments, we use a discretization of 100 points for the feasible storage levels of Shasta Lake (from $S_1^{\min}$ to C_1) and apply the resulting stepsize to discretize the storage levels of all other reservoirs. State updates are then rounded to this grid.

5.1. Calibration of the confidence parameter θ

For practical application, the value of the confidence parameter θ needs to be fixed before our robust policy can be implemented. The motivation for adopting a distributionally robust optimization framework is a given mistrust in the nominal uncertainty model $\mathbb{P}$ to accurately reflect the true uncertainty. Given the prohibitive costs and time required to tune θ in the field, it is practical to tune it in a simulator to ensure good out-of-sample performance of our robust policy. To achieve this, we follow a common approach in the literature (see, e.g., Kim and Lim 2016, Bertsimas et al. 2018, Esfahani and Kuhn 2018) by adapting a form of K-fold cross-validation from the statistical learning literature (see the book of Hastie et al. 2009) to our prescriptive setting.

To generate a sufficiently large number of (hypothetical) yearly inflow data, we first aggregate the daily inflow data on a monthly basis and label the resulting monthly data as either wet or dry. Bootstrapping monthly data instead of daily data allows us to account for the serial correlation observed in daily data (i.e., a rainy day is more likely to follow a rainy day). For the labeling, we rely on the official water year classification provided by the California Department of Water Resources, summarizing “Wet” and “Above normal” as wet, and “Below normal”, “Dry” and “Critical” as dry. The data then correspond to roughly one-third wet years and two-thirds dry years. This 1:2 ratio is also reflected in our generated yearly inflows. To generate a wet year, we sample the inflow for, say, January only from those January months in the data labeled as wet (each month contained in a wet water year is labeled as wet). Summing up monthly inflows from May to the following April, each sampled in this way, then gives a simulated wet year’s inflow. This bootstrapping approach allows generating an arbitrary number of hypothetical yearly inflows.

The details of the cross-validation algorithm are provided in EC.1.5. The algorithm requires a few hyperparameters, and we choose values ($K = 20, n_K = 30$ and $L = 1000$) to obtain numerically stable results. The output for candidate values of $\theta \in \{1000, 2000, 3000, 4000, 5000, 6000\}$ (and a fine grid for $\theta < 1000$) is shown in Figure 4.

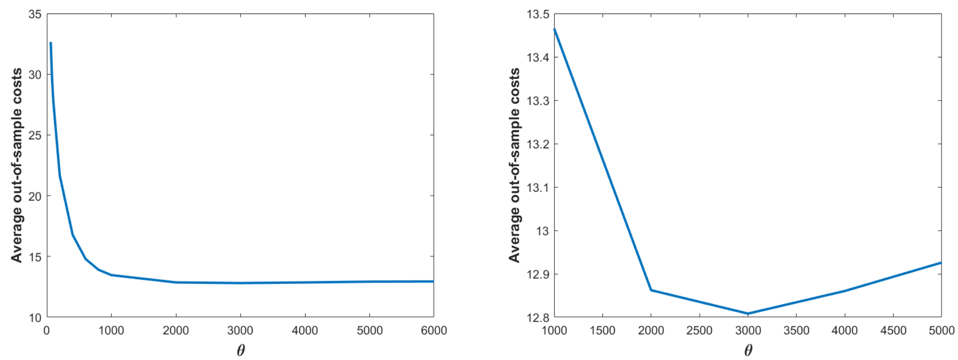


Figure 4 Results of the calibration algorithm for the confidence parameter θ. Left: Average out-of-sample costs (in \$100M) for a wide range of θ values. Right: Zoomed in to the range $\theta \in [1000, 5000]$.

The left plot in Figure 4 shows that average costs increase sharply as θ decreases below a value of $\theta = 1000$. For $\theta \geq 6000$, the simulation error would exceed the observed effect of increasing θ. Small levels of θ lead to overly conservative policies, which satisfy only a small portion of the demand. While this results in short cycles, the high costs within the cycles dominate, causing poor performance. Conversely, for large θ values the robust model approximates the nominal or ambiguity-neglecting model that fully relies on the nominal inflow distribution $\mathbb{P}$. The right plot in Figure 4 shows that $\theta^* = 3000$ minimizes the out-of-sample cost.

The confidence parameter θ serves to balance (expected) shortage costs, expressed in dollars, and the deviation of the worst-case inflow distribution from the nominal model, as measured by KL-divergence. If KL-divergence is based on the natural logarithm, the appropriate unit of measurement is the natural unit of information (symbol: nat), according to the International System of Quantities defined by the ISO/IEC 80000–13:2008 standard (see, e.g., the discussion in Stratonovich 2020, p. 4). Thus, for each nat of divergence from the nominal inflow distribution $\mathbb{P}$, a penalty of \$$\theta$ is incurred in our model. Notice that the calibrated θ value depends on the shortage cost function (and, implicitly, the demand), reflecting that it plays a regularization role.

5.2. Benchmark study

Recall from Section 2.1 that a cycle is defined as a pair of years in which all reservoirs are considered full on May 1, without this event occurring in any intermediate year. Based on this definition, the last completed cycle in the data spans the period from 2006 to 2019: on May 1, 2006, Shasta, Trinity, and Folsom were at 98%, 99%, and 95% of their capacities, respectively. These levels were reached or exceeded again (within a tolerance of one percentage point) for the first time on May 1, 2019 (98%, 98%, 96%). We use this period to back-test our model and policy.

We assume that all reservoirs are full at the beginning of the cycle. Then, we calculate the changes in storage levels from the real data (reflecting yearly inflows, outflows, and evaporation losses) for this 13-year period, but replace the observed releases during the dry seasons with the releases prescribed by our policy proposed in Section 4.2. This allows us to benchmark the hypothetical performance of our robust policy ($x^{(\theta)}$) against the policy that has been applied in practice (henceforth referred to as the "current policy" $x^{(\text{curr})}$). For our robust policy, we fix $\theta = 3000$ ($/nat) based on the above calibration study. To analyze the gain from robustness, we also consider the policy $x^{(\mathbb{P})}$ that neglects ambiguity by only considering the shortage cost in our model. As additional benchmarks, we include the most conservative policy $x^{(0)} \equiv 0$ ("no release"), and the most myopic policy $x^{(D)}$ ("satisfy demand") that always satisfies 100% of the demand if feasible, or the maximum feasible amount otherwise: $x^{(D)}(S) = \min\{D, \sum_{i=1}^{N}(S_i - S_i^{\min})\}$.

To assess the "in-model" performance of $x^{(\theta)}$, we compute the relative gap $(1 - \text{LB}/\text{UB})$ between the lower bound ("LB") on the optimal (worst-case) expected costs from EC.1.2; and the upper bound from Section 4.3 ("UB") on the worst-case expected costs, when releases during the dry season follow our robust policy. We observe a 14.72% gap. This means that the worst-case expected cycle costs associated with the optimal policy amount to more than 85.28% of the costs associated with our robust policy, suggesting that the latter is relatively close to optimal.

The plots in Figure 5 illustrate the evolution of relative storage levels as a percentage of reservoir capacity, when releases are made according to the current policy versus our robust policy. Minimum levels and starting levels, where the latter also represent the target levels to complete a cycle, are indicated by horizontal lines. The plots track the storage levels at the beginning of each dry season and wet season. At each beginning of a dry season (May 1), which are marked by dashed vertical lines, it is assessed whether or not a cycle is complete. Completed cycles are indicated by asterisks on the horizontal x-axis.

Comparing the two plots, it is unsurprising to observe significantly higher storage levels when releases follow our robust policy (lower plot). However, it is noteworthy that the current policy (upper plot) seems to keep relative storage levels much closer among the different reservoirs than our policy does. This discrepancy is intentional for our policy, as it bases release decisions not only on the size of a reservoir but also on its expected inflow. In particular, Folsom Lake has a much larger catchment area relative to its capacity compared to the other reservoirs. For instance, recall from above that Trinity Lake has more than 2.5 times the capacity of Folsom Lake, but on average receives only less than 50% of its inflow. Aiming at filling all reservoirs simultaneously, our policy

Figure 5 Back-testing results (2006–2019): Evolution of relative storage levels (w.r.t. reservoir capacity), observed at the beginning of each dry season (May 1, dashed vertical lines) and wet season (Oct 1), when releases follow the current policy (upper) versus our robust policy (lower).

therefore prioritizes using water from Folsom Lake to satisfy downstream demand. Indeed, we observed that our policy always exhausted the entire ex-ante maximum feasible release amount from Folsom Lake over any dry season. In all scenarios in which the other two reservoirs did not reach their capacity, it was hence expected to observe the lowest relative storage level for Folsom Lake. The lower plot illustrates that this strategy generally worked well, with one exception in 2015, when both Shasta and Trinity reached their capacities but Folsom did not. However, the robust policy was able to keep Folsom above 35% of its capacity the following summer, whereas it hit the minimum level under the current policy (and so did Trinity).

The costs of a cycle are given by the sum of yearly shortage costs over the cycle. The lower plot in Figure 5 shows that eight cycles are completed for our robust policy $x^{(\theta)}$ over the evaluation period. The average costs of these cycles amount to 15.80 ($100M). In comparison, the total cost over the single cycle that is completed when releases are made according to the current policy $x^{\text{(curr)}}$ is 26.18 ($100M). Thus, the average cycle cost associated with $x^{(\theta)}$ is about 40% less than the cost associated with $x^{\text{(curr)}}$. When neglecting ambiguity, i.e., for $x^{(\mathbb{P})}$, higher releases would lead to fewer completed cycles (six in total) with an average cost of 17.62 ($100M). In other words, the benefit of robustness, i.e. implementing $x^{(\theta)}$ instead of $x^{(\mathbb{P})}$, amounts to a 10% cost reduction.

For $x^{(0)}$, an analogous plot to Figure 5 would only show a trivial picture, as a cycle is completed after every year. The average cycle cost for $x^{(0)}$ thus equals the yearly shortage costs of $\kappa \cdot D = 35.85$ ($100M). Hence, compared to $x^{(0)}$, our robust policy would reduce costs by about 56% . The plot for the policy that always tries to satisfy total demand ($x^{(D)}$) is also omitted because it has a pattern very similar to the current policy, with one (important) difference: for $x^{(D)}$ a cycle is not completed on May 1, 2019 (or afterward). The total cost over the evaluation period when releasing according to $x^{(D)}$ is 30.69 ($100M). Therefore, compared to $x^{(D)}$, our robust policy would reduce costs by approximately 48.5% . Note that $x^{\text{(curr)}}$ has an inherent informational advantage over the robust policy $x^{(\theta)}$ and all other policies considered in our model because the releases implemented by $x^{\text{(curr)}}$ were determined (in reality) gradually over the dry season, whereas in the model the release for the entire dry season must be decided at once on May 1. Hence, the performance estimates with respect to $x^{\text{(curr)}}$ are arguably conservative.

To summarize, when exposed to historical inflow scenarios from the evaluation period, our policy outperforms current practice in terms of average cycle costs. It also outperforms three other sensible benchmark policies. This analysis highlights the key trade-off that drives our policy: to keep expected cycle costs low, it is neither advisable to act extremely cautiously (saving all available water for the future), nor to act myopically (blindly satisfying immediate demand). The proposed policy balances this trade-off while being robust to climate uncertainty.

5.3. Sensitivity analysis and robustness checks

This subsection tests our main assumptions and estimated parameter values.

5.3.1. Different demand levels True (downstream) demand cannot be directly observed from the data, which is a common issue in various business contexts. The previous selection of the demand parameter D was based on average historical releases. Given that only censored data is available, we now test the sensitivity of our results under higher demand levels.

Table 1 shows the results of this analysis. The first column lists the tested demand levels D, relative to the baseline value $D_0 = 4{,}481{,}756$ (af) used in Section 5.2. The range of values considered corresponds approximately to two times the standard deviation in the historical water release data. The second column shows the corresponding confidence parameter θ, as determined by our calibration algorithm (c.f. Section 5.1). Columns 3–5 present the gaps between the performance bounds from Section EC.1.2 (LB) and Section 4.3 (UB). The remaining columns report the results for the actual inflows from 2006 to 2019, comparing our policy (including and neglecting ambiguity) to the no-release benchmark $x^{(0)}$ and the current policy $x^{(\text{curr})}$. We omit the results for the "satisfy demand" policy $x^{(D)}$ since a cycle would not be completed on May 1, 2019, making it difficult to compare to the other policies.

D	θ^*	LB	UB	Gap	Cycles		Avg. Cycle Cost	Cost Reduction		
					$x^{(\theta)}$	$x^{(\mathbb{P})}$	$x^{(\theta)}$	vs. $x^{(\mathbb{P})}$	vs. $x^{(0)}$	vs. $x^{(\text{curr})}$
100%	3000	12.91	15.14	14.7%	8	6	15.80	-10%	-56%	-40%
110%	3000	18.02	20.89	13.8%	9	8	20.27	-6%	-49%	-64%
120%	4000	22.27	24.01	7.2%	10	9	24.44	-3%	-43%	-74%
130%	4000	26.71	28.25	5.4%	10	10	30.40	+9%	-35%	-77%
140%	8000	30.11	31.08	3.1%	10	10	35.07	+5%	-30%	-80%

Table 1 Sensitivity analysis with respect to the downstream demand parameter D (given relative to D_0). All values for LB, UB, and average cycle costs are given in \$100M, θ^* is measured in \$/nat.

In Table 1, we first observe that the calibrated confidence parameter θ^* increases with the demand D. This is not surprising given that larger demand levels lead to higher costs, and θ regularizes immediate shortage costs against penalties for deviating from the nominal distribution. However, the variation in θ^* across the analyzed demand levels is relatively small, especially up to 130% of the baseline level, which indicates that the value of θ^* calibrated in Section 5.1 is relatively stable with respect to the demand level.

For all tested demand levels, we observe suboptimality gaps between 3% and 15%. This observation is consistent across a wide range of demand values, which is further evidence that the proposed robust policy is close to optimal. Indeed, the 14.7% gap observed in Section 5.2 would decrease if the true demand level D turned out to be higher.

The results using the actual inflows from 2006 to 2019 are favorable for the robust policy $x^{(\theta)}$ as well. First, note that the ambiguity-neglecting policy $x^{(\mathbb{P})}$ is expected to perform well when tested using the historical sample path. This is by construction of the $x^{(\mathbb{P})}$ policy, which only considers the cycle shortage cost, and because the data used to estimate the nominal distribution $\mathbb{P}$ contains the inflows from 2006 to 2019. Hence, it is quite remarkable that $x^{(\theta)}$ outperforms $x^{(\mathbb{P})}$ for demand levels up to 120% of the baseline value D_0, as shown in Table 1. The better performance of the robust policy $x^{(\theta)}$ is because it completes more cycles than the ambiguity-neglecting policy $x^{(\mathbb{P})}$, while keeping yearly shortage costs relatively low. For larger values of D, $x^{(\mathbb{P})}$ performs better than $x^{(\theta)}$ (as expected) because both policies complete an equal number of cycles and $x^{(\theta)}$ releases less water leading to higher shortage costs, but these scenarios represent severe cases of demand censoring (30% or more), which are less likely.

For the no-release policy $x^{(0)}$, a cycle was completed after each year (13 in total), with average costs growing linearly with demand. The smaller number of completed cycles but larger release amounts observed for the robust policy $x^{(\theta)}$ are associated with significantly lower average cycle costs. Finally, compared to the current policy $x^{(\text{curr})}$, the cost reduction of the robust policy is even more significant than initially suggested in Section 5.2.

5.3.2. Uncertain demand The above sensitivity analysis assumes a true constant demand level, known at the time of policy design and applied across all years in the evaluation period. We now examine the situation when only expected demand is known a priori, but actual demand is sampled over the evaluation period. The goal is to test the performance of the robust policy $x^{(\theta)}$ when committing to the releases upfront (based on a given demand level) while observing shortage costs based on realizations of random demand.

We again consider the period from May 1, 2006, to May 1, 2019, during which the current policy completes one cycle. To determine our policy, we fix the baseline demand $D_0 = 4{,}481{,}756$ (af) as well as the confidence parameter $\theta = 3000$ ($/nat). Then, we test our policy against the benchmark policies, with random yearly demand D sampled from a Normal distribution with mean $\mu_D = D_0$ and standard deviation $\sigma_D = 901{,}195$ (af), truncated to the interval $[0, 2D_0]$ to avoid negative demand. The value of σ_D is estimated as the empirical standard deviation of total dry season releases since the year 2000. In each period, the releases $x^{(\text{curr})}(S), x^{(\theta)}(S), x^{(\mathbb{P})}(S), x^{(D_0)}(S) = \min\{D_0, \sum_{i=1}^{N}(S_i - S_i^{\min})\}$, and $x^{(0)}(S) \equiv 0$ are based on the fixed value D_0 and the state S of the system. Then, the demand $\tilde{D}$ is sampled, and shortage costs $c(x) = \kappa \cdot (\tilde{D} - \sum_{i=1}^{N} x_i)^+$ are determined for each policy. Of course, the sampled demand $\tilde{D}$ might be above or below D_0.

Running a Monte-Carlo simulation with $10{,}000$ replications, the average cycle cost for the current policy $x^{(\text{curr})}$ amounts to 45.67 ($100M), while the cost for the robust policy $x^{(\theta)}$ is only 16.34 ($100M), which corresponds to a reduction of 64%. On average, the robust policy completes eight cycles. When neglecting ambiguity, the policy $x^{(\mathbb{P})}$ completes six cycles (on average), incurring an average cost of 18.80 ($100M). Hence, there is a 13% gain from following the robust approach. The benchmark policy $x^{(D_0)}$ performs poorly, not completing a single cycle and accruing a cost of 62.80 ($100M). Finally, the no-release policy $x^{(0)}$ has an average cost of 35.83 ($100M), which is more than twice the cost of the robust policy.

As an additional benchmark, consider an anticipative policy $x^{(\tilde{D})} = \min\{\tilde{D}, \sum_{i=1}^{N}(S_i - S_i^{\min})\}$, which observes the sampled demand $\tilde{D}$ at the beginning of a dry season and then satisfies it if feasible. Although this policy has a clear informational advantage over the robust policy, blindly satisfying demand would not be advisable. Indeed, the average cycle cost associated with $x^{(\tilde{D})}$ is 26.44 ($100M), corresponding to 1.11 cycles on average, so it performs far worse than $x^{(\theta)}$.

5.3.3. Two demand streams In the previous experiments, we accounted a constant cost of $\$800$ for each acre-foot of water shortage. However, when multiple demand streams are present, the impact of a supply shortage may vary depending on which stream it affects. This means different shortage costs could apply. Although accurately estimating demand and shortage costs becomes even more challenging with multiple demand streams, our model (together with all our established results) is not limited to a single demand stream. Therefore, we now test our model under the assumption of two demand streams with different shortage costs.

Note that having multiple demand streams creates two competing goals: (i) achieving a shorter cycle, as when there is a single demand stream; and (ii) satisfying the high-priority demand streams, i.e., those with higher shortage costs, in the future. Thus, there is an additional trade-off between satisfying non-essential needs today and leaving part of that demand unmet to save water for high-priority needs tomorrow.

The downstream demand D is now split into an essential component $D' = 0.2D$ and a non-essential component $D'' = 0.8D$. This split is chosen based on information obtained from the operating agency (White 2022). The essential component D' reflects the demand for public health and safety, with $\kappa' = 800$ ($), as before. For the non-essential component, we assume 70% lower costs, i.e., $\kappa'' = 240$ ($). The ratio between κ' and κ'' reflects the relation between estimated shortage costs for residential and agricultural customers, as suggested by McCann (2022). Given $\kappa' > \kappa''$, the cost function can be written as $c(x) = \kappa' \cdot (D' - \min\{x, D'\}) + \kappa'' \cdot (D'' - (x - \min\{x, D'\}))^+$.

D	$x^{(\theta)}$			Average cycle cost reduction				
	Avg. fill rate		Avg. cost	vs. $x^{(\mathbb{P})}$	vs. $x^{(\text{curr})}$	vs. $x^{(D_0)}$	vs. $x^{(D')}$	vs. $x^{(0)}$
	D'	D''	($100M)					
μ_D	100%	63%	4.74	−10%	−40%	−49%	−45%	−70%
$\mathcal{N}(\mu_D, \sigma_D)$	100%	68%	4.90	−13%	−64%	−73%	−46%	−69%

Table 2 Fill rate of essential demand and average cycle costs for two demand streams.

Table 2 summarizes the results under the assumption of a constant demand $D = D_0$ (first row), and in the setting of Section 5.3.2 with a Normally distributed demand around D_0 (second row). We include an additional benchmark policy $x^{(D')} = \min\{D', \sum_{i=1}^{N}(S_i - S_i^{\min})\}$, which only satisfies essential demand D' (if feasible). The second and third columns in Table 2 show the average fill rates for the two demand streams when applying the robust policy $x^{(\theta)}$. Notably, essential demand is always satisfied 100% of the time, across all 13 years and all 10,000 paths sampled in the Monte-Carlo simulation to generate the results in the second row. In contrast, non-essential demand is not fully met to achieve shorter cycles. This strategy proves successful compared to the other benchmark policies. In particular, the fifth column demonstrate that a more prudent choice of the D'' fill rate—$D = \mu_D$: 63% ($x^{(\theta)}$) vs. 68% ($x^{(\mathbb{P})}$); $D \sim \mathcal{N}(\mu_D l \sigma_D)$: 68% ($x^{(\theta)}$) vs. 73% ($x^{(\mathbb{P})}$)—can lead to significantly reduced average cycle costs.

5.3.4. Choice of the confidence parameter The confidence parameter θ plays an important role in our robust model. Hence, in this section we examine the performance of the proposed heuristic policy $x^{(\theta)}$ under different θ choices.

Table 3 shows a sensitivity analysis in the vicinity of $\theta^* = 3000$ ($/nat), which is the calibrated value from Section 5.1. We observe that the number of completed cycles decreases in θ. This is because lower confidence levels (i.e., smaller θ values) lead to more cautious release decisions and more completed cycles. In contrast, average cycle costs are not monotonic in θ, attaining their minimum at $\theta = 2000$ ($/nat). The corresponding cost savings compared to the current policy are about 44%. This finding supports the reliability of the calibration algorithm, as the ex-ante choice

θ (in \$/nat)	Completed Cycles	Avg. Cycle Cost (in \$100M)	Cost Reduction		
			vs. $x^{(\text{curr})}$	vs. $x^{(0)}$	vs. $x^{(\mathbb{P})}$
1000	10	17.42	-34%	-51%	-1%
2000	9	14.71	-44%	-59%	-17%
3000	8	15.80	-40%	-56%	-10%
4000	7	17.21	-34%	-52%	-2%
5000	6	19.61	-25%	-45%	$+11\%$

Table 3 Sensitivity analysis of the benchmark study results, for varying choices of θ.

of $\theta^* = 3000$ (\$/nat) is close to the ex-post optimal choice, capturing more than 90% of the ex-post cost savings potential.

For very low confidence levels (i.e., $\theta \downarrow 0$), the robust policy approximates the no-release policy $x^{(0)}$. Although a cycle is then completed after each year, it incurs very high yearly shortage costs, making such overly conservative choice of θ unattractive. For high confidence levels (i.e., large θ values), the number of completed cycles does not drop below six. Even the ambiguity-neglecting policy $x^{(\mathbb{P})}$ based on the nominal model, which releases considerably more than the robust policy for $\theta = 5000$ (\$/nat), results in six completed cycles for the given inflow scenarios. Hence, $x^{(\mathbb{P})}$ incurs smaller average cycle costs than $x^{(\theta)}$ for any $\theta \geq 5000$ (\$/nat), but both costs converge for θ sufficiently large. Notably, these results do not imply that $x^{(\theta)}$ with $\theta \geq 5000$ (\$/nat) would perform worse than $x^{(\mathbb{P})}$ for all sets of inflow scenarios. Indeed, for any sample path, there is a threshold beyond which $x^{(\mathbb{P})}$ outperforms $x^{(\theta)}$. For the given historical sample path, this threshold is at $\theta = 5000$ (\$/nat). However, the calibration algorithm determined a value of $\theta^* = 3000$ (\$/nat) for which the robust policy significantly outperforms $x^{(\mathbb{P})}$.

6. Conclusion

We have presented a model for the sustainable management of a system of water reservoirs, all located in one river basin and collectively supplying downstream demand. The largest reservoir system in California served as our motivation and case study. Using tools from distributionally robust optimization and stochastic dynamic programming, we have incorporated major features such as climate uncertainty and a stochastic planning horizon in a tractable manner. Our theoretical analysis revealed structural results, allowing us not only to derive insights into the optimal policy but also to suggest a computationally tractable heuristic that overcomes the curse of dimensionality, together with performance bounds. Using real data from the Sacramento River Basin, we demonstrated the strength of the robust policy with suboptimality gaps below 15%. In a back-testing benchmark study, we demonstrated the potential of the robust policy to significantly reduce average shortage costs over "full-to-full" cycles, usually by more than 40% compared to the current policy and other benchmarks.

The obtained results are based on the assumption that inflows into individual reservoirs are perfectly correlated. In the case of California, this assumption is supported by very high correlations observed in historical data (see Figure 1). For other river basins with less correlation, the relaxation of this assumption might be an interesting direction for future research. We note that the majority of our results could be extended to a model with multiplicative white noise that is reservoir-specific. However, the resulting bounds would be weaker and solving the value iteration problem would become computationally more challenging. Most critically, one would have to rethink Proposition 7, which does not extend naturally. Hence, exploring other heuristics could be worthwhile.

We believe that the results contained in this paper have the potential to impact the practice of water resources management in two significant ways. First, we hope that the demonstrated economic potential to significantly reduce average cycle costs can influence practitioners to consider alternative planning horizons, performance metrics, and the impact of climate uncertainty. Second, our model may serve as a decision support tool for policymakers, for instance, related to the actively debated question of whether it is preferable to extend the capacity of existing reservoirs or to add newly built infrastructure (see, e.g., Barringer 2022). Given reliable cost estimates for such plans, our model could help examine the operational impact of the options under consideration. Another example is the question of centralizing reservoir management at the level of river basins, where the operational benefits of coordination could be assessed using our model. Numerous use cases for our model seem to exist, but each warrants a study of its own, going beyond the scope of the present paper.

In conclusion, the sustainable management of water reservoir systems is critical for ensuring water security, especially in the face of climate uncertainty. Our model provides a robust framework for strategically optimizing reservoir operation over dry seasons, offering both theoretical insights and practical guidelines. Future research can build on our framework to explore additional aspects of water resource management. We hope that our work will ultimately contribute to more efficient and sustainable water management practices.

Acknowledgments

The authors are grateful to the United States Bureau of Reclamation, in persona Deputy Regional Director Kristin White, for valuable discussions and feedback on our modeling assumptions.

References

American Society of Civil Engineers, ASCE . 2021. Dams 2021. *2021 Report Card for America's Infrastructure* .

Barringer, Felicity. 2022. Does drought-prone California need another reservoir? *'& the West', a magazine presented by Stanford University's Bill Lane Center for the American West* .

Bertsekas, Dimitri P., John N. Tsitsiklis. 1989. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific.

Bertsekas, Dimitri P., John N. Tsitsiklis. 1991. An Analysis of Stochastic Shortest Path Problems. *Mathematics of Operations Research* **16**(3) 580–595.

Bertsimas, Dimitris, Vishal Gupta, Nathan Kallus. 2018. Data-driven robust optimization. *Mathematical Programming* **167**(2) 235–292.

Brown, David B., James E. Smith, Peng Sun. 2010. Information Relaxations and Duality in Stochastic Dynamic Programs. *Operations Research* **58**(4, Part 1) 785–801.

Chun, Myung J. 2022. Shasta Lake at 38% capacity heading into the hottest months of the year. `https://www.latimes.com/california/story/2022-07-22/shasta-lake-at-38-percent-capacity-heading-into-the-hottest-months-of-the-year`. Accessed 30-Aug-2024.

Donsker, Monroe D., S. R. Srinivasa Varadhan. 1983. Asymptotic evaluation of certain Markov process expectations for large time. IV. *Communications on Pure and Applied Mathematics* **36**(2) 183–212.

Elleuch, Mohamed, Frikha Ahmed, Francisco Silva Pinto. 2022. A Review On Water Resources Management Modeling Using Operational Research and Decision Aid Sciences. *2022 International Conference on Decision Aid Sciences and Applications (DASA)*. 1565–1571.

Esfahani, Peyman Mohajerin, Daniel Kuhn. 2018. Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming Series A* **171** 115–166.

Fahlbusch, H. 2009. Early dams. *Proceedings of the Institution of Civil Engineers - Engineering History and Heritage* **162**(1) 13–18.

Gauvin, Charles, Erick Delage, Michel Gendreau. 2017. Decision rule approximations for the risk averse reservoir management problem. *European Journal of Operational Research* **261**(1) 317–336.

Gauvin, Charles, Erick Delage, Michel Gendreau. 2018a. A stochastic program with time series and affine decision rules for the reservoir management problem. *European Journal of Operational Research* **267**(2) 716–732.

Gauvin, Charles, Erick Delage, Michel Gendreau. 2018b. A successive linear programming algorithm with non-linear time series for the reservoir management problem. *Computational Management Science* **15**(1) 55–86.

Georgakakos, Aris P., Konstantine Georgakakos. 2007. Integrated Forecast and Reservoir Management for Northern California: System Development and Initial Demonstration. Tech. rep., Hydrologic Research Center and Georgia Water Resources Institute.

Georgakakos, K. P., T. M. Modrick, J. A. Sperfslage, C. R. Spencer, R. J. Banks, A. P. Georgakakos, M. Kistenmacher, X. Liu. 2018. Integrated Forecast and Reservoir Management (INFORM). *Final Report Prepared for the State of California, Department of Water Resources* .

Giuliani, M., Jonathan R. Lamontagne, P. M. Reed, Andrea Castelletti. 2021. A State-of-the-Art Review of Optimal Reservoir Control for Managing Conflicting Demands in a Changing World. *Water Resources Research* **57**.

González-Trejo, J. I., O. Hernández-Lerma, L. F. Hoyos-Reyes. 2002. Minimax Control of Discrete-Time Stochastic Systems. *SIAM Journal on Control and Optimization* **41**(5) 1626–1659. doi:10.1137/S0363012901383837. URL `https://doi.org/10.1137/S0363012901383837`.

Hastie, Trevor, Robert Tibshirani, Jerome H. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics, Springer.

Hu, Zhaolin, L. Jeff Hong. 2013. Kullback-Leibler Divergence Constrained Distributionally Robust Optimization. *Available at Optimization Online* .

Jeuland, Marc, Justin Baker, Ryan Bartlett, Guillaume Lacombe. 2014. The costs of uncoordinated infrastructure management in multi-reservoir river basins. *Environmental Research Letters* **9** 105006.

Kim, Michael Jong, Andrew E. B. Lim. 2016. Robust Multiarmed Bandit Problems. *Management Science* **62**(1) 264–285.

Le, Phong, James Randerson, Rebecca Willett, Stephen Wright, Padhraic Smyth, Clement Guilloteau, Antonios Mamalakis, Efi Foufoula-Georgiou. 2023. Climate-driven changes in the predictability of seasonal precipitation. *Nature Communications* **14**.

Martinez-Villalobos, Cristian, J. David Neelin. 2019. Why Do Precipitation Intensities Tend to Follow Gamma Distributions? *Journal of the Atmospheric Sciences* **76** 3611–3631.

McCann, Richard. 2022. Benefit/Cost Assessment of CVWD's CAPP. Tech. rep., M.Cubed.

Mintzberg, Henry. 1989. *The structuring of organizations*. Springer.

Murata, Akihiko, Shun-ichi I. Watanabe, Hidetaka Sasaki, Hiroaki Kawase, Masaya Nosaka. 2020. Assessing Goodness of Fit to a Gamma Distribution and Estimating Future Projection on Daily Precipitation Frequency Using Regional Climate Model Simulations over Japan with and without the Influence of Tropical Cyclones. *Journal of Hydrometeorology* **21**(12) 2997–3010.

Park, Jangho, Güzin Bayraksan. 2023. A multistage distributionally robust optimization approach to water allocation under climate uncertainty. *European Journal of Operational Research* **306**(2) 849–871.

Stanton, Elizabeth A., Ellen Fitzgerald. 2011. California Water Supply and Demand: Technical Report. Tech. rep., Stockholm Environment Institute - U.S. Center.

Stratonovich, Ruslan L. 2020. *Theory of Information and its Value*. Springer.

White, Kristin. 2022. Personal Communication, 12/06/2022. Deputy Regional Director over California operations for the US Bureau of Reclamation.

Zhang, Can, Atalay Atasu, Turgay Ayer, L. Beril Toktay. 2020. Truthful Mechanisms for Medical Surplus Product Allocation. *Manufacturing & Service Operations Management* **22**(4) 735–753.

EC.0. Additional notation for the EC

Some parts of both EC.1 and EC.2 require the introduction of additional notation. For some vector y, we denote by $\pi^a(y)$ and $\pi^d(y)$ the permutations of indices such that the values of y are sorted in ascending and descending order, respectively. In this way, we can also sort one vector according to another one. For instance, let $x = [10, 11, 12, 13]$ and $y = [8, 6, 9, 7]$. Then $\pi^a(y) = [2, 4, 1, 3], \pi^d(y) = [3, 1, 4, 2], x_{\pi^a(y)} = [11, 13, 10, 12], x_{\pi^d(y)} = [12, 10, 13, 11]$. To access elements of the sorted vectors, we use the notation $x_{\pi(y),i}$ such that, e.g., $x_{\pi^a(y),1} = 11$ and $x_{\pi^d(y),2} = 10$. When considering the sorted values of y directly, then we abbreviate the i-th order statistic by $y_{\pi^a,i} := y_{\pi^a(y),i}$ and $y_{\pi^d,i} := y_{\pi^d(y),i}$.

EC.1. Supplemental material

This section contains auxiliary results and detailed constructions/algorithms to complement the main text of the paper.

EC.1.1. Auxiliary results for Section 3.3

The following results are used in the buildup to Proposition 7.

LEMMA EC.1. *If for two states $S^{(1)}$ and $S^{(2)}$ it holds almost everywhere (i.e., for all inflow scenarios Γ) that*

$$\sum_{i=1}^{N} \min\left\{C_i, S_i^{(1)} + \xi_i\right\} \geq \sum_{i=1}^{N} \min\left\{C_i, S_i^{(2)} + \xi_i\right\}, \tag{EC.1}$$

then

$$\max_{i=1,\ldots,N} \delta_i(S_i^{(1)}) \leq \max_{i=1,\ldots,N} \delta_i(S_i^{(2)}). \tag{EC.2}$$

If strict inequality holds in (EC.1) *with positive probability, then strict inequality holds in* (EC.2).

LEMMA EC.2. *Let S be a balanced state. Consider a release vector x^* such that $\frac{1}{(1-\beta_i)\alpha_i} x_i^* = \frac{1}{(1-\beta_j)\alpha_j} x_j^*$ for all $1 \leq i, j \leq N$. Then, for any other release vector x with $\sum_i x_i = \sum_i x_i^*$, it holds almost everywhere that $\sum_i \min\{C_i, S_i + \xi_i - x_i^*\} \geq \sum_i \min\{C_i, S_i + \xi_i - x_i\}$. Strict inequality holds with positive probability.*

COROLLARY EC.1. *Let the state S of a system be given. Consider two feasible release decision vectors x and x^* with $\sum_{i=1}^N x_i = \sum_{i=1}^N x_i^*$, such that the condition*

$$\sum_{i=1}^{N} \min\{C_i, S_i + \xi_i - x_i^*\} \geq \sum_{i=1}^{N} \min\{C_i, S_i + \xi_i - x_i\} \tag{EC.3}$$

holds almost everywhere and strict inequality holds with a positive probability. Then

$$c(x^*) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_\theta(\min\{C, S+\xi-x^*\})}{\theta}}\right]\right) < c(x) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_\theta(\min\{C, S+\xi-x\})}{\theta}}\right]\right).$$

The interpretation of Corollary EC.1 is that if condition (EC.3) holds for a given pair of feasible release decision vectors x and x^*, which correspond to an equal total amount of water to be released from the system, then x^* is a better decision than x.

EC.1.2. Tightening the lower bound of Section 4.1

Any given state $S^{(1)}$ of the aggregate reservoir corresponds to at least one feasible state $S^{(N)}$ of the N-reservoirs system, such that $\sum_i S_i^{(N)} = S^{(1)}$. If $S^{(N)}$ is ambiguous, then the candidate states can be ranked with respect to their favorability. In particular, a system that is in a balanced (post decision) state either reaches the terminal state or it can store the entire total inflow over the upcoming wet season. Thus, a balanced system is most favorable, as it stores at least as big a part of the total inflow as a system in any other state $S^{(N)}$ with $\sum_i S_i^{(N)} = S^{(1)}$. If there is no feasible balanced state $S^{(N)}$ with $\sum_i S_i^{(N)} = S^{(1)}$ (due to the minimum levels), then there is a most favorable feasible weakly-balanced state with this property. Given $S^{(1)}$ together with a release decision $X(S^{(1)})$, the structure of an optimal policy known from Section 3.3 allows to determine the most favorable post-decision state in the corresponding N-reservoirs system by determining the most favorable pre-decision state $\hat{S}^{(N)}$ with $\sum_i \hat{S}_i^{(N)} = S^{(1)}$, and then splitting up $X(S^{(1)})$ among the reservoirs in an optimal way. The technical construction of such a *most favorable* state $\hat{S}^{(N)}$ with given total amount of stored water $S^{(1)}$ is detailed below the Proposition.

PROPOSITION EC.1. *Consider the setting of Proposition 8, but let the state update for the aggregated problem be given by*

$$S_{t+1}^{(1)} = \min\left\{C^{(1)}, S_t^{(1)} + \sum_{i=1}^{N} \min\left\{(1-\beta_i)\alpha_i\Gamma, C_i - \hat{S}_i^{(N)}(S_t^{(1)})\right\} - X(S_t^{(1)}) - e^{(1)}\right\},$$

where $\hat{S}_i^{(N)}(S_t^{(1)})$ denotes the i-th component of the most favorable state $\hat{S}^{(N)}$ with $\sum_{i=1}^N \hat{S}_i^{(N)} = S^{(1)}$, as constructed below. Denote by $\hat{v}_\theta^{(1)}, v_\theta^{(1)}$, and $v_\theta^{(N)}$ the optimal value function to this problem, the one of Proposition 8, and the one corresponding to the N-reservoirs system, respectively. Then, it holds for any state $S^{(N)}$ of the system of N reservoirs, that

$$v_\theta^{(1)}\left(\sum\nolimits_i S_i^{(N)}\right) \le \hat{v}_\theta^{(1)}\left(\sum\nolimits_i S_i^{(N)}\right) \le v_\theta^{(N)}\left(S^{(N)}\right) .$$

On the construction of the most favorable state. Let the total amount of stored water $S^{(1)}$ in the system of N reservoirs be given. A balanced state $S = (S_1, \dots, S_N)$ such that $\sum_{i=1}^N S_i = S^{(1)}$, is uniquely determined by setting

$$S_i = \frac{\alpha_i(1-\beta_i)}{\sum_{j=1}^N \alpha_j(1-\beta_j)}\left(S^{(1)} + \frac{\sum_{j\neq i}\alpha_j(1-\beta_j)}{\alpha_i(1-\beta_i)}(C_i+e_i) - \sum_{j\neq i}(C_j+e_j)\right) .$$

If such a state S results in a set of indices $\mathcal{J} \subset \{1,\dots,N\}$ with $S_j < S_j^{\min}$ for $j \in \mathcal{J}$, then setting

$$\begin{aligned} &\tilde{S}^{(1)} := S^{(1)} - \sum_{j\in\mathcal{J}} S_j^{\min} \\ &\forall j \in \mathcal{J} : S_j := S_j^{\min} \\ &\forall i \notin \mathcal{J} : S_i := \frac{\alpha_i(1-\beta_i)}{\sum_{k\notin\mathcal{J}}\alpha_k(1-\beta_k)}\left(\tilde{S}^{(1)} + \frac{\sum_{k\notin\mathcal{J},k\neq i}\alpha_k(1-\beta_k)}{\alpha_i(1-\beta_i)}(C_i+e_i) - \sum_{k\notin\mathcal{J},k\neq i}(C_k+e_k)\right) \end{aligned} \tag{EC.4}$$

gives a weakly-balanced state S with $\sum_{i=1}^N S_i = S^{(1)}$. If this state again leads to a (new) set of indices $\mathcal{J}'$ with $S_{j'} < S_{j'}^{\min}$ for all $j' \in \mathcal{J}'$, then the procedure can iteratively be repeated with an updated set $\mathcal{J} = \mathcal{J} \cup \mathcal{J}'$.

EC.1.3. On the construction of the least favorable state in Section 4.3

Let the total amount of stored water $S^{(1)}$ in the system of N reservoirs be given. Starting at the beginning of a cycle and releasing according to the policy of Section 4.2, the most unbalanced state S with $\sum_{i=1}^N S_i = S^{(1)}$, which could be reached when all reservoirs had unlimited capacity, would be the result of inflow after the system was in the post-decision state $S^{\min} = (S_1^{\min}, \dots, S_N^{\min})$. This state S would be given by

$$S_i = S_i^{\min} + \alpha_i(1-\beta_i)\cdot\frac{S^{(1)} - \sum_{j=1}^N (S_j^{\min} - e_j)}{\sum_{j=1}^N \alpha_j(1-\beta_j)} - e_i \tag{EC.5}$$

for all $i = 1, \ldots, N$. If such a state S results in a set of indices $\mathcal{J} \subsetneq \{1, \ldots, N\}$ with $S_j > C_j$ for all $j \in \mathcal{J}$, then S cannot be reached from $(S_1^{\min}, \ldots, S_N^{\min})$ in a feasible way due to the different capacities. One therefore needs to adjust the preceding post-decision state. Assuming a discrete state space with stepsize Δ_i for reservoir i, this adjustment to determine a preceding post-decision state S' that can lead to the most unbalanced state S in a feasible way due to inflow realization, can be made by performing the following iterative procedure:

1: Starting point: S according to (EC.5); $\exists \mathcal{J} \subsetneq \{1, \ldots, N\}$ with $S_j > C_j$ for all $j \in \mathcal{J}$
2: Init $S_i^{\min\prime} := S_i^{\min}, \delta_i' := \frac{1}{\alpha_i(1-\beta_i)} \left(C_i - (S_i^{\min\prime} - e_i)\right) \forall i = 1, \ldots, N$
3: Init $k := 2$
4: **while** $|\mathcal{J}| > 0$ **do**
5: **if** $S_{\pi^a(\delta'),k}^{\min\prime} + \Delta_k < C_k + e_k - \alpha_k(1-\beta_k) \cdot \delta'_{\pi^a,k-1}$ **then**
6: Update $S_{\pi^a(\delta'),k}^{\min\prime} := S_{\pi^a(\delta'),k}^{\min\prime} + \Delta_k$
7: **else**
8: Update $S_{\pi^a(\delta'),k}^{\min\prime} := C_k + e_k - \alpha_k(1-\beta_k) \cdot \delta'_{\pi^a,1}$
9: Set $k := k+1$
10: **end if**
11: For all $i = 1, \ldots, N$, compute S_i' according to (EC.5) based on $S_i^{\min\prime}$
12: Update $\mathcal{J} := \{j = 1, \ldots, N : S_j' > C_j\}$
13: **end while**
14: **return** S'

The "else" statement in Step 7 ensures that $\delta_k' \leq \delta'_{\pi^a,1}$ for all $k = 1, \ldots, N$. The notation used in the algorithm is defined in EC.0.

EC.1.4. An upper bound based on information relaxation

Upper bounds for maximization problems can generally be obtained by enlarging the set of feasible solutions. For our problem (7), this can be achieved by relaxing the nonanticipativity constraint on the transition distributions. This idea is usually referred to as *information relaxation* (see Brown et al. 2010 for a general introduction). Our approach is adopted from Kim and Lim (2016).

Given our discrete state space, optimizing over transition distributions $\mathbb{Q}_{S_t, x(S_t)}$ in (7) is equivalent to optimizing over likelihood ratios

$$y_t(j) := \frac{\mathbb{Q}[S_{t+1} = j | S_t, x(S_t)]}{\mathbb{P}[S_{t+1} = j | S_t, x(S_t)]}.$$

Let $\mathcal{Y} := \{(y_t)_{t\geq 0} : y_t \geq 0, \sum_{j\in\mathcal{S}} y_t(j)\mathbb{P}[S_{t+1}=j|S_t,x(S_t)] = 1\ \forall S_t \in \mathcal{S}, t\geq 0\}$ denote the set of all admissible sequences of such likelihood ratios. Then, problem (7) can be written as

$$\overline{v}_\theta(x,C) = \max_{y\in\mathcal{Y}}\ \mathbb{E}^{\mathbb{P}}\left[\sum_{t=0}^{\tau-1} z_t \cdot c_\theta(S_t,x(S_t),y_t) + z_\tau\cdot 0 \middle| S_0 = C\right] \tag{EC.6a}$$

$$\text{s.t. } z_0 = 1, z_t = \prod_{k=0}^{t-1} y_k(S_{k+1}), \tag{EC.6b}$$

where $c_\theta(S_t,x(S_t),y_t) := c(x(S_t)) - \theta\sum_{j\in\mathcal{S}} y_t(j)\log(y_t(j))\mathbb{P}[S_{t+1}=j|S_t,x(S_t)]$. Define the filtration $\mathcal{G}_t := \sigma(\{S_0,x(S_0),\ldots,S_t,x(S_t)\})$. Then $\mathcal{G}_\infty$ corresponds to the complete record of states and release decisions until the terminal state is reached. Let $\overline{\mathcal{Y}}$ be the set of all admissible, $\mathcal{G}_\infty$-measurable sequences of likelihood ratios. Essentially, $\overline{\mathcal{Y}}$ corresponds to the set of processes corresponding to likelihood ratios on a given path $(s_0,x(s_0),s_1,x(s_0)\ldots,s_\tau)$. Then, an upper bound for problem (EC.6) is given by:

$$\overline{v}_\theta(x,C) \leq \max_{y\in\overline{\mathcal{Y}}}\left\{\mathbb{E}^{\mathbb{P}}\left[P(y)|S_0=C\right] \text{ s.t. (EC.6b)}\right\} = \mathbb{E}^{\mathbb{P}}\left[\max_{y\in\overline{\mathcal{Y}}}\{P(y) \text{ s.t. (EC.6b)}\}\middle| S_0 = C\right], \tag{EC.7}$$

where $P(y) := \sum_{t=0}^{\tau-1} z_t\cdot c_\theta(s_t,x(s_t),y_t)$. To tighten the upper bound in (EC.7), a term

$$\lambda(h,y) := \sum_{t=0}^{\tau-1} z_t\left(y_t(S_{t+1})h\left(x(S_t),S_{t+1}\right) - \sum_{j\in\mathcal{S}} y_t(j)h(x(S_t),j)\mathbb{P}[S_{t+1}=j|S_t,x(S_t)]\right),$$

with a bounded function $h:\mathcal{X}\times\mathcal{S}\to\mathbb{R}, |h(\cdot,\cdot)|\leq c_h<\infty$, can be introduced to, intuitively speaking, penalize the exploitation of being clairvoyant. The resulting penalized problem is given by

$$\mathbb{E}^{\mathbb{P}}\left[\max_{y\in\overline{Y}}\{P(y)-\lambda(h,y) \text{ s.t. (EC.6b)}\}\right]. \tag{EC.8}$$

The following lemma shows that the penalty term has expectation zero.

LEMMA EC.3. *For any $y\in\mathcal{Y}$, the penalty function $\lambda(h,y)$ satisfies*

$$\mathbb{E}^{\mathbb{P}}[\lambda(h,y)|z_0,y_0] = 0.$$

Therefore, also the penalized problem (EC.8) yields an upper bound on $\overline{v}_\theta(x,C)$.

In contrast to the discounted infinite horizon setting used in Kim and Lim (2016), our stochastic shortest path model does not seem to allow showing Lemma EC.3 based on the Bounded Convergence Theorem. Our proof in Section EC.2 is therefore based on martingale theory instead. Given Lemma EC.3, the following proposition can then be shown in a similar way as Theorem 3 in Kim and Lim (2016).

PROPOSITION EC.2. *For a given policy $x(\cdot)$, let a path $(s_0, x(s_0), s_1, x(s_0) \ldots, s_\tau)$ be given. Define the constant*

$$\beta_t(f) := \Bigg[\exp\left(\frac{f - h\left(x(s_t), s_{t+1}\right)\left(1 - \mathbb{P}[S_{t+1} = s_{t+1} | s_t, x(s_t)]\right)}{\theta \cdot \mathbb{P}[S_{t+1} = s_{t+1} | s_t, x(s_t)]}\right) \cdot \mathbb{P}[S_{t+1} = s_{t+1} | s_t, x(s_t)] + \sum_{j \neq s_{t+1}} \exp\left(\frac{h\left(x(s_t), j\right)}{\theta}\right) \cdot \mathbb{P}[S_{t+1} = j | s_t, x(s_t)]\Bigg]^{-1}.$$

Then, the optimal objective value f_0^λ for the problem $\max_{y \in \overline{Y}} \{P(y) - \lambda(h, y)$ s.t. (EC.6b)$\}$ can be obtained by the following backwards recursion:

$$f_\tau^\lambda = 0$$

$$f_t^\lambda = c(x(s_t)) - \theta \cdot \log\left(\beta(f_{t+1}^\lambda)\right) \text{ for } t = \tau - 1, \ldots, 0\,.$$

For a given policy $x(\cdot)$, a path $(s_0, x(s_0), \ldots, s_\tau)$ of states and releases can easily be simulated under the nominal distribution $\mathbb{P}$. Given such a simulated path, Proposition EC.2 provides a recursion scheme to compute the optimal value for the pathwise maximization problem $\max_{y \in \overline{Y}} \{P(y) - \lambda(h, y)$ s.t. (EC.6b)$\}$. The value of the upper bound (EC.8) can thus be estimated by a standard Monte-Carlo simulation. A feasible choice for the penalty function would, for example, be given by setting $h(\cdot)$ equal to the value function of the aggregated system studied in Sec. 4.1.

EC.1.5. A cross-validation algorithm to calibrate the confidence parameter θ

Fix values for the hyperparameters K, n_K, the number of simulation runs L, and candidate values $\{\theta_1, \ldots, \theta_M\}$
Follow the bootstrapping method of Sect. 5.1 to generate a set Ω of $n_K \cdot K$ hypothetical yearly inflows.
Partition Ω into K subsets $\Omega_1, \ldots, \Omega_K$, such that $|\Omega_k| = n_K$ for all $k = 1, \ldots, K$.
for $m = 1, \ldots, M$ **do**
  for $k = 1, \ldots, K$ **do**
    Set $\mathcal{V}_k := \Omega_k$ and $\mathcal{T}_k := \cup_{j \neq k} \Omega_j$.
    Fit a (gamma mixture) distribution $\mathbb{P}_k$ based on the data in $\mathcal{T}_k$.
    Determine the policy $x_{m,k}$ according to the heuristic of Section 4.2, based on $\theta = \theta_m$ and $\mathbb{P} = \mathbb{P}_k$.
    Simulate L cycles for the policy $x_{m,k}$, by randomly drawing inflows (with replacement) from the validation set $\mathcal{V}_k$. Denote the simulated cycle costs by $\phi_{k,1}(\theta_m), \ldots, \phi_{k,L}(\theta_m)$.
    Compute the average cycle cost $\phi_k(\theta_m) := \frac{1}{L}\sum_{\ell=1}^{L} \phi_{k,\ell}(\theta_m)$ based on $\mathcal{V}_k$ and $\mathcal{T}_k$ (and θ_m).
  end for
  Compute $\phi(\theta_m) := \frac{1}{K}\sum_{k=1}^{K} \phi_k(\theta_m)$ as an estimate of the average out-of-sample cost associated with θ_m.
end for
return $\theta = \operatorname{argmin}_{m=1,\ldots,M} \phi(\theta_m)$.

EC.1.6. On the estimation of the inflow constants α_i in Section 5

Consider N reservoirs $i = 1, \ldots, N$. Let yearly inflow data $\widehat{\Gamma}_{t,i}$ be given for each reservoir i over $t = 1, \ldots, T$ years. To obtain such $\widehat{\Gamma}_{t,i}$, we aggregated the available daily data on a yearly basis (from May to April in the following year), for $T = 24$ years (2000–2023). We then estimated $\alpha_i, i = 1, \ldots, N$, according to the following procedure:

```
1: Init data as a T-dimensional vector.
2: for t = 1,...,T do
3:     data(t) = Γ̂_{t,i} / Γ̂_{t,1}
4: end for
5: return α_i := (1/T) Σ_{t=1}^{T} data(t)
```

By construction, $\alpha_1 = 1$ holds. For all $i \geq 2$, α_i reflects the average yearly inflow relative to the first reservoir (Shasta Lake).

EC.1.7. On the estimation of the outflow constants β_i in Section 5

Consider N reservoirs together with yearly inflow data $\widehat{\Gamma}_{t,i}$ as in Section EC.1.6 above. For each $t = 1, \ldots, T$, let outflow data $B_{t,i}$ be given for the respective rainy season. For instance, $\widehat{\Gamma}_{1,1}$ represents the aggregate inflow into Shasta Lake from 2000/05/01 to 2001/04/30, while $B_{1,1}$ represents the aggregate outflow from Shasta Lake from 2000/10/01 to 2001/04/30. We then estimated $\beta_i, i = 1, \ldots, N$, according to the following procedure:

```
1: Init data as a T-dimensional vector.
2: for t = 1,...,T do
3:     data(t) = B_{t,i} / Γ̂_{t,i}
4: end for
5: return β_i := (1/T) Σ_{t=1}^{T} data(t)
```

Then β_i reflects the average outflow from reservoir i during a wet season (October to April in the following year), relative to the inflow over the entire year (May to April in the following year) containing that wet season.

EC.2. Proofs

Proof of Lemma 1. We assume a discrete state space here. This is to directly relate to the definition of stochastic shortest path problems presented in Bertsekas and Tsitsiklis (1989), and allows a constructive proof using standard linear algebra arguments.

For the (post decision) state transition matrix P, let the element in the upper left corner correspond to remaining in the absorbing state $\overline{S}$, and the element in the lower right corner correspond to remaining in the state where $S_i^{\min}$ holds for all i. Order all other states (along each axes) in such a way that no state dominates any preceding one in all dimensions. Then, the transition matrix satisfies the following properties: (i) the first element equals 1; (ii) all entries in the first column as well as on the main diagonal are strictly positive; (iii) the rows of the matrix $P - \lambda I$ are linearly independent, for all eigenvalues λ of P and I denoting the identity matrix. Thus, the matrix P is diagonalizable and we may write the t-th power of P as $P^t = VD^tV^{-1}$, where D is the matrix of eigenvalues and V a matrix of (columnwise) corresponding eigenvectors. The largest eigenvalue of P equals 1. All other eigenvalues λ_i, corresponding to the elements on the main diagonal of P starting from row 2, satisfy $0 < \lambda_i < 1$ and therefore $\lim_{t\to\infty} \lambda_i^t = 0$. Thus, $\lim_{t\to\infty} D^t = \mathrm{diag}(1, 0, \ldots, 0)$. The matrix V is a lower triangular matrix, where all elements in the first column have the same value. The inverse of a lower triangular matrix is another lower triangular matrix, whose main diagonal consists of the reciprocals of the original diagonal elements. Thus, $\lim_{t\to\infty} P^t = V \lim_{t\to\infty} D^t V^{-1}$ is a matrix where all elements in the first column are one and all other elements are zero. This proves the assertion that, in the limit, the absorbing state is reached with probability one from any given state. □

Proof of Proposition 1. Follows directly from Prop. 3 in Bertsekas and Tsitsiklis (1989). □

Proof of Proposition 2. Let $c_\theta(x(S_t), \mathbb{Q}_t) := c(x(S_t)) - \theta D_{\mathrm{KL}}(\mathbb{Q}_t \| \mathbb{P})$. With this notation, the robust model in Equation (4) can be expressed as follows:

$$v_\theta(S) = \min_{x \geq 0} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=0}^{\tau} c_\theta(x(S_t), \mathbb{Q}_t) \middle| S_0 = S \right] \tag{EC.9}$$

$$\text{s.t. } S_{t+1,i} = \min\{S_{t,i} + \xi_{t,i} - x_i(S_t), C_i\} \geq S_i^{\min} \; a.s. \; \forall i \leq N, \forall t \geq 0,$$

where $\mathcal{Q} = \{\mathbb{Q}_t \in \Pi, t \geq 0\}$ and $\Pi = \{\mathbb{Q} = \mathbb{Q}(S, x(S)) : S \in \mathcal{S}, \mathbb{Q} \ll \mathbb{P}\}$.

Let $\mathbb{Q}^* = (\mathbb{Q}_0^*, \mathbb{Q}_1^*, \mathbb{Q}_2^*, \ldots)$ denote the optimal $\mathbb{Q}$ in Equation (EC.9). By the well-known Donsker and Varadhan's variational formula (see Donsker and Varadhan 1983), the inner stage-wise, conditional maxima contained (implicitly) in the objective of (EC.9), are attained by the distributions $\mathbb{Q}_t^*$ with Radon-Nikodym derivative

$$\frac{d\mathbb{Q}_t^*}{d\mathbb{P}} = \frac{e^{\frac{\mathbb{E}^{\mathbb{Q}^*}\left[\sum_{k=t+1}^{\tau} c_\theta(x(S_k),\mathbb{Q}_k^*)\middle|S_{t+1}\right]}{\theta}}}{\mathbb{E}^{\mathbb{P}}\left[e^{\frac{\mathbb{E}^{\mathbb{Q}^*}\left[\sum_{k=t+1}^{\tau} c_\theta\left(x(S_k),\mathbb{Q}_k^*\right)\middle|S_{t+1}\right]}{\theta}}\middle|S_t, x(S_t)\right]}, \tag{EC.10}$$

such that $\mathbb{Q}_t^*(S_{t+1} = \overline{S}) > 0$ holds for all S_t and feasible $x(S_t)$ due to Assumption 1 on $\mathbb{P}$. Then, it follows analogously to Lemma 1 that all policies are proper under the $\mathbb{Q}^*$ distribution.

We now show that a minimax Bellman equation holds for (EC.9). Define a sequence $(v_\theta^{(k)}(S))_{k\geq 0}$ for each state S, where $v_\theta^{(k)}(S) := \min_{0\leq x\leq S-S^{\min}} \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\left[\sum_{t=0}^{\min\{\tau,k\}} c_\theta(x(S_t),\mathbb{Q}_t)\middle|S_0 = S\right]$ denotes the optimal worst-case expected costs of reaching the target state $\overline{S}$ in at most k steps, with the state-update being defined as in (EC.9). For this k-step finite-horizon cost approximation, a minimax Bellman equation is available (see González-Trejo et al. (2002, Theorem 3.1)):

$$v_\theta^{(k)}(S_t) = \min_{0\leq x\leq S_t - S^{\min}} \sup_{\mathbb{Q}_t\in\Pi} c_\theta(x(S_t),\mathbb{Q}_t) + \mathbb{E}^{\mathbb{Q}_t}[v_\theta^{(k-1)}(S_{t+1})].$$

It follows from a straightforward induction argument that $v_\theta^{(k)}(S)$ is monotonically increasing in k. Since all feasible policies x in (EC.9) are proper under the corresponding worst-case distributions, we may choose any feasible policy $\tilde{x}$ to get a uniform upper bound

$$v_\theta^{(k)}(S) \leq \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\left[\sum_{t=0}^{\tau} c_\theta(\tilde{x}(S_t),\mathbb{Q}_t)\middle|S_0 = S^{\min}\right] < \infty \tag{EC.11}$$

for all k and S. Hence, $v_\theta(S) = \lim_{k\to\infty} v_\theta^{(k)}(S)$ holds for all $S \in \mathcal{S}$ by the Monotone Convergence Theorem. Uniform convergence then holds by Dini's Theorem under the given assumptions. Hence,

$$\begin{aligned} v_\theta(S_t) &= \lim_{k\to\infty} \min_{0\leq x\leq S_t-S^{\min}} \sup_{\mathbb{Q}_t\in\Pi} c_\theta(x(S_t),\mathbb{Q}_t) + \mathbb{E}^{\mathbb{Q}_t}[v_\theta^{(k-1)}(S_{t+1})] \\ &= \min_{0\leq x\leq S_t-S^{\min}} \sup_{\mathbb{Q}_t\in\Pi} c_\theta(x(S_t),\mathbb{Q}_t) + \mathbb{E}^{\mathbb{Q}_t}[v_\theta(S_{t+1})], \end{aligned} \tag{EC.12}$$

where interchanging the limit and expectation is justified by the Bounded Convergence Theorem based on the uniform bound given in (EC.11).

Again, all feasible policies are proper in (EC.12) (under the worst-case distributions) due to Assumption 1 on $\mathbb{P}$. Hence, uniqueness of the solution to the minimax Bellman equation (EC.12) follows from Bertsekas and Tsitsiklis (1991, Proposition 1) and the Banach Fixed-Point Theorem.

Finally, a direct application of the Donsker and Varadhan variational formula yields

$$\begin{aligned} v_\theta(S_t) &= \min_{0\le x\le S_t - S^{\min}} c(x) + \sup_{\mathbb{Q}_t\in\Pi} \mathbb{E}^{\mathbb{Q}_t}\Big[v_\theta(\min\{S_t+\xi-x, C\})\Big] - \theta\cdot D_{\mathrm{KL}}(\mathbb{Q}_t\|\mathbb{P}) \\ &= \min_{0\le x\le S_t - S^{\min}} c(x) + \theta\log\left(\mathbb{E}^{\mathbb{P}}\left[e^{\frac{v_\theta(\min\{S_t+\xi-x,C\})}{\theta}}\right]\right). \end{aligned}$$

□

Proof of Proposition 3. For ease of notation, define the state update function $g(S,x,\xi) := \min\{C, S+\xi-x\}$. Denote the density of the inflow distribution $\mathbb{P}$ by $p(\cdot)$. Notice from a simple case differentiation that for a convex function $h(y)$ which is decreasing on $[S^{\min}, C]$, it holds that $\tilde{h}(y) := h(\min\{C,y\})$ is also convex and non-increasing. We now show the convexity and monotonicity of the optimal value function in an inductive way, by using the value iteration scheme. Define $f_{v_\theta}(S,x) := c(x) + \theta\log\left(\mathbb{E}\left[e^{\frac{v_\theta(g(S,x,\xi))}{\theta}}\right]\right)$, such that for the optimal value function $v^*_\theta(\cdot)$ it holds $v^*_\theta(S) = \min_x f_{v^*_\theta}(S,x)$. Let $v_k(\cdot)$ be a convex, decreasing function. Then, for any pair of states $S_1, S_2 \le C$, feasible release decisions x_1, x_2, respectively, and $\lambda\in[0,1]$, it holds that

$$\begin{aligned} & f_{v_k}(\lambda S_1 + (1-\lambda)S_2, \lambda x_1 + (1-\lambda)x_2) \\ =\ & c(\lambda x_1 + (1-\lambda)x_2) + \theta\log\left(\int e^{\frac{v_k(g(\lambda S_1+(1-\lambda)S_2, \lambda x_1+(1-\lambda)x_2, \xi(\omega)))}{\theta}} p(\omega)\,d\omega\right) \\ \le\ & \lambda c(x_1) + (1-\lambda)c(x_2) + \theta\log\left(\int e^{\lambda\frac{v_k(g(S_1,x_1,\xi(\omega)))}{\theta} + (1-\lambda)\frac{v_k(g(S_2,x_2,\xi(\omega)))}{\theta}} p(\omega)\,d\omega\right) \\ =\ & \lambda c(x_1) + (1-\lambda)c(x_2) + \theta\log\left(\int \left(e^{\frac{v_k(g(S_1,x_1,\xi(\omega)))}{\theta} + \log(p(\omega))}\right)^{\lambda}\left(e^{\frac{v_k(g(S_2,x_2,\xi(\omega)))}{\theta} + \log(p(\omega))}\right)^{(1-\lambda)} d\omega\right) \\ \le\ & \lambda c(x_1) + (1-\lambda)c(x_2) + \theta\log\left(\left(\int e^{\frac{v_k(g(S_1,x_1,\xi(\omega)))}{\theta} + \log(p(\omega))} d\omega\right)^{\lambda}\cdot\left(\int e^{\frac{v_k(g(S_2,x_2,\xi(\omega)))}{\theta} + \log(p(\omega))} d\omega\right)^{1-\lambda}\right) \\ =\ & \lambda\cdot f_{v_k}(S_1,x_1) + (1-\lambda)\cdot f_{v_k}(S_2,x_2), \end{aligned}$$

i.e., the function $f_{v_k}(\cdot,\cdot)$ is jointly convex. The first inequality followed directly from the convexity of $c(\cdot)$ and $v_k(\cdot)$ (and the fact that the logarithm is a monotonic function), and the second inequality followed from Hölder's inequality. Joint convexity of $f_{v_k}(\cdot,\cdot)$ is a sufficient condition for the convexity of $v_{k+1}(S) := \min_x f_{v_k}(S,x)$:

$$\begin{aligned} \lambda v_{k+1}(S_1) + (1-\lambda)v_{k+1}(S_2) &= \lambda f_{v_k}(S_1, x^*_{k,1}) + (1-\lambda) f_{v_k}(S_2, x^*_{k,2}) \\ &\ge f_{v_k}(\lambda S_1 + (1-\lambda)S_2, \lambda x^*_{k,1} + (1-\lambda)x^*_{k,2}) \\ &\ge \min_x f_{v_k}(\lambda S_1 + (1-\lambda)S_2, x) = v_{k+1}(\lambda S_1 + (1-\lambda)S_2), \end{aligned}$$

where $x^*_{k,i} \in \operatorname{argmin}_x f_{v_k}(S_i, x)$, for $i = 1, 2$. Moreover, if $v_k(S_1) > v_k(S_2)$ holds for any $S_1 > S_2$, then it also holds

$$\begin{aligned} v_{k+1}(S_1) &= c(x^*_{k,1}) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_k(g(S_1, x^*_{k,1}, \xi))}{\theta}}\right]\right) > c(x^*_{k,1}) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_k(g(S_2, x^*_{k,1}, \xi))}{\theta}}\right]\right) \\ &\geq \min_x c(x) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_k(g(S_2, x, \xi))}{\theta}}\right]\right) = v_{k+1}(S_2). \end{aligned}$$

Hence, starting from a convex, decreasing function $v_0(\cdot)$, performing the value iteration $v_{k+1}(S) = \min_x f_{v_k}(S, x)$ preserves the convexity and monotonicity property in each iteration step. Since the value iteration is guaranteed to converge to the optimal value function $v^*(S)$ by Proposition 2, the assertions are shown. □

Proof of Proposition 4 . It follows from the proof of Proposition 3, that the function $f(y) := \theta \log\left(\mathbb{E}\left[e^{\frac{v(\min\{C, y+\xi\}}{\theta}}\right]\right)$ is convex and decreasing on $[S^{\min}, C]$. Thus, its derivative $f'(y)$ is increasing. If $f'(S^{\min}) > -\kappa$, define $S^* := S^{\min}$. If $f'(S^{\max}) < -\kappa$, define $S^* := S^{\max}$. If $f'(S^{\min}) \leq -\kappa$ and $f'(S^{\max}) \geq -\kappa$, define S^* as the argument of $f(\cdot)$ for which $f'(S^*) = -\kappa$ holds. Then, $f'(S) < -\kappa$ holds for all $S < S^*$ and $f'(S) > -\kappa$ holds for all $S > S^*$. Let $S^{\min} \leq S_1 < S_2 \leq S^*$. Then it holds $(f(S_2) - f(S_1))/(S_2 - S_1) < -\kappa$, which implies

$$f(S_2) - f(S_1) < -\kappa(S_2 - S_1) = \kappa(D - (S_2 - S_1)) - \kappa(D - 0) \leq c(S_2 - S_1) - c(0),$$

which is equivalent to $c(0) + f(S_2) < c(S_2 - S_1) + f(S_1)$. Hence, $x = 0$ is better than any other feasible decision given the state S_2, which shows the optimality of releasing zero in this case.

Now let $S^* \leq S_1 \leq S_2 \leq \min\{S^* + D, C\}$. For $S_1 > S^*$, it holds $(f(S_1) - f(S^*))/(S_1 - S^*) > -\kappa$, which implies

$$f(S_1) - f(S^*) > -\kappa(S_1 - S^*) = \kappa(D - (S_2 - S^*)) - \kappa(D - (S_2 - S_1)),$$

which is equivalent to $\kappa(D - (S_2 - S_1)) + f(S_1) > \kappa(D - (S_2 - S^*)) + f(S^*)$. Hence, the optimal release decision in this case is given by $x^*(S_2) = S_2 - S^*$.

For $C \geq S_2 > S^* + D$, the same argument applies for the optimality of $x^*(S_2) = D$ when replacing S^* by $S_2 - D$. For any $x > D$, it holds $c(D) = c(x)$ and hence $c(D) + f(S_2 - D) < c(x) + f(S_2 - x)$ by the monotonicity of f. □

Proof of Proposition 5. As in the proof of Proposition 3 above, we exploit the value iteration scheme. Assume a value function $v_{\theta,k}(S)$, which is decreasing in θ, for any given state S. For ease of notation, define the state update function $g(S,x,\xi) := \min\{C, S+\xi-x\}$ and the objective function $f_S^k(x,\theta) := c(x) + \sup_{\mathbb{Q}\ll\mathbb{P}}\{\mathbb{E}^{\mathbb{Q}}[v_{\theta,k}(g(S,x,\xi))] - \theta D_{\mathrm{KL}}(\mathbb{Q}\|\mathbb{P})\}$ in the dynamic programming formulation (5). Then, the value iteration is given by $v_{\theta,k+1}(S) = \min_x f_S^k(x,\theta)$. For a given value of $\theta=\theta_0$ and a given decision x, let $\mathbb{Q}_0^{k,*}(x)$ be the optimal distribution for the maximization in the definition of $f_S^k(x,\theta)$. By (EC.10), such a distribution $\mathbb{Q}_0^{k,*}(x)$ exists uniquely and the supremum is attained. Now consider $\theta_1 > \theta_0$. If there existed a distribution $\mathbb{Q}_1(x)$ such that $-\theta_1 D_{\mathrm{KL}}(\mathbb{Q}_1(x)\|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}_1(x)}[v_{\theta_1,k}(g(S,x,\xi))] \geq -\theta_0 D_{\mathrm{KL}}(\mathbb{Q}_0^{k,*}(x)\|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}_0^{k,*}(x)}[v_{\theta_0,k}(g(S,x,\xi))]$ holds, then, as KL divergence is non-negative and $v_{\theta,k}(\cdot)$ is decreasing in θ,

$$\begin{aligned}
-\theta_0 D_{\mathrm{KL}}(\mathbb{Q}_1(x)\|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}_1(x)}[v_{\theta_0,k}(g(S,x,\xi))] &> -\theta_1 D_{\mathrm{KL}}(\mathbb{Q}_1(x)\|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}_1(x)}[v_{\theta_0,k}(g(S,x,\xi))] \\
&> -\theta_1 D_{\mathrm{KL}}(\mathbb{Q}_1(x)\|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}_1(x)}[v_{\theta_1,k}(g(S,x,\xi))] \\
&\geq -\theta_0 D_{\mathrm{KL}}(\mathbb{Q}_0^*(x)\|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}_0^*(x)}[v_{\theta_0,k}(g(S,x,\xi))],
\end{aligned}$$

which contradicts the optimality of $\mathbb{Q}_0^*(x)$ for $\theta=\theta_0$. Thus, for any given x, the term $\sup_{\mathbb{Q}\ll\mathbb{P}}\{-\theta D_{\mathrm{KL}}(\mathbb{Q}\|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}}[v_{\theta,k}(g(S,x,\xi))]\}$ is decreasing in θ. Define $x_{\theta_i}^* := \operatorname{argmin}_x f_S^k(x,\theta_i)$ for $i=0,1$, with $\theta_0 < \theta_1$. Then,

$$v_{\theta_0,k+1}(S) = f_S^k(x_{\theta_0}^*,\theta_0) > f_S^k(x_{\theta_0}^*,\theta_1) \geq \min_x f_S^k(x,\theta_1) = v_{\theta_1,k+1}(S),$$

which shows that the value iteration preserves the monotonicity of the value function. □

Proof of Lemma EC.1. Suppose $\max_{i=1,\dots,N}\delta_i(S_i^{(1)}) > \max_{i=1,\dots,N}\delta_i(S_i^{(2)})$ holds. Consider

$$\Gamma \in \left[\max_{i=1,\dots,N}\delta_i(S_i^{(2)}), \max_{i=1,\dots,N}\delta_i(S_i^{(1)})\right).$$

Then

$$\begin{aligned}
\sum_{i=1}^{N}\min\left\{C_i, S_i^{(2)}+\xi_i\right\} = \sum_{i=1}^{N} C_i &> \sum_{i=1}^{N-1} C_{\pi^a(\delta(S^{(1)}+\xi)),i} + (S^{(1)}+\xi)_{\pi^a(\delta(S^{(1)}+\xi)),N} \\
&\geq \sum_{i=1}^{N}\min\left\{C_i, S_i^{(1)}+\xi_i\right\},
\end{aligned}$$

which contradicts the assumption in (EC.1). □

Proof of Lemma EC.2. The proof uses the notation introduced in EC.0. We start by checking how much water we might lose (due to spilling requirements resulting from exceeding the capacity), when using the different release decisions. Consider an inflow realization Γ such that

$\xi_1 \le C_1 - (S_1 - x_1^*)$. Then none of the reservoirs exceeds its capacity. For x, on the other hand, there might be a loss resulting from a reservoir exceeding its capacity. In any case, the total amount of water in the system when releasing according to x^* will be larger than (or equal to) the total amount of water when following x.

Now consider Γ such that $\xi_1 > C_1 - (S_1 - x_1^*)$. Let $K := \max\{k \le N : (S+\xi-x-C)_{\pi^d,k} \ge 0\}$. Then it holds

$$
\begin{aligned}
&\sum_{i=1}^{N} (S_i+\xi_i-x_i-C_i)^+ - (S_i+\xi_i-x_i^*-C_i)^+ \\
=&\sum_{k=1}^{K} (x^*-x)_{\pi^d(S+\xi-x-C),k} - \sum_{k=K+1}^{N} (S+\xi-x^*-C)_{\pi^d(S+\xi-x-C),k} \\
\ge&\sum_{k=1}^{K} (x^*-x)_{\pi^d(S+\xi-x-C),k} - \sum_{k=K+1}^{N} [(S+\xi-x^*-C)-(S+\xi-x-C)]_{\pi^d(S+\xi-x-C),k} \\
=&\sum_{k=1}^{K} (x^*-x)_{\pi^d(S+\xi-x-C),k} + \sum_{k=K+1}^{N} (x^*-x)_{\pi^d(S+\xi-x-C),k} \\
=&\sum_{i=1}^{N} x_i^* - x_i \\
=&0\,,
\end{aligned}
$$

where the inequality is strict if $K < N$. Using this inequality for the amount of water that potentially needs to be spilled when following x vs. x^*, we obtain that

$$
\begin{aligned}
\sum_{i=1}^{N} \min\{C_i, S_i+\xi_i-x_i^*\} &= \sum_{i=1}^{N} (S_i+\xi_i-x_i^*) - (S_i+\xi_i-x_i^*-C_i)^+ \\
&= \sum_{i=1}^{N} (S_i+\xi_i-x_i) - (S_i+\xi_i-x_i^*-C_i)^+ \\
&\ge \sum_{i=1}^{N} (S_i+\xi_i-x_i) - (S_i+\xi_i-x_i-C_i)^+ \\
&= \sum_{i=1}^{N} \min\{C_i, S_i+\xi_i-x_i\}
\end{aligned}
$$

holds for all Γ such that $\xi_1 > C_1 - (S_1 - x_1^*)$, with strict inequality holding for some. $\square$

Proof of Proposition 6. The proof is based on induction, using the value iteration scheme. Start with a function v_0 that satisfies the statement as well as $v_0(C) = 0$.

Denote the optimal policy given the state $S^{(2)}$ by $x^*_{(2)}$ and the corresponding total release by $X^* = \sum_{i=1}^N x^*_{(2),i}$. Consider any feasible release vector $x_{(1)}$ for $S^{(1)}$ with $\sum_{i=1}^N x_{(1),i} = X^*$. The property (6) then carries over to the state vectors $(\min\{C, S^{(1)} + \xi_i - x_{(1)}\})$ and $(\min\{C, S^{(2)} + \xi_i - x^*_{(2)}\})$ for all Γ, where the $\min$ operator is applied componentwise. By the induction assumption, it then follows for all Γ that

$$v_k\left(\min\left\{C, S^{(1)} + \xi - x_{(1)}\right\}\right) \leq v_k\left(\min\left\{C, S^{(2)} + \xi - x^*_{(2)}\right\}\right). \tag{EC.13}$$

Since the value function $v_k(\cdot)$ is non-negative and the confidence parameter $\theta > 0$, (EC.13) directly translates into

$$e^{\frac{v_k\left(\min\left\{C,S^{(1)}+\xi-x_{(1)}\right\}\right)}{\theta}} \leq e^{\frac{v_k\left(\min\left\{C,S^{(2)}+\xi-x^*_{(2)}\right\}\right)}{\theta}}. \tag{EC.14}$$

If strict inequality holds in (6) for the state vectors $S^{(1)} - x_{(1)}$ and $S^{(2)} - x^*_{(2)}$, then by Lemma EC.1 it implies

$$\max_{i=1,\dots,N} \delta_i(S_i^{(1)} - x_{(1),i}) < \max_{i=1,\dots,N} \delta_i(S_i^{(2)} - x^*_{(2),i}). \tag{EC.15}$$

Then

$$\mathbb{P}\left[\min\left\{C, S^{(1)} + \xi - x_{(1)}\right\} = C\right] = \mathbb{P}\left[\min\left\{C, S^{(2)} + \xi - x^*_{(2)}\right\} = C\right] + \mathbb{P}\Big[\Gamma \in \Omega\Big], \tag{EC.16}$$

where the set

$$\Omega := \left[\max_{i=1,\dots,N} \delta_i(S_i^{(1)} - x_{(1),i}), \max_{i=1,\dots,N} \delta_i(S_i^{(2)} - x^*_{(2),i})\right)$$

has a positive probability. For $\Gamma \geq \max_{i=1,\dots,N} \delta_i(S_i^{(2)} - x^*_{(2),i})$ it holds $v_k(\min\{C, S^{(1)} + \xi - x_{(1)}\}) = v_k(\min\{C, S^{(2)} + \xi - x^*_{(2)}\}) = 0$. For all $\Gamma \in \Omega$, on the other hand, it holds that $v_k(\min\{C, S^{(1)} + \xi - x_{(1)}\}) = 0$ but $v_k(\min\{C, S^{(2)} + \xi - x^*_{(2)}\}) > 0$, which means that

$$e^{\frac{v_k(\min\{C,S^{(1)}+\xi-x_{(1)}\})}{\theta}} = 1 < e^{\frac{v_k(\min\{C,S^{(2)}+\xi-x^*_{(2)}\})}{\theta}}. \tag{EC.17}$$

Combining (EC.14) with (EC.16) and (EC.17), we get

$$\theta \log\left(\mathbb{E}\left[e^{\frac{v_k\left(\min\left\{C,S^{(1)}+\xi-x_{(1)}\right\}\right)}{\theta}}\right]\right) \leq \theta \log\left(\mathbb{E}\left[e^{\frac{v_k\left(\min\left\{C,S^{(2)}+\xi-x^*_{(2)}\right\}\right)}{\theta}}\right]\right), \tag{EC.18}$$

with a strict inequality if the inequality in (6) is strict.

Since $\sum_{i=1}^{N} x_{(1),i} = \sum_{i=1}^{N} x^*_{(2),i}$, the immediate costs for unsatisfied demand

$$c\left(x_{(1)}\right) = c\left(x^*_{(2)}\right) \tag{EC.19}$$

coincide for $x_{(1)}$ and $x^*_{(2)}$. By combining (EC.18) and (EC.19), it follows that

$$\begin{aligned} v_{k+1}(S^{(2)}) &= c\left(x^*_{(2)}\right) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_k\left(\min\left\{C, S^{(2)}+\xi-x^*_{(2)}\right\}\right)}{\theta}}\right]\right) \\ &\geq c\left(x_{(1)}\right) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_k\left(\min\left\{C, S^{(1)}+\xi-x_{(1)}\right\}\right)}{\theta}}\right]\right) \\ &\geq \min_{x \in \mathcal{X}(S^{(1)})} c\left(x_i\right) + \theta \log\left(\mathbb{E}\left[e^{\frac{v_k\left(\min\left\{C, S^{(1)}+\xi-x\right\}\right)}{\theta}}\right]\right) \\ &= v_{k+1}(S^{(1)}), \end{aligned} \tag{EC.20}$$

where the inequality is strict if strict inequality holds in (6) with a positive probability. □

Proof of Corollary EC.1. (Follows from the proof of Proposition 6.) Given that: (i) it holds $\sum_i x_i = \sum_i x^*_i$, (ii) the parameter θ is positive, (iii) the logarithm is a positive monotonic function on $[1, \infty)$, and (iv) its arguments are greater than or equal to one, the assertion is reduced to the expected values on both sides of the inequality. By Proposition 6, it holds

$$v_\theta\left(\min\left\{C, S+\xi-x^*\right\}\right) \leq v_\theta\left(\min\left\{C, S+\xi-x\right\}\right)$$

almost everywhere and strict inequality holds with positive probability. Together with $\theta > 0$, this pathwise inequality for all scenarios and strict inequality for some scenarios then directly imply the assertion for the expected values. □

Proof of Corollary 1. Follows from Corollary EC.1. □

Proof of Proposition 7. Using the notation introduced in EC.0, conditions (ii) and (iii) in the statement of the Proposition can be reformulated as follows:

(ii) there is an index $K \leq N$ such that

- for all $K < i \leq N$ it holds $x^*_{\pi^a(\delta),i} = 0$;
- $\exists \mathcal{I} \subseteq \{1, \ldots, K\}$ such that:
  - —$\forall i, j \in \mathcal{I}$:
    - $\delta_{\pi^a(\delta(S)),i}\big((S-x^*)_{\pi^a(\delta(S)),i}\big) = \delta_{\pi^a(\delta(S)),j}\big((S-x^*)_{\pi^a(\delta(S)),j}\big)$;
  - —$\forall k \in \{1, \ldots, K\} \backslash \mathcal{I}$:

* $(S-x^*)_{\pi^a(\delta(S)),k} = S^{\min}_{\pi^a(\delta(S)),k}$
* $\delta_{\pi^a(\delta(S)),k}\big((S-x^*)_{\pi^a(\delta(S)),k}\big) \le \delta_{\pi^a(\delta(S)),i}\big((S-x^*)_{\pi^a(\delta(S)),i}\big)$ for all $i \in \mathcal{I}$;

(iii) there is no other vector x' satisfying (i) and (ii) for an index $K < K' \le N$.

For any feasible $x' \not\equiv x^*$ with $\sum_{i=1}^K x'_{\pi^a(\delta(S)),i} = \sum_{i=1}^N x'_i = X^*$, it follows directly from the proof of Lemma EC.2 that

$$\sum_{i=1}^N \min\{C_i, S_i + \xi_i - x^*_i\} \ge \sum_{i=1}^N \min\{C_i, S_i + \xi_i - x'_i\}$$

holds almost everywhere and therefore x^* is favorable by Corollary EC.1. Now consider some feasible x' with $\sum_{i=1}^N x'_i = X^*$ but $\sum_{i=1}^K x'_{\pi^a(\delta(S)),i} < X^*$. For any inflow $\Gamma \le \max_{i=1,\dots,K} \delta_{\pi^a(\delta(S)),i}\big((S-x^*)_{\pi^a(\delta(S)),i}\big)$, no water needs to be spilled when releasing according to x^* and therefore $\sum_{i=1}^N \min\{C_i, S_i+\xi_i - x^*_i\} \ge \sum_{i=1}^N \min\{C_i, S_i+\xi_i-x'_i\}$ holds. Now suppose $\Gamma > \max_{i=1,\dots,K} \delta_{\pi^a(\delta(S)),i}\big((S-x^*)_{\pi^a(\delta(S)),i}\big)$. There is a (nonempty) index set $\mathcal{J} \subseteq \{1,\dots,K\}$ such that $x'_{\pi^a(\delta(S)),i} < x^*_{\pi^a(\delta(S)),i}$ for all $i \in \mathcal{J}$ and $x'_{\pi^a(\delta(S)),i} \ge x^*_{\pi^a(\delta(S)),i}$ for all $i \in \{1,\dots,N\}\setminus\mathcal{J}$. This implies that:

- For each $i \in \mathcal{J}$, water needs to be spilled for both x^* and x':

$$\sum_{i\in\mathcal{J}} \big((S+\xi-x'-C)_{\pi^a(\delta(S)),i}\big)^+ - \big((S+\xi-x^*-C)_{\pi^a(\delta(S)),i}\big)^+ = \sum_{i\in\mathcal{J}} (x^*-x')_{\pi^a(\delta(S)),i}\,.$$

- For each $i \in \mathcal{J}^c := \{1,\dots,N\}\setminus\mathcal{J}$, either both strategies require to spill or only x^*:

$$\sum_{i\in\mathcal{J}^c} \big((S+\xi-x'-C)_{\pi^a(\delta(S)),i}\big)^+ - \big((S+\xi-x^*-C)_{\pi^a(\delta(S)),i}\big)^+ \ge \sum_{i\in\mathcal{J}^c} (x^*-x')_{\pi^a(\delta(S)),i}\,.$$

Using the resulting inequality for the sum over all $i = 1,\dots,N$, we get

$$\begin{aligned}
&\sum_{i=1}^N \min\{C_i,(S+\xi-x^*)_i\} - \min\{C_i,(S+\xi-x')_i\}\\
&= \sum_{i=1}^N (S+\xi-x^*)_i - ((S+\xi-x^*-C)_i)^+ - (S+\xi-x')_i + ((S+\xi-x'-C)_i)^+\\
&= \sum_{i=1}^N (x'_i - x^*_i) + \sum_{i=1}^N ((S+\xi-x'-C)_i)^+ - ((S+\xi-x^*-C)_i)^+\\
&\ge \sum_{i=1}^N (x'_i - x^*_i) + \sum_{i=1}^N (x^*_i - x'_i)\\
&= 0\,,
\end{aligned}$$

from which it follows that

$$\sum_{i=1}^{N} \min\{C_i, (S+\xi-x^*)_i\} \geq \sum_{i=1}^{N} \min\{C_i, (S+\xi-x')_i\}$$

also holds for any inflow $\Gamma > \max_{i=1,\ldots,K} \delta_{\pi^a(\delta(S)),i}((S-x^*)_{\pi^a(\delta(S)),i})$. By Corollary EC.1, x^* is therefore better than x' also in this case.

The subcondition involving $S^{\min}$ ensures feasibility of x^*. Moreover, if there existed an index $i \in \{1,\ldots,K\}\backslash\mathcal{I}$ such that $(S-x^*)_{\pi^a(\delta(S)),i} > S^{\min}_{\pi^a(\delta(S)),i}$, then, using the same argument as above based on Corollary EC.1, one can construct a better strategy that does satisfy the condition. □

Proof of Proposition 8. The proof is based on a pathwise argument. The initial state is not limited to $S_i^{(N)} \equiv C_i$, the result holds for all states. Let $S_0^{(1)} := \sum_{i=1}^N S_{0,i}^{(N)}$. Denote the optimal release decision for the N-reservoirs system by $x^{(N)}(S_0^{(N)})$. Define the corresponding release from the aggregate reservoir as $X(S_0^{(1)}) := \sum_{i=1}^N x_i^{(N)}(S_0^{(N)})$. By construction, $X(S_0^{(1)})$ is feasible and the shortage costs for the initial stage coincide: $c(X(S_0^{(1)})) = c(\sum_{i=1}^N x_i^{(N)})$.

As for the realized total first-stage inflow $\xi_0^{(1)} = \sum_{i=1}^N \xi_{0,i}$, we distinguish two cases:

(i) $\xi_0^{(1)} < C^{(1)} - (S_0^{(1)} - X)$:

$$\min\{C^{(1)}, S_0^{(1)} - X + \xi_0^{(1)}\} = S_0^{(1)} - X + \xi_0^{(1)} \;\geq\; \sum_{i=1}^{N} \min\{C_i, S_{0,i}^{(N)} - x_i^{(N)} + \xi_{0,i}\}\,;$$

(ii) $\xi_0^{(1)} \geq C^{(1)} - (S_0^{(1)} - X)$:

$$\min\{C^{(1)}, S_0^{(1)} - X + \xi_0^{(1)}\} = C^{(1)} = \sum_{i=1}^{N} C_i \;\geq\; \sum_{i=1}^{N} \min\{C_i, S_{0,i}^{(N)} - x_i^{(N)} + \xi_{0,i}\}\,.$$

In both cases, we see that the single, aggregate reservoir ends up at time 1 with at least as much water as the aggregate amount of water contained in the N-reservoirs system. If $S_0^{(N)}$ is an unbalanced state, then there is a positive probability that the aggregate reservoir ends up with strictly more water than the N-reservoirs system. At time 1, it is then again feasible to replicate the N-reservoirs system's optimal policy and release an amount of $X(S_1^{(1)}) := \sum_{i=1}^N x_i^{(N)}(S_1^{(N)})$ from the aggregate reservoir.

The argument applies iteratively along each path of replicated releases and inflow realizations, until the aggregate reservoir reaches its capacity. Until this event occurs, the accumulated shortage costs for the two systems coincide, based on identical total releases. By construction, there is no path where the N-reservoirs system reaches its terminal state before the aggregate reservoir exceeds its capacity. If the initial state $S_0^{(N)}$ is unbalanced, then for some paths (i.e., with positive probability), the aggregate reservoir exceeds its capacity at a point in time when the N-reservoirs system has not reached its terminal state yet.

When replicating the optimal release policy of the N-reservoirs system, the expected total shortage costs for the aggregate reservoir (until it reaches its capacity) are therefore smaller than or equal to the expected total shortage costs for the N-reservoirs system (until it reaches its terminal state). This consequently also applies to the optimal expected costs for the aggregate reservoir, which proves the assertion. □

Proof of Proposition 9. The bound follows analogously to the proof of Prop. 8. The key is to observe that along each path of inflow realizations Γ_t, the N-reservoirs system at each time t collects at least as much inflow as the virtual aggregated reservoir constructed in this proposition.

By construction of the restricted inflow based on the least favorable state, the given policy for the aggregated setting is still a proper policy. By Lemma 3.1 of Bertsekas and Tsitsiklis (1989), the value of the policy evaluation problem (7) for the aggregated reservoir is therefore given by the unique fix point of (8). □

Proof of Proposition EC.1. Analogous to the proof of Proposition 8, a pathwise argument can be used. The key is that along each path of inflow realizations Γ_t, the aggregated reservoir of Prop. 8 in each step receives at least as much inflow as the aggregated reservoir of Prop. EC.1, and the latter one receives as least as much inflow as the corresponding N-reservoir system. □

Proof of Lemma EC.3. For any given $K \geq 0$, define

$$\lambda^K(h,y) := \sum_{t=0}^{\min\{\tau-1,K\}} z_t \left(y_t(S_{t+1})h(x(S_t),S_{t+1}) - \mathbb{E}_{\mathbb{P}}\left[y_t(S_{t+1})h(x(S_t),S_{t+1})|\mathcal{F}_t\right]\right)$$

and

$$\tilde{\lambda}^K(h,y) := \sum_{t=0}^{K} z_t \left(y_t(S_{t+1})h(x(S_t),S_{t+1}) - \mathbb{E}_{\mathbb{P}}\left[y_t(S_{t+1})h(x(S_t),S_{t+1})|\mathcal{F}_t\right]\right).$$

It follows from applying the law of iterated expectation, that $\mathbb{E}[\tilde{\lambda}^K(h,y)|z_0,y_0] = 0$ holds. By the law of total probability, this also translates into $\mathbb{E}[\lambda^K(h,y)|z_0,y_0] = 0$ for any given K. It remains to show that this also holds in the limit as $K \to \infty$.

First, notice that the end of a cycle has finite expectation:

$$\mathbb{E}[\tau] = \sum_{n=1}^{\infty} n \cdot \mathbb{P}[\Gamma \geq \max_{i=1,\dots,N} \delta_i(S_n)] \cdot \prod_{t=1}^{n-1} \mathbb{P}[\Gamma < \max_{i=1,\dots,N} \delta_i(S_t)] \leq \sum_{n=1}^{\infty} n \cdot c_\xi^{n-1} = \frac{1}{(c_\xi - 1)^2} < \infty\,, \tag{EC.21}$$

since $c_\xi := \mathbb{P}[\Gamma < \max_{i=1,\dots,N} \delta_i(S^{\min})\}]$ is bounded away from 1 by Assumption 1.

Define

$$\Lambda_n := \sum_{t=0}^{n-1} z_t \left(y_t(S_{t+1}) h\left(x(S_t), S_{t+1}\right) - \mathbb{E}_{\mathbb{P}}\left[y_t(S_{t+1}) h(x(S_t), S_{t+1}) | \mathcal{F}_t\right]\right)$$

to consider $\lambda(h, y)$ as a stopped stochastic process $\Lambda = (\Lambda_n)_{n \geq 1}$, stopped at time τ. Then Λ_n and z_n are measurable with respect to the sigma algebra $\mathcal{F}_n := \sigma(S_n, y_{n-1})$. It follows directly that Λ is a martingale:

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[\Lambda_{n+1} | \mathcal{F}_n] &= \mathbb{E}_{\mathbb{P}}\left[\Lambda_n + z_n \cdot (y_n(S_{n+1}) h(S_{n+1}) - \mathbb{E}_{\mathbb{P}}\left[y_n(S_{n+1}) h(S_{n+1}) | \mathcal{F}_n\right]) | \mathcal{F}_s\right] \\ &= \Lambda_n + z_n \cdot (\mathbb{E}_{\mathbb{P}}\left[y_n(S_{n+1}) h(S_{n+1}) | \mathcal{F}_n\right] - \mathbb{E}_{\mathbb{P}}\left[y_n(S_{n+1}) h(S_{n+1}) | \mathcal{F}_n\right]) = \Lambda_n\,.\end{aligned} \tag{EC.22}$$

Since $h(\cdot, \cdot)$ is bounded and $y_t, z_t \in (0, 1]$, using the triangle inequality, it follows that the increments of λ are uniformly bounded:

$$\begin{aligned}|\Lambda_{n+1}(\omega) - \Lambda_n(\omega)| &\leq |y_n(S_{n+1}(\omega)) h(x(S_n(\omega)), S_{n+1}(\omega))| + |\mathbb{E}_{\mathbb{P}}\left[y_t(S_{n+1}) h(x(S_n), S_{n+1}) | \mathcal{F}_n\right]| \\ &\leq 2 c_h\,.\end{aligned} \tag{EC.23}$$

Given (EC.21), (EC.22), and (EC.23), it follows from Doob's Optional Stopping Theorem that $\mathbb{E}[\lambda(h, y)] = \mathbb{E}[\Lambda_\tau] = \mathbb{E}[\Lambda_1] = 0$. □