ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research®

RESEARCH ARTICLE
10.1029/2023WR036724

Reinforcement Learning of Multi-Timescale Forecast Information for Designing Operating Policies of Multi-Purpose Reservoirs

D. Zanutto[1,2], A. Ficchì[1], M. Giuliani[1], and A. Castelletti[1]

[1]Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy, [2]KWR Water Research Institute, Nieuwegein, The Netherlands

Key Points:

- Multi-timescale hydro-meteorological forecast information helps in multi-purpose reservoir operations
- A new methodology is proposed to learn the most valuable information from multiple forecasts and jointly design operating policies
- With the available operational products, more skillful short-range lead times are preferred to ideally more informative extended-range ones

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Castelletti,
andrea.castelletti@polimi.it

Citation:

Zanutto, D., Ficchì, A., Giuliani, M., & Castelletti, A. (2025). Reinforcement learning of multi-timescale forecast information for designing operating policies of multi-purpose reservoirs. *Water Resources Research*, *61*, e2023WR036724. https://doi.org/10.1029/2023WR036724

Received 21 NOV 2023
Accepted 30 JAN 2025

Author Contributions:

Conceptualization: D. Zanutto, A. Ficchì, M. Giuliani, A. Castelletti
Data curation: D. Zanutto
Formal analysis: D. Zanutto
Funding acquisition: A. Castelletti
Investigation: D. Zanutto, A. Ficchì
Methodology: D. Zanutto, A. Ficchì, M. Giuliani, A. Castelletti
Project administration: A. Castelletti
Resources: A. Castelletti
Software: D. Zanutto
Supervision: A. Ficchì, A. Castelletti
Validation: D. Zanutto, A. Ficchì
Writing – original draft: A. Ficchì, M. Giuliani, A. Castelletti

Abstract Hydrological forecasts have significantly improved in skill over recent years, encouraging their systematic exploitation in multipurpose reservoir operations to improve reliability and resilience to extreme events. Despite the growing availability of multi-timescale forecasts, there is still a lack of transparent and integrated methods for selecting the most suitable forecast products, variables, and lead times for specific operational challenges. In this work, we propose a holistic approach based on Reinforcement Learning (RL) to design multipurpose dam operating policies informed by available multi-timescale forecast products. Our approach extends the traditional Evolutionary Multi-Objective Direct Policy Search method by parametrizing both the operating policy and the forecast information extraction process. We compare our RL approach with a state-of-the-art two-step procedure in which the forecast selection and processing are performed before the policy optimization. We demonstrate the value of the method for the multipurpose operation of Lake Como (Italy) by considering multi-timescale forecasts from short to seasonal lead times to manage flood- and drought-related operational objectives. Our approach identifies solutions achieving an 18% improvement in hypervolume indicator compared to policies not informed by forecasts and a 6% improvement over those designed using the two-step reference methodology. These improvements are accompanied by increased flexibility in policy design and trade-off analysis by directly extracting forecast information within the multi-objective optimization. This study demonstrates the feasibility and benefits of integrating policy design with forecast information extraction, particularly when multiple operational forecasts are available.

1. Introduction

Water reservoir operations face challenges posed by hydroclimatic variability and more frequent climate extremes (Stevenson et al., 2022), as well as by an increasing pressure driven by economic and population growth. To cope with these evolving physical and socioeconomic changes, adaptive reservoir management can use increasingly skillful hydrometeorological and seasonal climate forecasts (Arnal et al., 2018; Bauer et al., 2015) in combination with feedback and feedforward control schemes to design optimal operations (Castelletti et al., 2008; Giuliani et al., 2021).

In recent years, hydrometeorological and climate forecasting have seen a marked improvement across a range of timescales, with more and more products available (e.g., Meehl et al., 2021; Wetterhall & Di Giuseppe, 2018). Several studies have shown that more skillfull forecasts lead to improved water management outcomes (e.g., Giuliani et al., 2020; Lee et al., 2022; Turner et al., 2017; Yang et al., 2020). Yet, the forecast skill–value relationship is complex and influenced by several factors, including the properties of the forecast, such as uncertainty, update frequency, and lead times (e.g., Doering et al., 2021; Guo et al., 2021; Torres et al., 2023; Turner et al., 2017; Yang et al., 2021; T. Zhao et al., 2011); the physical characteristics of the system, such as the hydrological variability, reservoir, size and operational constraints; and the priorities and risk aversion of the decision-makers (e.g., Anghileri et al., 2016; Bertoni et al., 2021; Giuliani et al., 2020; Peñuela et al., 2020; Semmendinger et al., 2022; Yang et al., 2021). Indeed, determining the most suitable forecast information for a given problem requires careful evaluation of several factors. The user has to make many crucial decisions, such as selecting the forecast product, its lead time, the variables to extract, and the type of information (e.g., deterministic or probabilistic). These choices often lack transparent reporting of operational rules and guidelines to support critical decisions, which are usually reservoir-specific or based on the operator experience (Turner et al., 2020). An additional challenge lies in the multi-objective nature of the problem, as different Pareto-optimal solutions may require different decisions.

Writing – review & editing: A. Ficchì, M. Giuliani, A. Castelletti

Among these decisions, one that has been studied extensively is the choice of the optimal forecast lead time for a single product. Studies by You and Cai (2008), T. Zhao et al. (2012), Denaro et al. (2017), Q. Zhao et al. (2019), and Turner et al. (2020) have all looked into this aspect, examining forecast horizon selection from either hypothetical perfect forecasts, synthetic (statistical) forecasts, or real forecasts. There is, however, a lack of studies investigating general methods to support multiple choices of forecast information extraction beyond the forecast horizon, especially when multi-timescale products with different characteristics are available. Moreover, only the work of Zaniolo et al. (2021b) investigated the relationship between the objective trade-off and the best input for the operating policies, demonstrating that different information is needed for different objective trade-offs; still, the only candidate inputs they considered were perfect inflow forecasts at different lead times. To respond to these challenges, we propose a new approach to support the joint extraction of forecast information (from real multi-timescale products) and policy design. While previous studies have treated these two inter-linked tasks as sequential steps (Giuliani et al., 2015; Yang et al., 2017; Zaniolo et al., 2021b), we demonstrate that integrating the information extraction within the policy design task is beneficial to respond to the multi-objective nature of water management problems.

There are several reservoir control algorithms that can be used to condition reservoir operating policies based on forecast (or exogenous) information. These algorithms can be classified into two categories: real-time control and offline policy design. The natural approach for adaptive water management based on forecasts is Model Predictive Control (MPC; Bertsekas, 2005), a class of real-time control algorithms that have been increasingly adopted for water reservoir operations over recent years (Castelletti et al., 2023). However, MPC is susceptible to two key challenges: (a) by directly feeding forecasts into a system's model, its performance is impacted by forecast biases and inaccuracies (e.g., Arsenault & Côté, 2019); (b) it is inherently single-objective, thus disregarding the multi-objective nature of water systems (Castelletti et al., 2023). To overcome these challenges, a solution is to use offline, data-driven, multi-objective approaches that can use forecasts, like Reinforcement Learning (RL) algorithms (Giuliani et al., 2021). In RL, agents learn (or design) an optimal operating policy collecting experience from an environment via iterative trial-and-error exploration of the state-decision space (Sutton & Barto, 2020). Applied to water reservoir control, the learning environment is based on historical data and on simulation models of the coupled river-reservoir systems, as exploring the real system is unfeasible (Rieker & Labadie, 2012; Sutton & Barto, 2020). The need to manage exogenous information in a data-driven manner and for multi-objective problems significantly narrows the field of the approaches that can be used in this context. Evolutionary Multi-Objective Direct Policy Search (EMODPS; Giuliani, Castelletti, et al., 2016) meets these requirements, as it relies on the parametrization of the policy within a given class of functions that can easily be extended to include exogenous signals (forecasts or observations), and a multi-objective optimization algorithm capable of generating a Pareto front of optimal solutions within a single run. Other RL algorithms that possess the required characteristics have also been successfully applied, for example, Multi-Objective Fitted Q-Iteration (e.g., Castelletti et al., 2013; Pianosi et al., 2013). However, as suggested by the review by Giuliani et al. (2021), EMODPS has significant computational and upscaling capabilities, as it works with an Approximation in Policy Space (Bertsekas, 2019; Powell, 2019), as opposed to MOFQI that works with an Approximation in the Value Space. Therefore, as EMODPS is the most promising multi-objective RL-based algorithm that uses exogenous information, we use it as the starting point for our study, and we advance it by embedding the forecast information extraction task in the policy parametrization.

Previous studies have proposed combining the task of exogenous information extraction with EMODPS within broader frameworks, for example, the Information Selection and Assessment (ISA; Giuliani et al., 2015) and the Selection of Information for NeuroEvolutionary Policy Search (SINEPS; Zaniolo et al., 2021b). However, these two frameworks combine the policy design algorithm with a preliminary feature selection algorithm that selects the most relevant information based on an ideal, pre-computed target solution (obtained by assuming perfect knowledge of the future and dynamic programming). This assumption limits their application to large complex (multi-reservoir) systems due to computational and scalability issues (curse of dimensionality). Moreover, the set of candidate forecast variables used in these approaches is usually restricted to only a sample of the whole explorable space due to the computational costs of the feature selection algorithms, potentially leading to sub-optimal policy performance.

To overcome these challenges, in this paper, we propose a more general and scalable framework based on Reinforcement Learning (RL) that jointly learns how to extract the most valuable information from a set of multi-timescale forecast products and how to use this information for advancing multi-purpose dam policy design.

Specifically, we extend EMODPS to explore a wider decision space, including, along with the traditional policy parameters, additional parameters representing different choices that specify the best candidate forecast product, lead time, and raw or processed variables to use. Our approach is versatile, adaptable, and inherently multi-objective, and it can be easily customized based on the number of forecast products involved and the necessary decisions, independently of the number of objectives and reservoirs. The parameters set can be expanded to include any candidate policy inputs, such as forecasts, raw observational data, or any processed variables derived by applying a parametrized function to these data (such as temporal aggregation or bias adjustment). The strength of this approach is that the information extraction is wholly integrated within the multi-objective policy design, producing a seamless RL approach that can extract the most valuable information for different Pareto optimal tradeoffs. Therefore, compared to existing approaches that first apply feature extraction algorithms and then implement a policy design algorithm (e.g., Giuliani et al., 2015; Yang et al., 2017; Zaniolo et al., 2021b), our new integrated learning approach is more flexible and widely applicable. Moreover, it has the advantage of directly optimizing the extraction of forecast information jointly with the policy within the same multi-objective problem formulation and in a single step. These novel, useful features differentiate our framework from the existing two-step approaches and bring three advantages: (a) eliminating the need to pre-compute target ideal solutions using dynamic programming, overcoming computational and scalability challenges for large systems; (b) eliminating the need for pre-sampling forecast information using feature selection algorithms, to help navigate across increasingly extensive and diverse multi-scale forecast data sets; (c) eliminating the need of multiple runs and steps, proposing an inherently multi-objective optimization that can jointly explore the decisions of both policy parameters and which forecast information to extract and use.

The proposed RL approach is demonstrated on the Lake Como system (Italy), a regulated lake operated primarily for flood control and water supply. The subalpine basin of the lake is characterized by mixed slow and fast dynamics resulting from its snow- and rain-dominated hydrology (Denaro et al., 2017). The lake's operations affect several stakeholders, including the population living along the lake shores, who are exposed to flood risk, and downstream users, including farmers and hydropower companies, who are served by the lake releases and are exposed to drought risk (Giuliani, Li, et al., 2016). In the context of such dual temporal dynamics, forecasts on both short and seasonal time scales may be valuable, as demonstrated by previous studies working on perfect forecasts only (Denaro et al., 2017; Zaniolo et al., 2021b). However, nowadays the lake operator has access to different operational products: short-term (i.e., 60-hr lead time) deterministic forecasts produced with locally calibrated models, as well as continental-scale sub-seasonal and seasonal forecasts of the Copernicus Emergency Management Service's European Flood Awareness System (EFAS) (Barnard et al., 2020; Wetterhall et al., 2020). These continental forecasts are ready-to-use but are known to have limited skill in the Alpine region (Wetterhall & Di Giuseppe, 2018), partly due to little model calibration and large biases. Thus, considering these operational multi-timescale forecast products, we design a set of forecast-informed policies for the multi-purpose operation of the dam situated at the outlet of the lake, jointly learning what information to extract from the forecast set. Given the approximate nature of the proposed RL approach, we compare our solutions with those obtained using the Information Selection and Assessment (ISA) framework (Giuliani et al., 2015), showing the benefits of our new approach.

2. Methods

2.1. Problem Formulation

We consider a general reservoir policy design problem defined over the simulation horizon H and concerning a multidimensional objective function $\mathbf{J}$ formulated as:

$$p^* = \arg\min_{p} \mathbf{J}\left(p,\ s_0,\ q_{[1,H]}\right) \tag{1}$$

where the initial storage s_0 is given, and $q_{[1,\ H]}$ is the trajectory of the external drivers (e.g., reservoir inflows). The policy p is defined as a closed-loop control law that determines the daily release decision $u_t = p(d_t,\ s_t,\ I_t)$ at each time step $t = 0,\ \dots,\ H - 1$ as a function of the day of the year (d_t), the reservoir storage (s_t), and a vector of exogenous information (I_t). The system evolves according to a state transition function that describes reservoir storage dynamics using a mass balance equation $s_{t+1} = s_t + q_{t+1} - r_{t+1}$, where q_{t+1} is the net inflow entering

10.1029/2023WR036724

the reservoir (i.e., the inflow and precipitation minus evaporation and other losses) and r_{t+1} the actual release, which is determined by the decision u_t (correcting it considering the minimum and maximum release constraints). In the adopted notation, the subscript of a variable indicates the time when its value is deterministically known: the variables q_{t+1} and r_{t+1} are known only at the end of the interval and hence appear with the subscript $t+1$.

Given the complexity of the climate and hydrological systems influencing the net inflow q_{t+1}, the latter is often modeled as a system disturbance: at each time step t, we have at our disposal a set $(\hat{Q})$ of different forecast products ($\gamma \in \Gamma$). Formally, each product can have one or more ensemble members ($i \in [1, \ldots, n_e^\gamma]$, where $n_e^\gamma = 1$ indicates a deterministic forecast), and predicts the net inflow from $q_{\tau+1}$ to $q_{\tau+LT^\gamma}$ (LT^γ is the forecast lead time). In this notation, the subscript $\tau \in \mathrm{T}^\gamma$ represents the time when the forecast has been issued (with T^γ representing the collection of forecast issue dates), as each product has a different update frequency that is not necessarily synchronized with the simulation time step t.

2.2. Forecast Evaluation

To gain insights that support the understanding of the information extraction and policy design results, we evaluate the accuracy of the different inflow forecast products. Here, we introduce the main metrics that we selected for a clear comparison of both deterministic and probabilistic forecasts considered in this study (see Section 3.2).

The Kling-Gupta Efficiency score (KGE; Gupta et al., 2009) is a deterministic metric that quantifies the overall accuracy of forecasts with respect to the corresponding observations. It is defined as:

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{2}$$

where r is the Pearson correlation coefficient between the two time series; $\alpha = \sigma_{pred}/\sigma_{obs}$ is a measure of the relative variability (ratio of standard deviations); and $\beta = \mu_{pred}/\mu_{obs}$ is the ratio between the mean values (i.e., bias indicator). KGE and its three components (r, β, α) have an ideal value of 1. Hence, the KGE score is 1 minus the Euclidean distance to the ideal point in this 3-dimensional space, and the KGE range is $(-\infty, 1]$. For the probabilistic forecasts, we use the ensemble mean for the computation of the KGE and its components.

The Continuous Ranked Probability Score (CRPS; Hersbach, 2000) is a probabilistic verification metric that compares the continuous cumulative distribution of an ensemble forecast with the distribution of observations. It reduces to the Mean Absolute Error (MAE) for a deterministic forecast, allowing a clear interpretation of the relative performance of ensemble and deterministic forecasts. The CRPS is defined as:

$$\mathrm{CRPS} = \frac{1}{H}\sum_{t=1}^{H}\int_{-\infty}^{+\infty}[P_{\hat{q}_t}(x) - P_{q_t}(x)]^2 \mathrm{d}x \tag{3}$$

where $P_{\hat{q}_t}$ and P_{q_t} are the cumulative distribution functions of the forecast and observation data at time step t. The CRPS varies in the range $[0, +\infty)$, with 0 representing a perfect prediction. While the KGE and its components are effective for assessing deterministic forecasts like the ensemble mean (but not fully capturing the probabilistic performance and reliability), the CRPS offers a comprehensive evaluation of the uncertainty and reliability in ensemble forecasts, accounting for their spread, outlier frequency and magnitude, and achieving a perfect score only with perfect deterministic forecasts (Hersbach, 2000).

To further assess the skill of the probabilistic forecasts in predicting low- and high-flow events (i.e., threshold exceedances), we also analyzed the ensemble forecast products with an unbiased version of the Brier Skill Score (UBSS; see Supporting Information S1 for its definition).

2.3. Joint Learning of Forecast Information and Operating Policy

In this section, we illustrate the proposed RL approach (Figure 1b) for the design of the optimal operations of a multipurpose reservoir leveraging the most valuable information (I_t) to be extracted from a set of forecasts available at time t $(\hat{Q}_t)$.

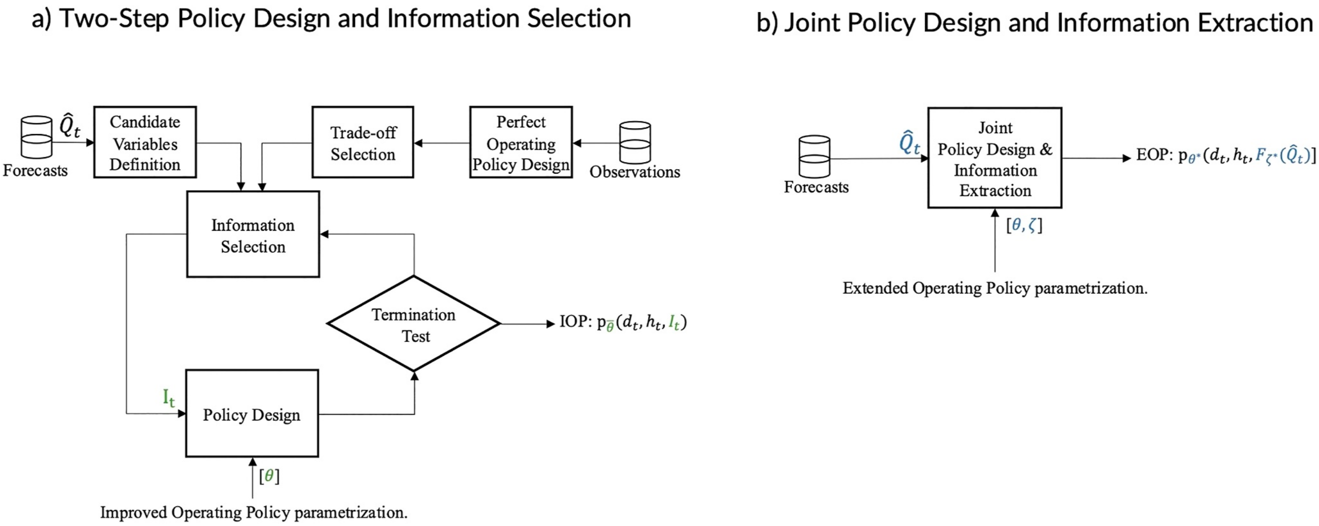

Figure 1. Conceptual flowcharts of (a) the benchmark framework (a generalization of the Information Selection and Assessment framework; Giuliani et al., 2015) and (b) the proposed methodology for the joint learning of forecast information and operating policy. Both approaches rely on the parametrization of the policy as a non-linear approximating network and the same input signals, that is, day of the year (d_t), lake level (h_t), and exogenous information (I_t) derived from the inflow forecasts $(\hat{Q}_t)$. Our new proposed approach embeds the necessary steps to extract the exogenous information as a combination of parametrized functions ($\mathcal{F}_\zeta(\cdot)$).

Specifically, we introduce a generic parametric function $I_t = \mathcal{F}_\zeta(\hat{Q}_t)$ representing the extraction of information from available forecasts, which can include the following operations: (a) selection of the forecast product ($\gamma \in \Gamma$); (b) selection of the effective forecast lead time ($\lambda < LT^\gamma$); (c) selection of the temporal aggregation operator of the forecasts over the selected lead time (Ψ_λ); (d) selection of the operator to deal with the forecast uncertainty (Ψ_{n_e}). Moreover, an implicit operation is always performed to use only the most recent forecast between those available at time t.

For this reason, the formulation of such a parametric information extraction function is then coupled with Direct Policy Search, which is based on the parameterization of the operating policy (p_θ) within a given family of functions and the exploration of the parameter space to find a parameterized policy that optimizes the vector of objective functions (Rückstiess et al., 2010). Combining these two formulations, the daily release decision is now determined as $u_t = p_\theta(d_t, s_t, \mathcal{F}_\zeta(\hat{Q}_t))$ and Problem 1 can be reformulated as finding the best parameters of an Extended Operating Policy (EOP) that will specify both forecast information extraction (ζ^*) and reservoir operation (θ^*):

$$[\theta^*, \zeta^*] = \arg\min_{[\theta, \zeta]} \mathbf{J} \quad \text{s.t. } \theta \in \Theta, \zeta \in Z \tag{4}$$

The parameters ζ do not directly affect the inner operating policy p_θ but indirectly influence its output by altering the input signal I_t and consequently the optimal policy parameters (θ^*). In other words, fixing the policy input signal by setting specific parameters of the information extraction function (e.g., $\zeta = \zeta_1$), yields a particular set of optimal inner-policy parameters ($\theta^*_{\zeta_1}$) that may differ significantly from the ones obtained if a different exogenous signal is extracted (i.e., if $\zeta = \zeta_2$ there is no guarantee that $\theta^*_{\zeta_1} = \theta^*_{\zeta_1}$). Our proposed setup expands the decision space by combining all possible policy parametrizations with all candidate input signals derived from multiple forecasts. As a consequence, while the decision space size increases, the Pareto-front also changes, now containing the actual non-dominated solutions obtained from all combinations of these two critical components (policy parametrizations and candidate input signals), allowing the optimization algorithm to search for these Pareto-optimal combinations directly. This feature offers an advantage compared to approaches that separate the two steps (e.g., ISA; see Figure 1a), that is, selecting the signal I_t first, then optimizing the policy parameters, and iteratively scouting for better combinations. However, this benefit may come at the expense of a larger decision space, which could be more challenging for the optimization algorithm to navigate; nonetheless, this added complexity may be balanced by the reduced computational demands due to bypassing the use of feature selection algorithms.

2.4. Benchmark Framework

Our methodology can accommodate multi-objective problems and incorporate forecast-derived exogenous information in a data-driven manner, with an offline policy design problem formulation. Therefore, we sought a benchmark approach that allows us to compare the results under the same assumptions. Among the existing frameworks in the literature that combine the task of exogenous information extraction with policy design, the Information Selection and Assessment (ISA) framework (Giuliani et al., 2015) appeared to be the best fit. With suitable algorithms, ISA can support a parametric policy design approach and exploit exogenous information (both forecast and observations) in a data-driven manner (Denaro et al., 2017). On the other hand, ISA is not inherently multi-objective, as to explore the relationship between the selected inputs and the objectives trade-offs, users would require multiple runs of the framework (see Figure 1a). Another viable benchmark option would be the Selection of Information for NeuroEvolutionary Policy Search (SINEPS; Zaniolo et al., 2021b), which could learn the best policy input set for the objective trade-offs dynamically. Still, its additional algorithmic complexity does not seem worth pursuing at this research stage.

Thus, the Information Selection and Assessment (ISA) framework (Giuliani et al., 2015) is used as the benchmark for the proposed RL approach. The first step in the ISA framework (see conceptual flowchart in Figure 1a, and detailed flowchart in Figure S2 of Supporting Information S1) is to estimate the expected value of perfect information, that is, the potential performance improvement that could be achieved under the ideal conditions of perfect knowledge of the future disturbance of the system (removing uncertainty on the future, when decisions are taken). We estimate this quantity by contrasting the performance of the Perfect Operating Policy (POP), obtained for example, via Dynamic Programming (Bellman, 1957), against that of a poorly informed baseline, that is, Basic Operating Policy (BOP), relying on a basic set of information (e.g., day of the year and lake storage). The gap between the perfect and basic operating policies reveals the performance improvement that could be achieved under the unrealistic assumption of perfect foresight of future system conditions, indicating the potential benefit of collecting accurate forecast information (Giuliani et al., 2015). In the case of a single-objective problem, this gap is the difference between the (scalar) performance of the policies, while in a multi-objective problem multiple metrics need to be considered to assess the difference between the Pareto fronts (for a review on the topic, see Maier et al., 2014). These metrics should account for the convergence of the final solution to the actual Pareto front that would result from an optimal method, the coverage of the non-dominated space (diversity), and the extent of the non-dominated front, that is, the degree over which the final solution evenly covers the solution space.

To reduce the gap between the POP and BOP, we design the Improved Operating Policies (IOPs) by extending the policy input space with relevant sample information I_t that is selected and used in a model-free fashion, as a surrogate of the future external drivers. As the set of possibly informative signals Ξ is generally large (including, for example, rainfall, inflows, water levels, etc.), automatic selection techniques offer a practical approach to identifying the most relevant subset of information (I_t) that better explains the optimal sequence of release decisions from the POP. For example, we can employ a tree-based algorithm (Giuliani et al., 2015) to rank the signals based on their contribution to explaining one trajectory of release decisions computed by the perfect policy for a selected trade-off between the different objectives $\left(\bar{u}^{POP}\right)$.

The ISA framework provides an approximate solution to close the POP/BOP gap, as the relative contribution of each individual component of the sample information vector to explain the selected POP release decisions may not align with the relative contribution to closing the gaps with the POP performance. Thus, the two steps of Information Selection and (informed) Policy Design (Figure 1a) are repeated until satisfactory results are achieved (Giuliani et al., 2015). Supporting Information S1 provides a more detailed explanation of the ISA framework (see Text S2) and the policy parametrization (see Figure S3).

3. Case Study, Data, and Experimental Setting

3.1. System Description

Lake Como is a large, regulated lake with an active storage capacity of more than 200 Mm^3. It is located in the Adda River Basin of Northern Italy (Figure 2).

The Adda basin upstream of the lake (4,552 km^2) has the typical hydrological regime of sub-alpine regions: inflows are higher during spring and autumn due to snow melt and precipitation, respectively, while lower inflows

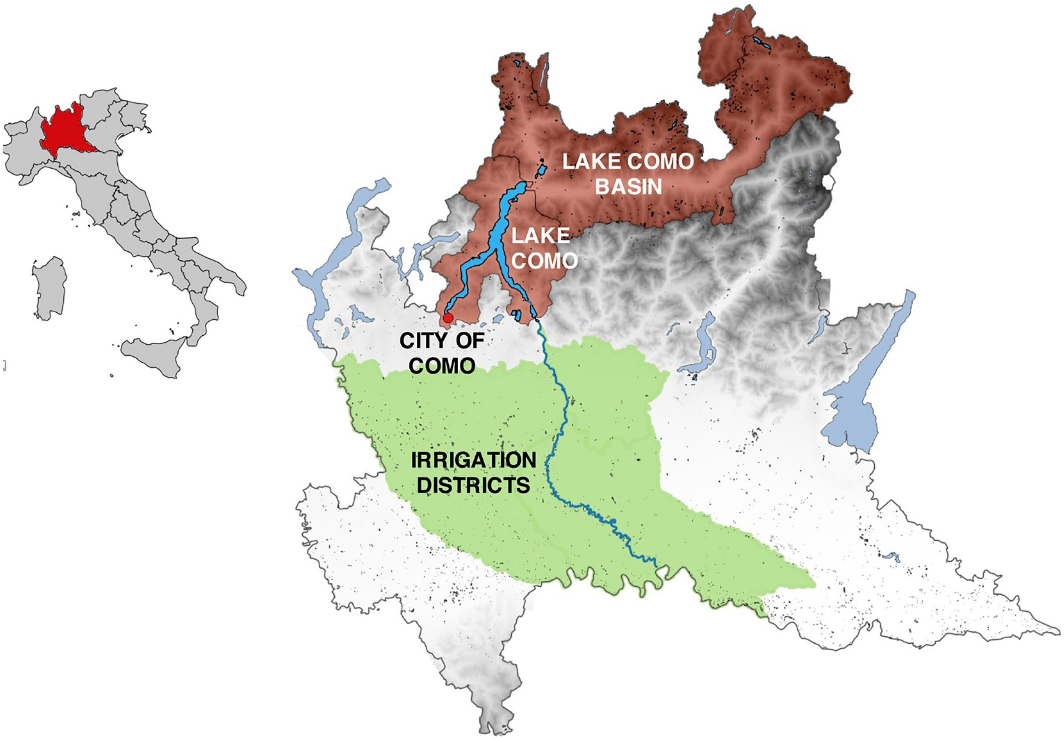

Figure 2. On the left, the Lombardy region is highlighted on a map of Italy. On the right, a physical map of Lombardy, comprising the Lake Como basin, in red, Lake Como, the city of Como, and the irrigation district downstream of the lake (Zaniolo et al., 2021b).

are observed during winter and summer. A number of small- to medium-sized hydropower plants influence the inflow to the lake, releasing more water in autumn and winter, when electricity tariffs are higher while reducing the summer release (Giudici et al., 2021). On the other hand, the downstream water demands would require different timing, as the cumulative summer demand generally exceeds the natural water availability, mainly due to the demands of a large irrigated area (1,400 km^2), with four agricultural districts and eight run-of-the-river hydropower plants (Giuliani, Li, et al., 2016). The time mismatch between the water demand and the inflow to the lake makes it necessary to regulate the outflow and increase the release when most needed downstream while storing more water in other seasons. However, especially during spring, this limits the buffer capacity of the reservoir and increases flood risk for the communities along the lake shores (e.g., the city of Como). Thus, two primary competing objectives have historically driven the regulation of the lake: (a) flood control to avoid floods affecting the city of Como and other populated areas on its shoreline, and (b) water supply to satisfy the demand of downstream agricultural districts and run-of-the-river hydropower plants. Also, recent changes in the hydrological regime have exacerbated the competition between these objectives and required the introduction of a "low lake level" objective. This new objective aims to prevent shallow lake levels during summer that are detrimental to several users, including navigation, tourism, and the environment (Yang et al., 2023).

Therefore, in this study, we formulate the objectives of Lake Como operation as follows:

1. Flood days: the average annual number of days when the lake level (h_t) is above the threshold $h^{flo} = 1.1$ m; it is computed as:

$$J^{flo} = \frac{1}{H/T} \sum_{t=0}^{H-1} g_{t+1}^{flo}; \quad g_{t+1}^{flo} = \begin{cases} 1 & \text{if } h_{t+1} > h^{flo} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where H is the simulation horizon (days), and T is the hydrological year (days).

2. Water supply deficit: the daily mean deficit considering the water released from the lake (r_{t+1}) and the water demand of the downstream users (w_t); it is computed as:

$$J^{def} = \frac{1}{H}\sum_{t=0}^{H-1} g_{t+1}^{def}; \quad g_{t+1}^{def} = \left(\max\left(w_t - \left(r_{t+1} - q^{MEF}\right), 0\right)\right)^{\beta_t} \tag{6}$$

where $q^{MEF} = 22$ m^3/s is the Minimum Environmental Flow (MEF) constraint ensuring adequate environmental conditions in the Adda River, and β_t is a time-varying exponent that penalizes with different importance the deficit during summer and winter. This parameter has been tuned to mimic the decision-making preferences of the operator, with the deficit squared during the summer ($\beta_t = 2$ from 1 April to 10 October), while the actual value ($\beta_t = 1$) is taken during winter.

3. Low lake levels days: the average annual number of days when the lake level (h_t) is below the threshold $h^{low} = -0.2$ m; it is computed as:

$$J^{low} = \frac{1}{H/T}\sum_{t=0}^{H-1} g_{t+1}^{low}; \quad g_{t+1}^{low} = \begin{cases} 1 & \text{if } h_{t+1} < h^{low} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The Lake Como basin has been the subject of various previous studies, which have delved into its intricate social and environmental dynamics and the significance of hydro-meteorological information for its operation. The reader can refer to earlier works to gain a more comprehensive understanding of the socio-economic context of Lake Como and understand the value of hydro-meteorological information for its regulation, (e.g., Anghileri et al., 2013; Denaro et al., 2017; Giuliani & Castelletti, 2016; Giuliani et al., 2019, 2020; Guariso et al., 1986; Zaniolo et al., 2021b). However, it is worth mentioning that our study assumes stricter parameters than the previous ones regarding the Lake Como system. New regulations have introduced a higher Minimum Environmental Flow (MEF) of 22 m^3/s, almost five times higher than the previous 5 m^3/s. Additionally, the long-lasting subsidence in Como (Nappo et al., 2020) has lowered the level of Piazza Cavour, the lake's lowest point on the shoreline. This phenomenon has caused the flood threshold to fall to 1.1 m, down from its former value of 1.24 m. As a result, Lake Como's operational storage capacity has dropped from 254 to 220 Mm^3. This new condition and the additional low-level objective have recently intensified the conflict between competing sectors (upstream stakeholders, lake shores population, and downstream users).

3.2. Observations and Forecasts

Daily time series of observed lake levels and releases are provided by the lake operator from 1946 (the start of regulation, after dam construction) to 2022. From these observations, the net inflow to the lake is estimated by inverting the mass balance equation of the lake storage. This is used both as the observed disturbance of the system for the simulation of the lake dynamics and as a perfect forecast generated from an ideal forecasting system.

We considered multi-timescale real hydrological forecasts, including short-term forecasts that the lake operator currently consults and sub-seasonal and seasonal operational products (freely-available) that may be considered for possible future formal decision-support systems:

- *Short-term deterministic* forecasts (*PRO*), provided by the local company PROGEA (2022) and obtained by feeding a locally calibrated hydrological model (TOPKAPI; Ciarapica & Todini, 2002) with short-term weather forecasts from COSMO (COnsortium for Small-scale MOdelling, 2023). They are single trajectories with an hourly time step and update frequency, a lead time of up to 60 hr, and initially available between 2014 and 2022. The hydrological model was calibrated with the same time series of observed lake levels available for this study, complemented by the river discharges measured at stations upstream of the lake (Fuentes, Samolaco).
- *Sub-seasonal probabilistic* re-forecasts (*EFRF*) produced by the European Flood Awareness System (EFAS, part of the Copernicus Emergency Management Service), over the whole European domain, by forcing the LISFLOOD (Knijff et al., 2010) hydrological model (uncalibrated for the Lake Como basin) with extended-range ensemble forecasts (ECMWF, 2022a). These ensemble forecasts have 11 members with a 6-hr time step,

a twice-weekly update frequency, a 46-day lead time, and are available over the period 1999–2018 (Barnard et al., 2020).

- *Seasonal probabilistic* re-forecasts (*EFSR*) produced by EFAS. Similarly to the sub-seasonal product, these ensemble forecasts are obtained by LISFLOOD but forced here with seasonal meteorological forecasts from the SEAS5 model (ECMWF, 2022b). They have 25 ensemble members, a daily time step, are issued on the first day of each month, have up to 6 months lead time, and are available over the period 1999–2019 (Wetterhall et al., 2020).

As observations represent the daily discharge, we pre-process all products to a daily resolution. The selected study period (and optimization horizon) is from 1 January 1999 to 31 December 2018 ($H = 7{,}305$ days, $T = 365.25$ days), for which observations, EFRF, and EFSR are available. As PROGEA's forecasts are available only between 2014 and 2022, we use a synthetic generation technique to extend them to cover the period 1999–2018 (see Nayak et al., 2018, for details on the method).

3.3. Experiment Settings of the Joint Learning of Forecast Information and Operating Policy Framework

The Extended Operating Policies (EOPs) comprise an operating policy p_θ which is parameterized as a non-linear approximating network through a combination of Gaussian Radial Basis Functions (RBF; Busoniu et al., 2011). RBF-based policies are chosen because in this context they have been shown to consistently outperform those based on Artificial Neural Networks (ANN) (Giuliani, Castelletti, et al., 2016) and represent the most common nonlinear parameterization in water resources operations (Giuliani et al., 2021). The number of basis functions (N) is set as the number of inputs (M) plus two ($N = M + 2$), where M depends on the experiment (i.e., on the forecast information extraction settings). The network presents a single output ($K = 1$) representing the optimal release of the dam. The policies incorporate as inputs the time information through two variables (sine and cosine of the day of the year). Additionally, the policy inputs include lake level (h_t) and a vector of exogenous information (I_t; with size M_I) that results from applying the information extraction function $\mathcal{F}_\zeta(\cdot)$ to the inflow forecasts $\left(\hat{Q}_t\right)$. The information extraction function $\mathcal{F}_\zeta(\cdot)$ specifies the forecast product selection (γ) and its effective lead time (λ). Moreover, to extract information from ensemble forecasts, two candidate uncertainty processing operators (Ψ_{n_e}) are considered in this study, that is, the mean ($\bar{q}^\gamma$) and the quantile of the ensemble $\left(\text{quantile}_\phi(\mathbf{q}^\gamma)\right)$, with quantiles ranging from 0 to 1 with steps of 0.25. For the forecast temporal aggregation over the selected lead time (Ψ_λ), we consider only the mean, that is, the processed forecast represents the cumulative/average discharge over the selected effective lead time (also called *aggregation time* for this reason). Finally, we employ the forecasts by seamless integration, that is, by always using the most recent forecast of each product and without applying any reconciliation technique or bias correction procedure.

All the parameters of an EOP ($[\theta, \zeta]$) are subject to optimization; these include the parameters of the operating policy (θ), that is, the weights for the combination of the basis functions, the bias on the output, the centroids, and the shape parameters of the basis, and the parameters of the information extraction function (ζ; the exact number of parameters depending on the experiment and number of signals M_I to extract from the forecasts).

The search for the optimal parameters of the Extended Operating Policies is performed using the self-adaptive Borg Multi-Objective Evolutionary Algorithm (Borg-MOEA) (Hadka & Reed, 2013), which outperforms other state-of-the-art MOEAs in solving multi-objective optimal control problems (Salazar et al., 2016). Each optimization is run for 2 million function evaluations over the period 1999–2018, with 20 random initializations (initial populations). Building upon prior knowledge from this case study that used similar parametrization (Denaro et al., 2017), we adopted settings that consistently achieve convergence after approximately 500,000 function evaluations.

3.4. Experiment Settings of the Benchmark Framework

The Perfect Operating Policy is computed with Deterministic Dynamic Programming (Bellman, 1957), and the three objectives are converted into a single one by the weighting method (Gass & Saaty, 1955).

The Basic and Improved Operating Policies are designed via Evolutionary Multi-Objective Direct Policy Search (EMODPS; Giuliani, Castelletti, et al., 2016). For a fair comparison against the proposed approach, we used an approximating network of the same size and type, and the same settings for the evolutionary algorithm, as described in Section 3.3.

10.1029/2023WR036724

The candidate variables set (Ξ) is built from the forecast products in a pre-processing phase by applying the same pre-processing function $\mathcal{F}_{\zeta}(\cdot)$ of the joint learning framework. However, this would result in a set with many candidate variables with redundant information that would confuse any automatic selection algorithm. Therefore, by selecting only some values of the temporal aggregation parameter ($\lambda \in [1, 3, 5, 7, 14, 21, \ldots]d$), only a subset of the whole space of combination Z is explored.

As there are too many candidate variables to feasibly explore them exhaustively, we use the Iterative Input Selection (IIS) algorithm (Galelli & Castelletti, 2013b) with Extremely Randomized Trees (Galelli & Castelletti, 2013a; Geurts et al., 2006) to select the set I of variables which, together with the time information (day of the year) and the lake storage, better explain one sequence of release decisions computed by the perfect policy. This iterative algorithm ranks candidate variables based on their ability to reduce the error (measured in terms of coefficient of determination) in predicting the optimal release trajectory using the aforementioned tree-based model. This helps us prioritize variables with the strongest explanatory power for future decision-making. A more detailed description of the IIS algorithm is present in Supporting Information S1 (see Text S3).

As in the original work on the ISA framework (Giuliani et al., 2015), we compare the policies using three metrics: (a) the hypervolume indicator (HV), that is, the dominated space captured by the solution set; (b) the minimum distance from a target solution (Dmin), that is, the distance in the normalized objective space between the target point and the closest solution of the set; and (c) the average distance (Davg), that is, the average distance in the normalized objective space between the target point and the points in the solution set.

4. Results and Discussion

4.1. Forecast Skill Assessment

A preliminary analysis is performed to quantify the accuracy of the available forecasts, comparing their relative performance against a simple benchmark (climatology). As more than one forecast product is available at lead times shorter than 1 month, comparing their performance may be beneficial to understand the results and steer the policy design in the right direction. We remind that the KGE score (Equation 2) is based on the ensemble mean for the probabilistic products (EFRF and EFSR), while for the deterministic forecast (PROGEA), the CRPS is equal to the MAE.

Figure 3 shows the forecast accuracy scores (KGE and its components, as well as CRPS) as a function of the Aggregation Time (i.e., the period over which forecasts are averaged). Although forecast accuracy is often analyzed with respect to the lead time at the original resolution (e.g., 6-hourly and daily values for EFRF and EFSR, respectively), the relationship with AT is more interesting here for two reasons: (a) the lake regulation is expected to benefit from information on the cumulative inflow (e.g., the average discharge over the next month), and so the period of aggregation of the inflow over a future horizon is the key informative parameter here, encompassing more than one forecast lead time; (b) the strategy used to inform the controller is by seamless integration, that is, by always using the most recent forecasts; hence further lead times may be merged or never used separately.

Climatology, defined as the cyclostationary mean of the observed inflows, is used to generate the benchmark scores against which the forecasts are evaluated. According to all scores (Figure 3), PROGEA forecasts outperform the EFAS products when cumulated over their short lead time (i.e., with AT of up to 3 days). Similarly, the sub-seasonal EFAS forecasts (EFRF) outperform their seasonal counterpart (EFSR) up to their maximum AT of 42 days. This result is in agreement with previous studies (Wetterhall & Di Giuseppe, 2018) and is motivated by the more frequent update of the initial condition of the sub-seasonal forecasting model, providing more accurate initial conditions.

However, the climatology benchmark outperforms the EFAS forecasts, either when aggregating over more than 1 month according to the KGE, or across all ATs according to the CRPS. This is mainly determined by the large bias of EFAS with respect to the local observations from which the climatology is derived (which has a higher impact on the CRPS than the KGE). This problem is an expected consequence of the lack of calibration of the LISFLOOD hydrological model (used in EFAS) in the region. However, the correlation component of the KGE suggests that EFAS forecasts are always more correlated with the observations than the climatology, and their performance peaks when cumulated over a horizon of around 14 days. Despite the relatively good correlation of

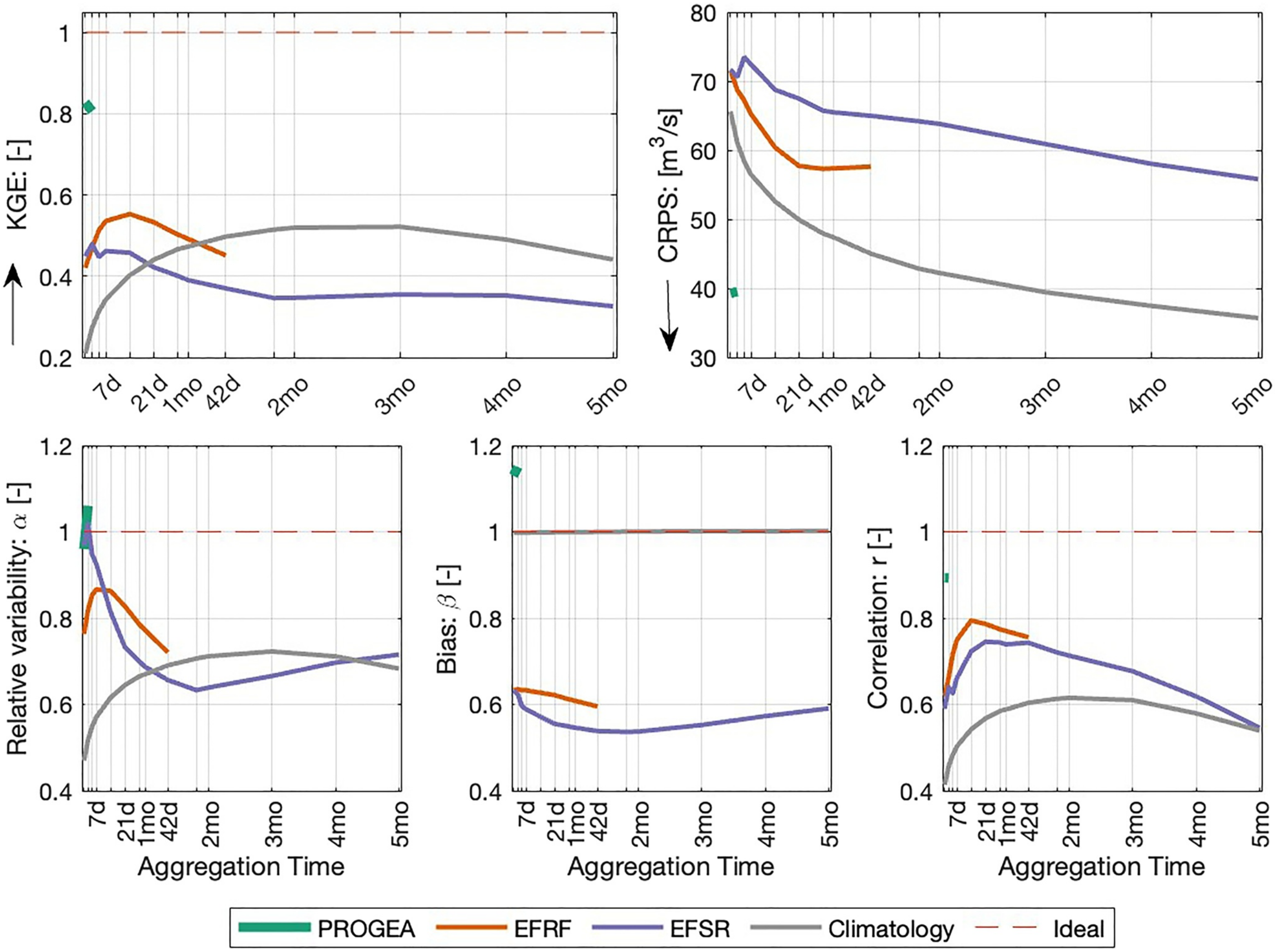

Figure 3. Kling-Gupta Efficiency (KGE), Continuous Rank Probability Score (CRPS), and three KGE components (Relative variability, Bias Ratio, and Correlation) as a function of Aggregation Time (horizon over which inflow is cumulated) for all the available forecast products and a climatology benchmark. The dashed red line indicates the ideal value for each score.

EFAS forecasts with observations, their strong dry bias ($\beta < 1$) and underestimated variability ($\alpha < 1$) penalize them, leading to lower overall performance than climatology according to the general metrics, whether considering the uncertainty of the ensemble (CRPS) or not (KGE). Further results on the unbiased Brier Skill Score (UBSS; see Figure S1 in Supporting Information S1) show that,in general, EFAS forecasts outperform the climatology benchmark at lead times longer than 7 days across different low-flow and high-flow thresholds, with a clear performance peak for ATs around 21 days for high flows. These results are in line with those of correlation (Figure 3) and suggest that EFAS forecasts provide more information than the climatology when focusing on inflow trends and anomalies while discarding the effect of the strong biases reflected by the KGE and CRPS.

4.2. Performance of Extended Operating Policies

Building on the forecast skill assessment reported in the previous Section, we perform a first experiment to verify the learning of the best Aggregation Time of a single fictitious seamless deterministic product ($\gamma = BS$; $\zeta = [\lambda]$; $n_\theta = 55$). This is called the best-skill product and combines the 3 available forecast products by selecting the forecast with the best KGE score for each AT. This means using PROGEA for the first 3 days, EFAS EFRF ensemble mean between 4 and 42 days, and EFAS EFSR ensemble mean from 43 days onwards. The performance of the EOPs is benchmarked against a set of Basic Operating Policies (BOPs) not informed by any forecast (formally $I_t = \varnothing$; $\zeta = [\ ]$; $n_\theta = 36$) and an upper-bound solution obtained by solving a deterministic problem, which is further discussed in the next Section.

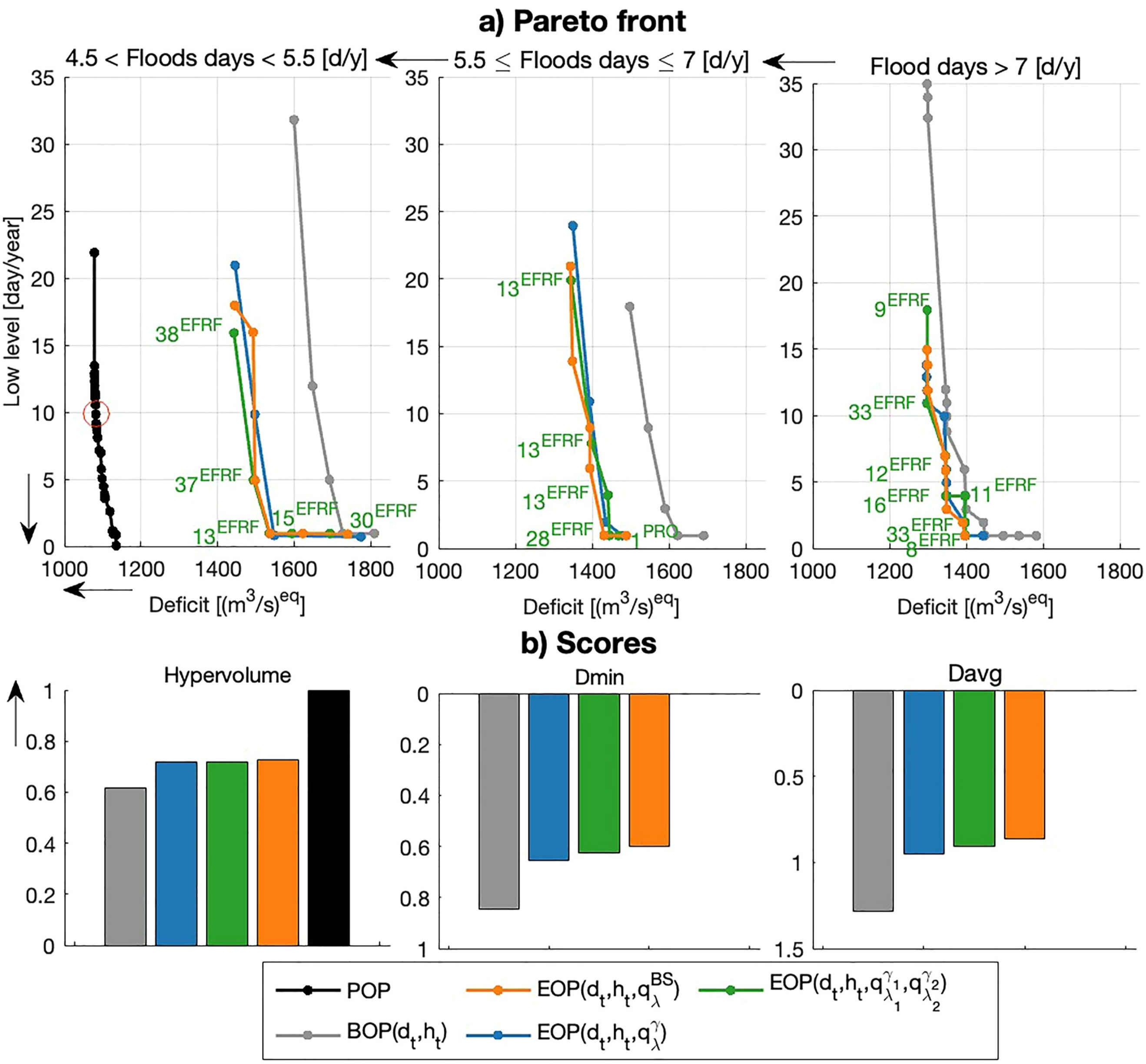

Figure 4. Performances obtained by three sets of Extended Operating Policies (EOPs), the Perfect Operating Policies (POPs), and the Basic Operating Policies (BOPs). The EOPs are informed by day of the year (d_t) , lake level (h_t) , and one between best-skill forecast product (q_λ^{BS}, with λ representing the selected Aggregation Time), all products processed into one input ($\bar{q}_\lambda^\gamma$, with γ representing the selected product), and all products processed into two inputs ($\bar{q}_{\lambda_1}^{\gamma_1}$, $\bar{q}_{\lambda_2}^{\gamma_2}$). The symbol $\bar{q}$ indicates that the average operator processes the ensemble of the probabilistic forecasts. Panels: (a) policies in three projections of the objective space for different levels of flood days; the arrows indicate the direction of preference, with the dominant solutions in the left-bottom corner of the leftmost plot; all solutions have designed policies informed by the PROGEA forecast with an Aggregation Time of 3 days (γ = PRO, $\lambda = 3$); when the EOP uses a second exogenous input, we report the chosen parameters of the forecast information extraction function (λ_2 and γ_2) in a label close to the solution (e.g., 38^{EFRF}, where $\lambda_2 = 38$ and γ_2 = EFRF); (b) forecast value quantified by the metrics described in Section 3.4 using the target POP solution marked by the red circle (hypervolume–HV, minimum and average distance from a target solution–Dmin, Davg).

The results also show that the EOPs successfully improve the performance of the BOPs, especially for solutions with less than 5.5 flood days per year (Figure 4). Interestingly, the policy design consistently selects the 3-day cumulative inflow, corresponding to using the PROGEA forecasts. This choice can be explained by the substantially higher accuracy of the PROGEA short-term forecasts with respect to the sub-seasonal and seasonal EFAS products (Figure 3). Using more skillful, shorter-term forecasts thus results in policies that outperform those using lower-skill, longer-term products. The results also indicate that it is possible to further improve the Water Deficit and Low-Level objectives by accepting more flood events (Figure 4). However, the forecast value decreases when moving to solutions with higher numbers of flood days because, in this case, the knowledge of future inflows is less critical for the lake operation, which can store water in favor of the other objectives without being limited by the increasing flood risk.

Given these promising results in learning the best AT, we run a second experiment in which the policy design simultaneously learns the best AT and the best forecast product ($\zeta = [\lambda, \gamma]$; $n_\theta = 55$). The performance of the resulting EOPs (Figure 4) is very similar to the solutions informed by the best-skill product, as the EOP design selects the PROGEA forecast again at 3 days AT to inform the lake operation. We can notice, however, that these solutions attain slightly higher values of the other metrics (Dmin and Davg), probably because of the larger parameter space to be explored than the first experiment.

As a third experiment, we now solve the EOP design problem while selecting two different forecast products with their respective ATs ($\zeta = [\lambda_1, \gamma_1, \lambda_2, \gamma_2]$; $n_\theta = 78$). In this case, the performance of the EOPs improves slightly from that of the EOPs informed by a single product but does not outperform the solutions informed by the best-skill product (Figure 4). However, it is worth analyzing the selected combinations of forecast products and ATs: all solutions rely on the PROGEA forecasts aggregated over 3 days, along with medium- to extended-range information (i.e., AT between 9 and 38 days for the EFRF product). Notably, the EOPs attaining the best performance in terms of flood control that are located in the compromise region of the Pareto front between Deficit and Low Level select an AT of 13 and 15 days, while the extreme solutions of the same Pareto front are associated to an AT of 38 and 30 days for the best Deficit and Best Low-Level solutions, respectively. This result is in disagreement with the findings of Denaro et al. (2017) and Zaniolo et al. (2021b), who, using perfect forecasts, found much longer ATs. In addition to the influence of larger errors in real forecast products at longer lead times, this difference is due to the reduction of the lake's active capacity due to the subsidence of the city of Como (see Section 3.1) and is in line with other studies that linked reservoir capacity with effective forecast horizon (Q. Zhao et al., 2019; Turner et al., 2020). Yet, the added value of this second medium- to extended-range information is marginal, and the forecast value is dominated by the information provided by the PROGEA product.

To confirm the impact of the reservoir capacity reduction, we run the EOP design with perfect forecasts (fourth experiment: $\gamma = PF$; $\zeta = [\lambda]$; $n_\theta = 55$). In this case, the preferred AT is between 20 and 25 days and the resulting EOP performance is substantially improved compared to all other EOPs due to the more accurate (perfect) information used to condition the lake operation (see Figure S4 in Supporting Information S1). The move toward longer selected ATs with perfect forecasts with respect to those chosen using real ones suggests that with real forecasts, the RL algorithm finds a trade-off between more skillful shorter-range forecasts and, ideally, more informative longer-range ones.

In the fifth (and sixth) experiment, we use the quantile operator (controlled by the parameter ϕ; $0 \leq \phi \leq 1$) instead of the mean operator to process the ensemble forecasts, leading to the same results as using the mean (see Figure S5 in Supporting Information S1). The similarities are not only limited to the fact that almost all solutions use the median ($\phi = 0.5$) as the preferred quantile from the ensemble: the algorithm learns the same forecast product and AT (PROGEA with AT of 3 days is the preferred option, followed by EFRF with AT between 30 and 40 days). In line with our skill analysis (e.g., see CRPS and UBSS), the additional forecast uncertainty information provided by the ensemble forecasts (EFAS's EFRF and EFSR) is not informative enough to be as easily used as the more skillful short-range deterministic forecast (PROGEA).

4.3. Benchmark Solutions

The results in Figure 4 are evaluated against a set of Basic and Perfect Operating Policies, representing a lower and upper bound of the system performance. The BOP solutions represent traditional operating policies that are not informed by any forecast. At the same time, the POPs are obtained by ideally removing the uncertainty of the inflow process and solving a deterministic problem (see Section 3.4). Given the BOP and POP sets, we also run the ISA framework (Giuliani et al., 2015) to validate the information selection results of the EOP design. Specifically, the Iterative Input Selection algorithm is used to extract the subset of variables that best explain the release decision of a target POP.

After the algorithm removes the influence of the day of the year and the lake storage (representative of the BOP inputs), it selects the extended-range forecast information, that is, the EFRF product with AT equal to 14 days (Figure 5b). The same trade-off between skill and value of the last EOP design appears here (see the end of

10.1029/2023WR036724

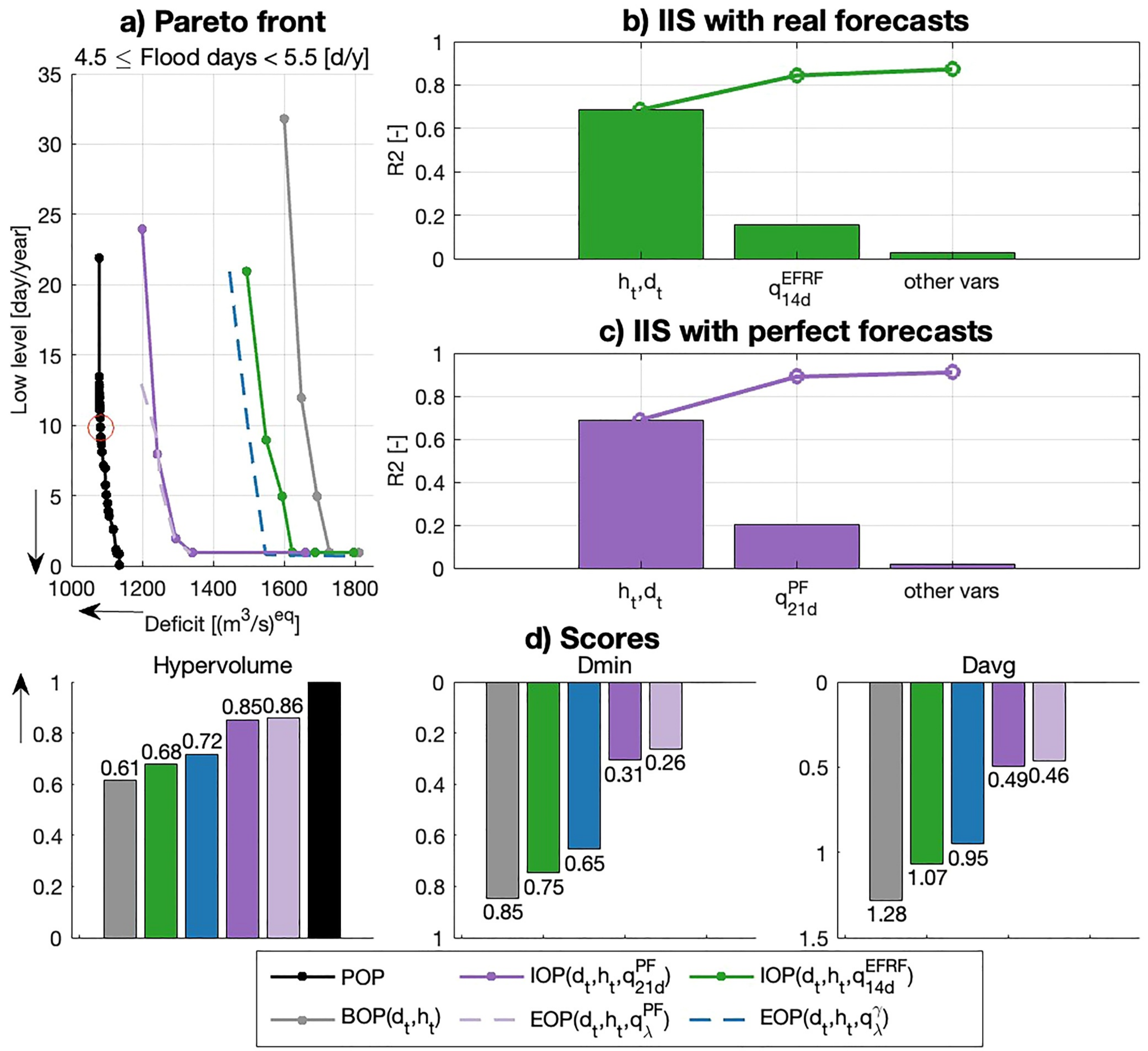

Figure 5. Performances obtained by two benchmark sets of Improved Operating Policies (IOPs) designed with the Input Selection and Assessment framework (ISA; Giuliani et al., 2015), two sets of Extended Operating Policies (EOPs) from our proposed methodology, the Perfect Operating Policies (POPs) and the Basic Operating Policies (BOPs). Panels: (a) policies in the objectives space for the lowest level of flood days; filled lines indicate the benchmark policies (POPs, BOPs, and IOPs); dashed lines represent the output of the proposed methodology (EOPs); the symbols in the legend represent the policies inputs: day of the year (d_t), lake level (h_t), forecast (q); the superscript on the forecast indicates the product: perfect forecast (PF), EFAS sub-seasonal forecast (EFRF), and product freely chosen by our joint learning algorithm (γ); the subscript on the forecast indicates the Aggregation Time selected in the ISA framework or freely chosen by our algorithm (λ); (b) Iterative Input Selection (IIS; Galelli & Castelletti, 2013b) results as the average cumulative performance, measured by the coefficient of determination (R2), of the model describing the optimal release trajectory of the selected POP solution (red circle) with the selected inputs; (c) same IIS results with perfect forecasts; (d) forecast value quantified by the metrics described in Section 3.4 using the target solution marked by the red circle (hypervolume–HV, minimum and average distance from a target solution–Dmin, Davg).

Section 4.2). The most informative aggregation time is above 21 days (see results with the perfect forecast in Figure 5c and Figure S4 in Supporting Information S1). However, when using real forecasts, the algorithm needs to find a compromise between the effective forecast horizon and skill: it reduces the considered horizon in exchange for increased accuracy.

10.1029/2023WR036724

Interestingly, the IIS prefers the EFRF to the PROGEA product, although we showed already that the latter is more valuable for reservoir operations (Figure 4). Indeed, the IOPs designed with the results of this selection procedure are dominated by the EOPs (Figures 5a and 5d). Moreover, the value of conditioning the IOPs with additional information on the forecast uncertainty (variance of the ensemble together with the mean of the ensemble as input) is marginal (see Figure S6 in Supporting Information S1).

From this analysis emerge three key findings: (a) the added value of the proposed approach with respect to other two-step approaches, such as the ISA framework: by optimizing the forecast processing during the policy design, we can directly extract the most valuable information for reservoir operations; (b) forecast skill is the main factor behind the choice of both the product and the effective forecast horizon, and this is evident especially in our RL approach; (c) information on forecast uncertainty can be considered as candidate policy inputs in both our RL approach and the benchmark (ISA); however, for this case study (setting/data set), simple forecast uncertainty information like the variance or quantiles from the ensemble do not add value to the policy performance. Moreover, our approach has other general and technical advantages: (a) it can employ real deterministic and ensemble forecasts out-of-the-box, as it extends EMODPS by also parametrizing the forecast pre-processing and information extraction; (b) it does not suffer from biases in the forecast, in contrast to model-based real-time control approaches (like MPC); (c) compared to the ISA benchmark, it is faster and reduces computational costs, as it does not require separate pre-processing and automatic information selection (see Figure 2), encompassing all steps in a single optimization run (which has negligible runtime overhead with respect to the canonical EMODPS); (d) differently from two-step methodologies (e.g., Giuliani et al., 2015; Zaniolo et al., 2021b), there is no need to compute and select a target solution (e.g., via DDP), exploring all the possible solutions in the multi-objective space directly.

4.4. Discussion on the Behavior of Alternative Policies

Looking at simulated trajectories of lake level and release under different policies can help us better understand the various contributions of selected forecast information. Figure 6 investigates the trajectories of the following policies: the target POP, the closest BOP, two IOPs that use real (EFRF) and perfect extended-range forecasts, and an EOP informed by a single forecast product. To avoid selecting extreme policies, the chosen solutions minimize flooding to a minimum of 4.45 days per year while minimizing the deficit and maintaining a reasonable distance from the target POP low levels (±3 days per year).

Thanks to the additional information about future inflows, all the IOP/EOPs maintain a higher average lake level than the BOP, saving more water before the summer irrigation period. This behavior is crucial to satisfy the two competing objectives of water supply and low lake level control. Then, all policies show similar release trajectories during summer, with only relatively small deviations observed mainly at the end of the irrigation season (September). A late summer release reduction appears only for the BOP and the IOP with EFRF forecasts to find a compromise with the low-level objective. This compromise leads to a sub-optimal water supply, showing the impact of the lack of accurate information by using real forecasts with respect to perfect ones.

Interestingly, all the forecast-informed policies lead to a lake level even higher than the target POP at the beginning of the irrigation period (April). This condition would be helpful to ensure that the water supply is satisfied during the summer period. However, the IOP, informed by real EFRF forecasts, releases an excessive amount of water to prevent flooding, negatively affecting both the deficit and low levels objectives. This negative effect is reduced for the EOP policy, as the simulated one uses the very skillful PROGEA forecasts. Thanks to the high reliability of the forecast, this policy can approach the flooding limit more safely, which, in the end, turns into a competitive advantage to decrease the water deficit. This behavior explains why PROGEA is the preferred forecast across all experiments despite not providing information on the forecast uncertainty.

4.5. Discussion on the Effects of Forecast Information on the Policy Performance

Our analysis shows that the superior skill of the PROGEA forecast product positively impacts policy performance across all experiments, dominating the forecast selection despite the shorter lead times available (3 days) and the lack of explicit information on forecast uncertainty (being PROGEA deterministic). Therefore, to better isolate the forecast value of the EFAS EFRF product, we design an EOP by allowing the selection of the best AT (λ) but forcing the use of this forecast product ($\gamma = EFRF; \zeta = [\lambda]; n_\theta = 55$). Results show that the performance of the resulting EOPs is lower than the one using the PROGEA forecasts (see Figure S4 in Supporting Information S1),

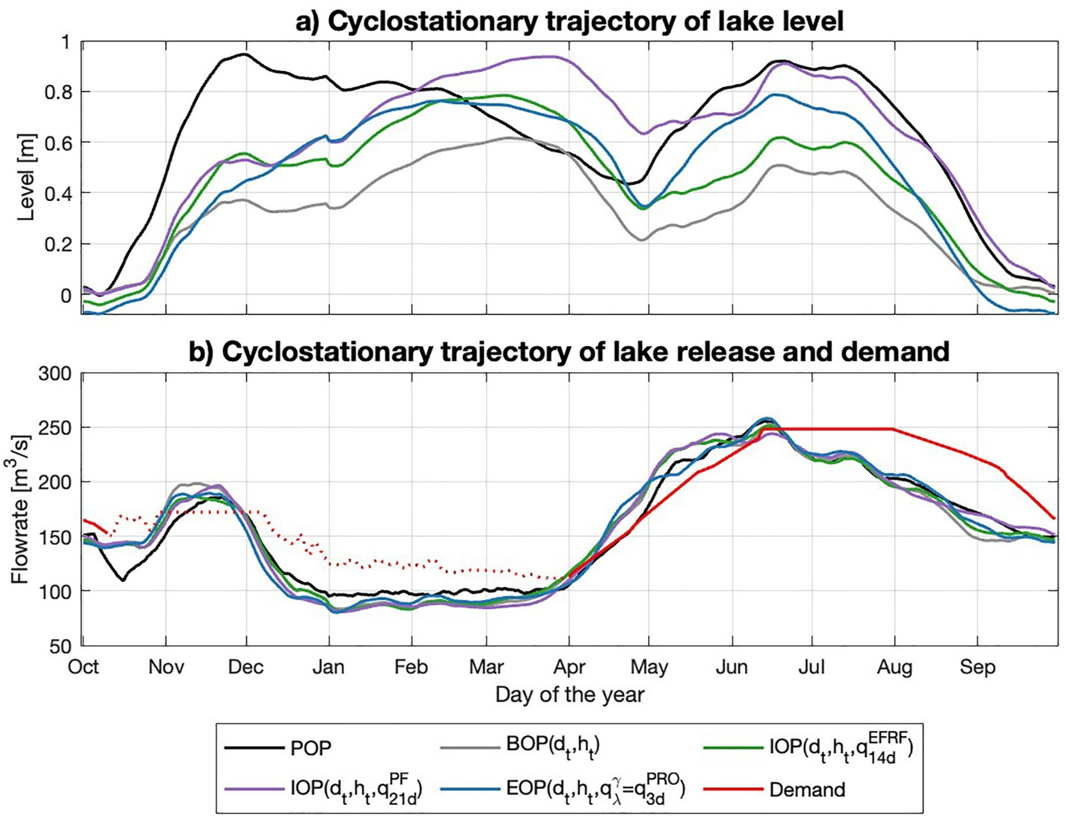

Figure 6. Analysis of the Lake Como level (a) and release (b) cyclostationary trajectories (with 5-d moving average) under different policies: the target Perfect Operating Policy, the closest Basic Operating Policy, two Improved Operating Policies informed by either perfect or real extended-range forecasts, an Extended Operating Policy informed by real forecasts. The red line indicates the amount of water to be released to reduce the deficit, that is, the sum of the water demand and MEF $\left(w_t + q^{MEF}\right)$, with a continuous line when the exponent for the deficit objective β_t is 2, dotted when 1.

confirming that our approach effectively learned the most valuable forecast information across multiple products (see Figure 4). Still, it is interesting to discuss the selected ATs when all forecasts are available or when only the EFRF ones are; in both cases, the EOPs (the resulting PROGEA-informed EOP and EFRF-informed EOP) select ATs that strike a balance between the most skillful lead-time for each product (1 day for PROGEA and 14 days for EFRF; see Figure 3) and the most valuable extended range lead-time selected in the ideal case of perfect forecasts (around 23 days; see Figure S4 in Supporting Information S1). The selection of an AT that compromises on the ideal extended range horizon in favor of a shorter but more skillful one suggests that with real forecasts, the proposed RL-based algorithm learns the trade-off between forecast skill and the information in the effective forecast horizon, selecting the one that leads to the best value of the Extended Operating Policy in the objective space.

When multiple forecasts are available over the same lead time, it is reasonable to expect that the optimization algorithm simply learns that a more skillful product should lead to a higher value. However, when a single forecast product is available across different lead times (with its skill continuously changing across this dimension), the most skilled signal that the algorithm can extract may differ from the one with the highest value for policy design. Indeed, moving from an ideally more informative extended range horizon to a more skillful shorter one happens independently from the method applied, and we observe it both in the EOP design (as mentioned above) and the ISA framework (see Section 4.3 and Figure 5). Nevertheless, the value of the forecast-informed exogenous signal extracted is realized only when paired with a policy optimized for that signal. As shown in the trajectories (Figure 6), different forecast products and skills lead to different operational strategies, for example, with the skillful short-term PROGEA, the levels are kept generally higher during summer and spring than with the extended range EFAS EFRF thanks to the higher reliability of the former. However, this trade-off can only be appreciated when these combinations compete directly in the objective space, which is the main strength of the proposed RL methodology.

5. Conclusions

In this paper, we propose a new general and flexible approach that can jointly learn how to extract the most valuable information from a set of multi-timescale forecasts and how to use this information to improve water systems operations. This is achieved by a Reinforcement Learning (RL) framework that is based on an extension of the Evolutionary Multi-Objective Direct Policy Search (EMODPS). Specifically, we augment the parameter space the evolutionary algorithm has to learn, including (a) the parameters governing the information extraction (e.g., forecast products and aggregation times) and (b) the control policy parameters. This new approach overcomes the limitations of previous algorithms for information selection and sequential policy design, such as the Input Selection and Assessment (ISA) framework. Unlike two-step methodologies, our RL approach designs control policies while simultaneously extracting the most effective forecast information, ensuring that the goals of the two tasks are aligned. In contrast, frameworks like ISA rely on (a) separate algorithms for information extraction and policy design, (b) a pre-computed target (or "perfect") policy—typically derived with Dynamic Programming, which requires observations of the inflows and becomes computationally infeasible for multi-reservoir systems, (c) a predefined set of candidate signals extracted from the forecasts that can be used as policy input, and (d) multiple executions of the framework to explore the trade-offs between objectives. Lake Como is used as a case study, as this water system is regulated to meet multiple objectives acting over different time horizons, primarily flood control and water supply, for which forecasts over both short and seasonal time scales may be valuable. Multiple operational forecast products are available to investigate this, including locally calibrated short-term deterministic forecasts (PROGEA), which are currently used by the lake operator, and continental locally uncalibrated sub-seasonal and seasonal ensemble forecasts (EFAS).

The Extended Operating Policies (EOPs) designed with our proposed approach outperform the Improved Operating Policies (IOPs) designed with the ISA benchmark. When using real forecasts, the EOP solutions in the Pareto front show a 6% improvement in hypervolume, a 13% reduction in the minimum distance from the target solution, and a 12% reduction in the average distance from the target solution. Also, our approach offers greater flexibility for policy design and trade-off analysis by producing a single Pareto front of solutions (policy parameters paired with forecast information) in a single run of the optimization algorithm. Overall, this study shows that integrating the policy design with forecast selection and processing operations is possible and beneficial. However, for this case study, a single exogenous variable is sufficient to extract most of the value from the forecasts. The test with two additional inputs led to such a slight improvement that it does not seem worth adding more than one input from the current forecast products for Lake Como. This should be reassessed with future advances in forecast skills. With real operational products, the controller prefers the more skillful short-range forecasts to, ideally, more informative extended-range ones. Still, the results show that the EMODPS framework can reach a good performance even with the locally uncalibrated extended-range forecasts, despite their bias for this case study, as the policies informed by the sub-seasonal EFAS forecasts produce a Pareto Front relatively close to the ones relying on the PROGEA forecasts, which are explicitly calibrated for the Adda Basin.

Future research efforts should primarily focus on exploring the use of the probabilistic information of ensemble forecasts within the proposed framework. This could be done in several ways, such as testing as candidate policy inputs more complex probabilistic descriptors than those tested (i.e., mean, quantiles, and variance), but also by leveraging the ability of the EMODPS framework to work with Monte Carlo simulations. Moreover, other operations that could be parametrized and included in the forecast information extraction layers of the EOPs include (a) combining multiple products with more sophisticated techniques (e.g., forecast reconciliation); (b) including bias-correction (also to further prove and quantify the low impact of bias in EMODPS); and (c) designing policies informed by variables with time-varying or event-dependent parameters. The EMODPS procedure can also benefit from a relaxation of the structure with which the control law has been parameterized (e.g., Neuro-EMODPS Zaniolo et al., 2021a).

Finally, the EMODPS procedure can easily be scaled up for many-objective optimization problems and multi-reservoir systems without requiring additional algorithms. However, the increase in complexity derived from the augmentation of the space of the parameters may lead to higher computational costs. Our results suggest that wise use of additional information on the forecasts (e.g., their skill) can help reduce the parameter space and speed up the process.

Data Availability Statement

Observations of lake inflows are published daily by Consorzio dell'Adda (2023, available for download under request to the lake regulator, https://www.laghi.net). Short-term forecast data by PROGEA S.r.l. are available under their EFFORTS system's data policy (PROGEA, 2022, http://www.progea.net/prodotti.php?c=Software&p=EFFORTS) with restrictions for commercial regulations and are not publicly accessible to the public. Researchers can gain access under agreements with PROGEA or by purchasing the data. Sub-seasonal (Barnard et al., 2020, https://doi.org/10.24381/cds.c83f560f) and seasonal forecasts (Wetterhall et al., 2020, https://doi.org/10.24381/cds.768eefc2) produced by the European Flood Awareness System are available for download on the Copernicus Climate Change Service Climate Data Store (CDS) under the CEMS-FLOODS data sets license. The Lake Como simulation and EMODPS implementation source code is available from Zenodo (Giuliani & Yang, 2022, GNU General Public License v2.0, https://doi.org/10.5281/zenodo.7401168). The Borg Many-Objective Evolutionary Algorithm is presented in this paper: Hadka and Reed (2013). The Iterative Input Selection algorithm is included in this paper: Galelli and Castelletti (2013b). The Dynamic Programming algorithm and the forecasts analysis scripts are available on Zenodo (Zanutto, 2023, GNU General Public License v3.0, https://doi.org/10.5281/zenodo.10123625).

Acknowledgments

The authors would like to thank PROGEA S.r.l., the European Centre for Medium-Range Weather Forecasts (ECMWF), and the Copernicus Emergency Management Service (CEMS) for producing the forecast products used in this study. We gratefully acknowledge the support of Gianni Pani and Gianluca Colliva from PROGEA for their data provision. We thank Guido Ascenso and two anonymous reviewers for the careful and constructive review of the manuscript. Dennis Zanutto is funded by the PhD scholarship of Politecnico di Milano and KWR. Andrea Ficchì, Matteo Giuliani, and Andrea Castelletti were partly funded by the EU Horizon 2020 project CLINT (Climate Intelligence: Extreme events detection, attribution and adaptation design using machine learning) under Grant Agreement 101003876. Andrea Ficchì also acknowledges support from the AXA Research Fund Fellowship on Coastal Livelihoods.

References

Anghileri, D., Castelletti, A., Pianosi, F., Soncini-Sessa, R., & Weber, E. (2013). Optimizing watershed management by coordinated operation of storing facilities. *Journal of Water Resources Planning and Management*, *139*(5), 492–500. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000313

Anghileri, D., Voisin, N., Castelletti, A., Pianosi, F., Nijssen, B., & Lettenmaier, D. P. (2016). Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments. *Water Resources Research*, *52*(6), 4209–4225. https://doi.org/10.1002/2015WR017864

Arnal, L., Cloke, H. L., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., et al. (2018). Skilful seasonal forecasts of streamflow over Europe? Hydrology and. *Earth System Sciences*, *22*(4), 2057–2072. https://doi.org/10.5194/hess-22-2057-2018

Arsenault, R., & Côté, P. (2019). Analysis of the effects of biases in ensemble streamflow prediction (ESP) forecasts on electricity production in hydropower reservoir management. *Hydrology and Earth System Sciences*, *23*(6), 2735–2750. https://doi.org/10.5194/hess-23-2735-2019

Barnard, C., Krzeminski, B., Mazzetti, C., Decremer, D., de Wiart, C. C., Harrigan, S., et al. (2020). Reforecasts of river discharge and related data by the European Flood Awareness System, version 4.0. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.c83f560f

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47–55. https://doi.org/10.1038/nature14956

Bellman, R. (1957). *Dynamic programming*. Princeton University Press.

Bertoni, F., Giuliani, M., Castelletti, A., & Reed, P. M. (2021). Designing with information feedbacks: Forecast informed reservoir sizing and operation. *Water Resources Research*, *57*(3), e2020WR028112. https://doi.org/10.1029/2020WR028112

Bertsekas, D. P. (2005). Dynamic programming and suboptimal control: A survey from ADP to MPC. *European Journal of Control*, *11*(4), 310–334. https://doi.org/10.3166/ejc.11.310-334

Bertsekas, D. P. (2019). *Reinforcement learning and optimal control* (1st ed., Vol. 1). Athena Scientific.

Busoniu, L., Ernst, D., De Schutter, B., & Babuska, R. (2011). Cross–Entropy optimization of control policies with adaptive basis functions. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, *41*(1), 196–209. https://doi.org/10.1109/TSMCB.2010.2050586

Castelletti, A., Ficchì, A., Cominola, A., Segovia, P., Giuliani, M., Wu, W., et al. (2023). Model predictive control of water resources systems: A review and research agenda. *Annual Reviews in Control*, *55*, 442–465. https://doi.org/10.1016/j.arcontrol.2023.03.013

Castelletti, A., Pianosi, F., & Restelli, M. (2013). A multiobjective reinforcement learning approach to water resources systems operation: Pareto Frontier approximation in a single run. *Water Resources Research*, *49*(6), 3476–3486. https://doi.org/10.1002/wrcr.20295

Castelletti, A., Pianosi, F., & Soncini-Sessa, R. (2008). Water reservoir control under economic, social and environmental constraints. *Automatica*, *44*(6), 1595–1607. https://doi.org/10.1016/j.automatica.2008.03.003

Ciarapica, L., & Todini, E. (2002). TOPKAPI: A model for the representation of the rainfall-runoff process at different scales. *Hydrological Processes*, *16*(2), 207–229. https://doi.org/10.1002/hyp.342

COnsortium for Small-scale MOdelling. (2023). Retrieved from https://www.cosmo-model.org

Consorzio dell'Adda. (2023). Retrieved from https://www.addaconsorzio.it

Denaro, S., Anghileri, D., Giuliani, M., & Castelletti, A. (2017). Informing the operations of water reservoirs over multiple temporal scales by direct use of hydro-meteorological data. *Advances in Water Resources*, *103*, 51–63. https://doi.org/10.1016/j.advwatres.2017.02.012

Doering, K., Quinn, J., Reed, P. M., & Steinschneider, S. (2021). Diagnosing the time-varying value of forecasts in multiobjective reservoir control. *Journal of Water Resources Planning and Management*, *147*(7), 04021031. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001386

ECMWF. (2022a). Medium-range weather forecasts. Retrieved from https://www.ecmwf.int/en/forecasts/documentation-and-support/medium-range-forecasts

ECMWF. (2022b). Seasonal weather forecasts. Retrieved from https://www.ecmwf.int/en/forecasts/documentation-and-support/long-range

Galelli, S., & Castelletti, A. (2013a). Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling. *Hydrology and Earth System Sciences*, *17*(7), 2669–2684. https://doi.org/10.5194/hess-17-2669-2013

Galelli, S., & Castelletti, A. (2013b). Tree-based iterative input variable selection for hydrological modeling. *Water Resources Research*, *49*(7), 4295–4310. https://doi.org/10.1002/wrcr.20339

Gass, S. I., & Saaty, T. L. (1955). Parametric objective function (part 2)—Generalization. *Journal of the Operations Research Society of America*, *3*(4), 395–401. https://doi.org/10.1287/opre.3.4.395

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*(1), 3–42. https://doi.org/10.1007/s10994-006-6226-1

Giudici, F., Anghileri, D., Castelletti, A., & Burlando, P. (2021). Descriptive or normative: How does reservoir operations modeling influence hydrological simulations under climate change? *Journal of Hydrology*, *595*, 125996. https://doi.org/10.1016/j.jhydrol.2021.125996

Giuliani, M., & Castelletti, A. (2016). Is robustness really robust? How different definitions of robustness impact decision-making under climate change. *Climatic Change*, *135*(3), 409–424. https://doi.org/10.1007/s10584-015-1586-9

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., & Reed, P. M. (2016). Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. *Journal of Water Resources Planning and Management*, *142*(2), 04015050. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000570

Giuliani, M., Crochemore, L., Pechlivanidis, I., & Castelletti, A. (2020). From skill to value: Isolating the influence of end user behavior on seasonal forecast assessment. *Hydrology and Earth System Sciences*, *24*(12), 5891–5902. https://doi.org/10.5194/hess-24-5891-2020

Giuliani, M., Lamontagne, J., Reed, P., & Castelletti, A. (2021). A state-of-the-art review of optimal reservoir control for managing conflicting demands in a changing world. *Water Resources Research*, *57*(12), e2021WR029927. https://doi.org/10.1029/2021WR029927

Giuliani, M., Li, Y., Castelletti, A., & Gandolfi, C. (2016). A coupled human-natural systems analysis of irrigated agriculture under changing climate. *Water Resources Research*, *52*(9), 6928–6947. https://doi.org/10.1002/2016WR019363

Giuliani, M., Pianosi, F., & Castelletti, A. (2015). Making the most of data: An information selection and assessment framework to improve water systems operations. *Water Resources Research*, *51*(11), 9073–9093. https://doi.org/10.1002/2015WR017044

Giuliani, M., & Yang, G. (2022). EILab-Polimi/LakeComo: First release of LakeComo_equity code. *Zenodo*. https://doi.org/10.5281/zenodo.7401169

Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., & Block, P. (2019). Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. *Water Resources Research*, *55*(11), 9133–9147. https://doi.org/10.1029/2019WR025035

Guariso, G., Rinaldi, S., & Soncini-Sessa, R. (1986). The management of Lake Como: A multiobjective analysis. *Water Resources Research*, *22*(2), 109–120. https://doi.org/10.1029/WR022i002p00109

Guo, Y., Yu, X., Xu, Y.-P., Chen, H., Gu, H., & Xie, J. (2021). AI-based techniques for multi-step streamflow forecasts: Application for multi-objective reservoir operation optimization and performance assessment. *Hydrology and Earth System Sciences*, *25*(11), 5951–5979. https://doi.org/10.5194/hess-25-5951-2021

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hadka, D., & Reed, P. (2013). Borg: An auto-adaptive many-objective evolutionary computing framework. *Evolutionary Computation*, *21*(2), 231–259. https://doi.org/10.1162/EVCO_a_00075

Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, *15*(5), 559–570. https://doi.org/10.1175/1520-0434(2000)015⟨0559:DOTCRP⟩2.0.CO;2

Knijff, J. M. V. D., Younis, J., & Roo, A. P. J. D. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, *24*(2), 189–212. https://doi.org/10.1080/13658810802549154

Lee, D., Ng, J. Y., Galelli, S., & Block, P. (2022). Unfolding the relationship between seasonal forecast skill and value in hydropower production: A global analysis. *Hydrology and Earth System Sciences*, *26*(9), 2431–2448. https://doi.org/10.5194/hess-26-2431-2022

Maier, H., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L., Cunha, M., et al. (2014). Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions. *Environmental Modelling & Software*, *62*(0), 271–299. https://doi.org/10.1016/j.envsoft.2014.09.013

Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblas-Reyes, F., et al. (2021). Initialized Earth System prediction from sub-seasonal to decadal timescales. *Nature Reviews Earth & Environment*, *2*(5), 340–357. https://doi.org/10.1038/s43017-021-00155-x

Nappo, N., Ferrario, M. F., Livio, F., & Michetti, A. M. (2020). Regression analysis of subsidence in the Como Basin (Northern Italy): New insights on natural and anthropic drivers from InSAR data. *Remote Sensing*, *12*(18). 2931. https://doi.org/10.3390/rs12182931

Nayak, M. A., Herman, J. D., & Steinschneider, S. (2018). Balancing flood risk and water supply in California: Policy search integrating short-term forecast ensembles with conjunctive use. *Water Resources Research*, *54*(10), 7557–7576. https://doi.org/10.1029/2018WR023177

Peñuela, A., Hutton, C., & Pianosi, F. (2020). Assessing the value of seasonal hydrological forecasts for improving water resource management: Insights from a pilot application in the UK. *Hydrology and Earth System Sciences*, *24*(12), 6059–6073. https://doi.org/10.5194/hess-24-6059-2020

Pianosi, F., Castelletti, A., & Restelli, M. (2013). Tree-based fitted Q-iteration for multi-objective Markov decision processes in water resource management. *Journal of Hydroinformatics*, *15*(2), 258–270. https://doi.org/10.2166/hydro.2013.169

Powell, W. B. (2019). A unified framework for stochastic optimization. *European Journal of Operational Research*, *275*(3), 795–821. https://doi.org/10.1016/j.ejor.2018.07.014

PROGEA. (2022). Retrieved from http://www.progea.net/chi-siamo.php

Rieker, J. D., & Labadie, J. W. (2012). An intelligent agent for optimal river-reservoir system management. *Water Resources Research*, *48*(9), W09550. https://doi.org/10.1029/2012WR011958

Rückstiess, T., Sehnke, F., Schaul, T., Wierstra, D., Sun, Y., & Schmidhuber, J. (2010). Exploring parameter space in reinforcement learning. Paladyn. *Journal of Behavioral Robotics*, *1*(1), 14–24. https://doi.org/10.2478/s13230-010-0002-4

Salazar, J. Z., Reed, P. M., Herman, J. D., Giuliani, M., & Castelletti, A. (2016). A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. *Advances in Water Resources*, *92*, 172–185. https://doi.org/10.1016/j.advwatres.2016.04.006

Semmendinger, K., Lee, D., Fry, L., & Steinschneider, S. (2022). Establishing opportunities and limitations of forecast use in the operational management of highly constrained multiobjective water systems. *Journal of Water Resources Planning and Management*, *148*(8), 04022044. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001585

Stevenson, S., Coats, S., Touma, D., Cole, J., Lehner, F., Fasullo, J., & Otto-Bliesner, B. (2022). Twenty-first century hydroclimate: A continually changing baseline, with more frequent extremes. *Proceedings of the National Academy of Sciences*, *119*(12), e2108124119. https://doi.org/10.1073/pnas.2108124119

Sutton, R. S., & Barto, A. (2020). *Reinforcement learning: An introduction* (2nd ed.). The MIT Press.

Torres, M. O., Tilmant, A., Medina, E. V., Anctil, F., & Ramos, M.-H. (2023). Hydropower system operation and the quality of short-term hydrologic ensemble forecasts. *Journal of Water Resources Planning and Management*, *149*(5), 04023016. https://doi.org/10.1061/JWRMD5.WRENG-5608

Turner, S., Bennett, J. C., Robertson, D. E., & Galelli, S. (2017). Complex relationship between seasonal streamflow forecast skill and value in reservoir operations. *Hydrology and Earth System Sciences*, *21*(9), 4841–4859. https://doi.org/10.5194/hess-21-4841-2017

Turner, S., Xu, W., & Voisin, N. (2020). Inferred inflow forecast horizons guiding reservoir release decisions across the United States. *Hydrology and Earth System Sciences*, *24*(3), 1275–1291. https://doi.org/10.5194/hess-24-1275-2020

Wetterhall, F., Arnal, L., Barnard, C., Krzeminski, B., Ferrario, I., Mazzetti, C., et al. (2020). Seasonal reforecasts of river discharge and related data by the European Flood Awareness System, v4.0. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.768eefc2

Wetterhall, F., & Di Giuseppe, F. (2018). The benefit of seamless forecasts for hydrological predictions over Europe. *Hydrology and Earth System Sciences*, *22*(6), 3409–3420. https://doi.org/10.5194/hess-22-3409-2018

Yang, G., Giuliani, M., & Castelletti, A. (2023). Operationalizing equity in multipurpose water systems. *Hydrology and Earth System Sciences*, *27*(1), 69–81. https://doi.org/10.5194/hess-27-69-2023

Yang, G., Guo, S., Liu, P., & Block, P. (2020). Integration and evaluation of forecast-informed multiobjective reservoir operations. *Journal of Water Resources Planning and Management*, *146*(6), 04020038. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001229

Yang, G., Guo, S., Liu, P., & Block, P. (2021). Sensitivity of forecast value in multiobjective reservoir operation to forecast lead time and reservoir characteristics. *Journal of Water Resources Planning and Management*, *147*(6), 04021027. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001384

Yang, G., Guo, S., Liu, P., Li, L., & Xu, C. (2017). Multiobjective reservoir operating rules based on cascade reservoir input variable selection method. *Water Resources Research*, *53*(4), 3446–3463. https://doi.org/10.1002/2016WR020301

You, J.-Y., & Cai, X. (2008). Determining forecast and decision horizons for reservoir operations under hedging policies. *Water Resources Research*, *44*(11), W11430. https://doi.org/10.1029/2008WR006978

Zaniolo, M., Giuliani, M., & Castelletti, A. (2021a). Neuro-evolutionary direct policy search for multiobjective optimal control. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(10), 5926–5938. https://doi.org/10.1109/TNNLS.2021.3071960

Zaniolo, M., Giuliani, M., & Castelletti, A. (2021b). Policy representation learning for multiobjective reservoir policy design with different objective dynamics. *Water Resources Research*, *57*(12), e2020WR029329. https://doi.org/10.1029/2020WR029329

Zanutto, D. (2023). *zannads/lakeComo_forecast: Release to Zenodo*. Zenodo. https://doi.org/10.5281/zenodo.10123625

Zhao, Q., Cai, X., & Li, Y. (2019). Determining inflow forecast horizon for reservoir operation. *Water Resources Research*, *55*(5), 4066–4081. https://doi.org/10.1029/2019WR025226

Zhao, T., Cai, X., & Yang, D. (2011). Effect of streamflow forecast uncertainty on real-time reservoir operation. *Advances in Water Resources*, *34*(4), 495–504. https://doi.org/10.1016/j.advwatres.2011.01.004

Zhao, T., Yang, D., Cai, X., Zhao, J., & Wang, H. (2012). Identifying effective forecast horizon for real-time reservoir operation under a limited inflow forecast. *Water Resources Research*, *48*(1), W01540. https://doi.org/10.1029/2011WR010623