



Informing the operations of water reservoirs over multiple temporal scales by direct use of hydro-meteorological data



Simona Denaro^a, Daniela Anghileri^b, Matteo Giuliani^a, Andrea Castelletti^{a,b,*}

^a Department of Electronics, Information and Bioengineering, Politecnico di Milano, Piazza L. da Vinci, 32, I-20133 Milano, Italy

^b Institute of Environmental Engineering, ETH Zurich, Stefano-Franscini-Platz 5, CH-8093 Zurich, Switzerland

ARTICLE INFO

Article history:

Received 22 September 2016

Revised 17 February 2017

Accepted 20 February 2017

Available online 21 February 2017

Keywords:

Water reservoirs

Hydrological forecast

Optimal operation

Input selection

Snow

ABSTRACT

Water reservoir systems may become more adaptive and reliable to external changes by enlarging the information sets used in their operations. Models and forecasts of future hydro-climatic and socio-economic conditions are traditionally used for this purpose. Nevertheless, the identification of skillful forecasts and models might be highly critical when the system comprises several processes with inconsistent dynamics (fast and slow) and disparate levels of predictability. In these contexts, the direct use of observational data, describing the current conditions of the water system, may represent a practicable and zero-cost alternative. This paper contrasts the relative contribution of state observations and perfect forecasts of future water availability in improving multipurpose water reservoirs operation over short- and long-term temporal scales. The approach is demonstrated on the snow-dominated Lake Como system, operated for flood control and water supply. The Information Selection Assessment (ISA) framework is adopted to retrieve the most relevant information to be used for conditioning the operations. By explicitly distinguishing between observational dataset and future forecasts, we quantify the relative contribution of current water system state estimates and perfect streamflow forecasts in improving the lake regulation with respect to both flood control and water supply. Results show that using the available observational data capturing slow dynamic processes, particularly the snow melting process, produces a 10% improvement in the system performance. This latter represents the lower bound of the potential improvement, which may increase to the upper limit of 40% in case skillful (perfect) long-term streamflow forecasts are used.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Most existing water reservoirs operating rules are conditioned upon a simple information set, including the day of the year and the storage, and, in few cases, also the previous day inflow (Hejazi et al., 2008). These operating rules are ordinarily designed to perform well under some reference boundary conditions, such as normal hydrological year, flood with a certain return period, given water demands, but can easily become ineffective when these conditions are not met. This situation is becoming more and more frequent due to the increasing variability and uncertainty of reservoir inflows and/or other key variables to the operations, such as water demand or energy price, which are negatively impacting on the expected system performance (e.g., Fowler et al., 2003; Christensen et al., 2004; Georgakakos et al., 2012). One option for increasing the overall reliability of water resources systems is to better ex-

plot the flexibility of operations by conditioning the existing operating rules upon additional information (e.g., Tejada-Guibert et al., 1995; Galelli and Soncini-Sessa, 2010; Gelati et al., 2011).

Previous studies have investigated the role of hydrological information in water management by considering previous period inflow (e.g., Gal, 1979; Maidment and Chow, 1981; Giuliani et al., 2016; Chang and Chang, 2001), current period inflow (e.g., Loucks et al., 1981; Stedinger et al., 1984; Karamouz and Houck, 1987), future inflow scenarios (e.g., Kelman et al., 1990; Kim and Palmer, 1997; Faber and Stedinger, 2001; Kim et al., 2007; Castelletti et al., 2008a), or more sophisticated indexes (e.g., the Palmer drought severity index (Hejazi and Cai, 2011)). This exogenous information, i.e., variables that are observable but are not endogenous in the original problem formulation and, hence, are not modeled, is generally assimilated by a decision model as forecasts about future hydro-climatic and water use conditions.

Skillful forecasts are an asset for improving decision making, but their operational value, i.e., the actual improvement achievable in the system's performance, depends also on other factors, such as the primary purpose of the decisions and the physical charac-

* Corresponding author.

E-mail address: andrea.castelletti@polimi.it (A. Castelletti).

teristics of the controlled system (e.g., Anghileri et al., 2016, and references therein). For example, for a reservoir operated for flood control, the key information to the operations is the time necessary to create a buffer volume for the flood. This is relatively easy to identify as it is related to the physical properties of the reservoir (e.g., reservoir capacity, spillway capacity) and a few hours/days ahead streamflow forecasts might prompt timely anticipatory operations (Valeriano et al., 2010; Raso et al., 2014). Instead, when the reservoir is operated with long-term operating rules, such as for water supply, it is not straightforward to understand which forecasted variable may be useful to design effective hedging rules (e.g., streamflow forecast at a certain point in the future, cumulative inflows over a certain lead time), and to obtain a skillful estimate of such variable (e.g., Zhao et al., 2012). In fact, forecasts are usually skillful over short lead time (from hours to days), but predictability tends to decrease on longer lead times. Seasonal forecasts may be skillful over several months in some specific areas where climate teleconnection, such as El Niño Southern Oscillation (ENSO) (e.g., Hamlet and Lettenmaier, 1999; Sharma et al., 2000; Block and Rajagopalan, 2007; Kalra et al., 2013; Zimmerman et al., 2016), or particular hydrological characteristics, such as snow-dominance (e.g., Koster et al., 2010; Mahanama et al., 2012), may enable predictability over longer lead times.

When considering multi-purpose systems, the picture becomes even more intricate as operations usually need to balance short- and long-term operating targets, e.g., flood control and water supply, (e.g., Sreekanth et al., 2012; Xu et al., 2015). In this case, valuable forecasts may comprise several processes with inconsistent dynamics and disparate levels of predictability. For example, flood control can benefit from hydro-meteorological forecasts over a short lead time, which are generally sufficiently accurate to characterize the fast dynamics of the peak flow and to predict its magnitude and timing. Conversely, water supply, often associated with slow dynamics evolving over longer time scales, would require medium to long-term hydrological forecasts. However, when available, such medium- to long-term forecasts are generally less accurate. Furthermore, the relative importance of different pieces of information may evolve in time as the operator's preference among the operational targets can change with specific temporary conditions, e.g., storms, dry spells (Amigoni et al., 2016). In these circumstances, the use of forecasts to inform decision making may be even more challenging (e.g., Zhao and Zhao, 2014).

An alternative to the adoption of forecasts is represented by the direct use of observational data about the current condition (i.e., state) of the water system. Despite being seldom adopted in water management applications (e.g., Guariso et al., 1984; Faber and Stedinger, 2001; Castelletti et al., 2010; Giuliani et al., 2015), this approach is largely used in several research domains (e.g., artificial intelligence, automotive control, robotics), where decision models are directly conditioned upon observational data to capture the variability of stochastic processes that cannot be accurately modeled/forecasted and would therefore produce detrimental effects on the performance of model-based decisions (e.g., Formentin et al., 2012; Hou and Wang, 2013). In fact, observations about the state of the system have been demonstrated to be sufficient for conditioning optimal operational decisions (Bertsekas, 1976).

However, when dealing with a water system, the state of the real system (not to be confused with the state of the modeled system) is a complex and heterogeneous entity. It includes the water available in the whole basin under multiple forms, including reservoir and lake storages, snow water equivalent, soil water content, and groundwater levels. Beyond the hydrological state, also the climate system has a role (e.g., Turner and Galelli, 2016). Air moisture, temperature, wind speed, and atmospheric pressure contribute to the formation of precipitation events and, as such, they should be also included in the water system state. Moreover, the

state of the real system is spatially distributed and can be potentially inaccessible to observations or measurements. As a consequence, it is unfeasible to have the full knowledge of it. In these cases, feature extraction techniques may be used to identify some observable variables, which represent a surrogate of the unknown real state (or part of it) and maintain some state characteristics that are important to properly control the water system (Bertsekas and Tsitsiklis, 1995). Feature extraction methods can be defined as a set of machine learning tools, which allow the extraction of the more relevant determinants of an observed process out of a large set of candidate variables (e.g., Guyon and Elisseeff, 2003).

In this paper, we quantify the potential improvement in the operations of multi-purpose water reservoirs with short-term and long-term targets by conditioning their operating policy upon exogenous information. In the analysis, we explicitly distinguish between observational data, representing the state of the system or its surrogate, and perfect streamflow forecasts. We use the Information Selection and Assessment (ISA) framework proposed by Giuliani et al. (2015) to compute the relative contribution of exogenous information to the system performance by automatically selecting the most valuable information for informing the operating policies. This automatic selection, performed via feature extraction techniques, allows solving complex problems with multiple objectives over different time dynamics that would challenge similar existing approaches (e.g., Liu et al., 2006).

First, we estimate the performance achievable by a basic information set, including the day of the year and the reservoir level. Then, we use the ISA framework to select the most informative state surrogates, which are incrementally included as additional arguments of the operating policy. The contribution of these different pieces of information represents a substantiation of the potential improvement associated to the use of observational data collected from the existing monitoring system. Such improvement is finally compared with the upper bound of the system performance obtained by conditioning the reservoir operation with inflow forecasts on different lead times. We use a retrospective streamflow dataset as perfect forecast, which is expected to approximate the use of a sophisticated forecasting system (Zhao et al., 2011) while also removing possible modeling biases in the construction of the forecasts.

We demonstrate our approach on Lake Como (Italy). The lake is regulated balancing multiple and conflicting needs, primarily flood control on the lake shores and water supply to downstream irrigated agriculture and several run-of-the-river power plants (Guariso et al., 1986). The lake catchment is located in the Italian Alps and is characterized by a mixed snow- and rain-dominated hydrology (e.g., Castelletti et al., 2010; Anghileri et al., 2011). The rain-dominated component has fast dynamics and affects mainly flood control. Autumn is the flood season, but some flooding events may occur in late spring, due to intense snow-melt peaks. The snow-dominated component, instead, has slow dynamics and sustains the lake inflow mainly in spring, while fading out during summer. The amount and timing of the snow-melt is relevant for hedging the water supply during the irrigation season, which goes from spring to autumn. The two operating objectives are conflicting because the lake operator needs to build up the reserve for water supply purposes while keeping the appropriate flood pool to buffer flood peaks. In such a context, both forecasts on short and seasonal lead times may be valuable. The predictability in this catchment is partially provided by snow accumulation and melting processes and spans from spring to early summer. Forecasts on longer lead times are instead less reliable. The existing monitoring system collects observations of precipitation, snow water equivalent, and storage of the several hydropower reservoirs located in the lake catchment, which represent candidate state surrogates to inform the lake operations. The multi-purpose nature

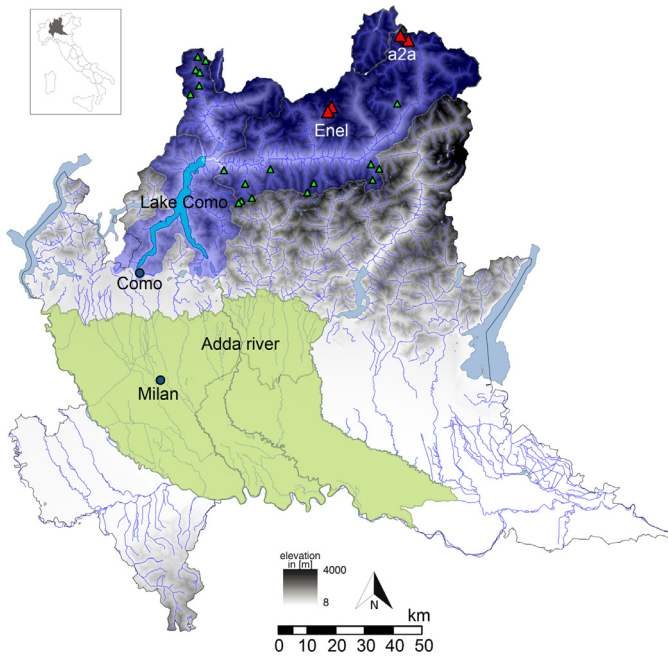


Fig. 1. Map of the study area: Lake Como, the watershed area (violet), and downstream agricultural districts (green). The triangles denote hydropower reservoirs (the red ones are the four reservoirs considered in this study). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of the reservoir, the complex catchment structure, and different processes' dynamics, comprising both fast and slow process, make Lake Como an interesting case study to test the proposed methodology.

Summarizing, the main novel contributions of the paper are: *i*) we identify the most informative exogenous dataset containing complementary information for designing an optimal reservoir operating policy able to balance short- and long-term operating objectives; *ii*) we design a multi-purpose reservoir operating policy directly conditioned upon the selected exogenous information, thus providing a demonstration of direct use of raw hydro-meteorological data to inform decision making in water management applications; *iii*) we quantify the relative contribution of observations of the state of the system, or its surrogate, and perfect forecasts of future water availability in improving reservoir operations.

The paper is organized as follows: the next section introduces the Lake Como study site, followed by the description of the methodologies adopted for quantifying the operational value of the considered exogenous information. Numerical results are then reported first on the analysis of observations of state surrogates and, then, on inflow forecasts. Final remarks and directions for further research are discussed in the last section.

2. Study site

2.1. System description

Lake Como (Fig. 1) is the third largest lake in Italy with a total volume of 23.4 km³, of which 254 Mm³ are regulated through a dam on the effluent Adda river. The lake is fed by a 4,552 km² watershed characterized by the mixed snow-rain dominated regime typical of the southern Alps. The annual average hydrograph (Fig. 2) shows relatively dry winters and summers, and higher flows in spring and autumn due to snow-melt and precipitation, respectively.

The lake releases are controlled since 1946 with the primary, twofold purpose of flood protection along the lake shores and water supply to the downstream users. The lake serves eight run-of-the-river hydropower plants and a dense network of irrigation canals, which distribute the water to four agricultural districts with a total surface of 1,400 km² mostly cultivated with maize. Historically, water availability has not been a major limiting factor to the economic development of this area. Rather, water governance was mostly concerned with flood risk management. Yet, in recent decades, climate change has been showing its potential negative impact in a number of situations (e.g., García-Herrera et al., 2010). For example, two severe droughts in 2003 and in 2005 generated acute crop failures and exacerbated the conflicts between agriculture and the other sectors (Anghileri et al., 2012).

The predicted increase in the frequency and intensity of water crises over the next years (Lehner et al., 2006) represents a major challenge to the operations of the Lake Como system. As the lake regulation capacity is relatively small, being the annual average inflow-to-capacity ratio approximately equal to 6% (Anghileri et al., 2011), the operator is able to shift only small volumes of water in time. To satisfy the summer water demand peak, the historical regulation operated the lake to be, approximately, at full capacity between June and July (see Fig. 2). The projected anticipation of the snow melt caused by increasing temperature, coupled with the predicted decrease of water availability in the summer period (Forzieri et al., 2014), would require storing additional water and for longer periods. Yet, this strategy would induce an increase of the flood risk associated to high water levels and, consequently, would exacerbate the conflict between the short-term flood protection target and the long-term water supply objective.

Although describing the fast dynamics of rainfall is generally a challenge in the highly spatially heterogeneous Alpine systems (Cherubini et al., 2002), collecting and using information on the slow dynamics of snow can potentially contribute to improving the lake regulation in terms of water supply. In fact, a large volume of water is accumulated as snow during the winter, as shown in Fig. 2 by the average annual pattern of Snow Water Equivalent (SWE). The associated snow-melt represents the main contribution to the lake's seasonal storage supporting the downstream water demand. The accurate characterization of snow dynamics, especially during the melting season, might be a valuable information to partially anticipate the summer water availability and, hence, aid the implementation of effective hedging rules.

To complement local observations from existing ground stations, which provide a very coarse and uneven coverage of the snow cover patterns and the associated water content, since 2006 the Regional Agency for Environmental Protection (ARPA - Agenzia Regionale per la Protezione dell'Ambiente) produces weekly estimates of SWE at the basin scale. These are obtained through a hybrid procedure combining snow height and temperature data from ground stations, manual measures of snow density in few specific locations, satellite retrieved data of snow cover from MODIS, and model outputs for spatially interpolating these data (Bellingeri et al., 2006).

Several artificial hydropower reservoirs are located in the upstream part of the catchment (green and red triangles in Fig. 1). Their total storage capacity sums up to 545 Mm³, more than twice the active capacity of the lake. The reservoirs' operation affects the seasonal inflow pattern to the lake, as they retain a fraction of the snow-melt in spring and summer, which will be released in the following winter when the energy prices are higher. The size of the reservoirs instead does not allow for inter-annual water transfer. The regulation of the lake, which is currently uncoordinated with respect to the operations of the hydropower reservoirs (Anghileri et al., 2012), could therefore benefit from using information that describes the conditions of these reservoirs. In this study, we con-

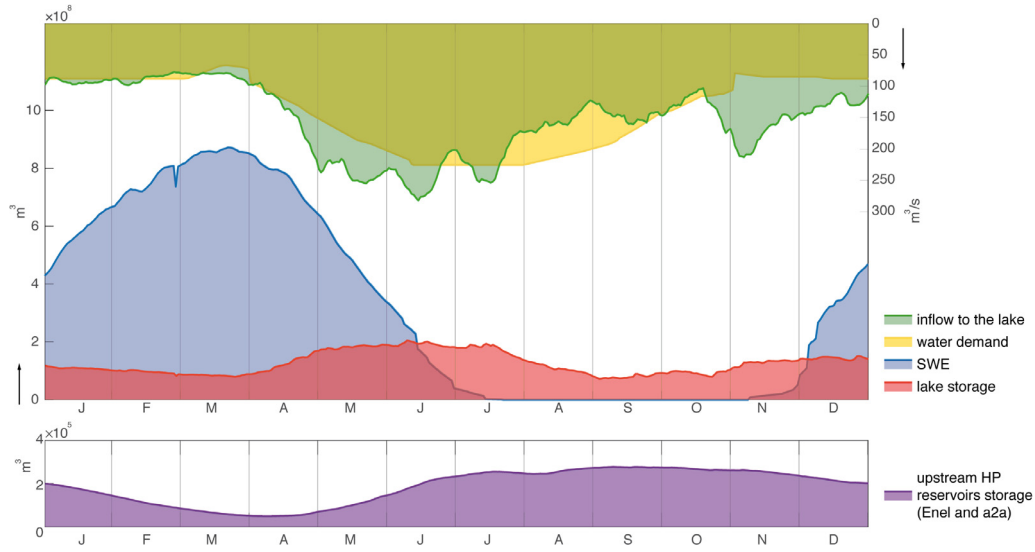


Fig. 2. Main components of the hydrological cycle in the study area. The patterns represent moving averages computed from observed data over the period 2006–2013.

sider the two major reservoir systems in the basin (red triangles in Fig. 1) operated by ENEL and a2a, which together represent more than half the total hydropower storage in the watershed.

2.2. The lake dam operation

The operations of Lake Como can be modeled as a daily operating policy, which provides the volume of water to be released over the next 24 hours within the zone of operation discretion delimited by physical and normative constraints.

As discussed in the above section, the lake regulation aims at satisfying the two main objectives of flood control and water supply. On the basis of previous works (Anghileri et al., 2011; Culley et al., 2016), we use the following objective formulations computed over the evaluation horizon H :

- Flood control: the annual average number of days [day/year] when a flood occurs, computed as

$$J^{flo} = \frac{1}{H} \sum_{t=0}^{H-1} g_{t+1}^{flo} \quad \text{where}$$

$$g_{t+1}^{flo} = \begin{cases} 1 & \text{if } h_{t+1} > \bar{h} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where h_t is the lake level and \bar{h} is the flood level threshold equal to 1.24 m, which delimits the occurrence of a flood in Como city.

- Water supply: the average squared daily water deficit [(m³/s)²/day], computed as

$$J^{irr} = \frac{1}{H} \sum_{t=0}^{H-1} g_{t+1}^{irr} \quad \text{where}$$

$$g_{t+1}^{irr} = \left(\max(w_t - (r_{t+1} - q^{ef}), 0) \right)^2 \quad (2)$$

where w_t is the daily water demand (yellow area in Fig. 2), r_{t+1} is the release from the lake, and q^{ef} is the environmental flow equal to 5 m³/s. The squared power minimizes crop vulnerability by penalizing higher shortages, which can compromise the crop growth, with respect to more frequent but smaller shortages, which are less dangerous to the crops (Hashimoto et al., 1982).

Designing the Lake Como optimal operation consists in finding the best operating policy p^* such that

$$p^* = \arg \min_p \mathbf{J} = |J^{flo}, J^{irr}| \quad (3)$$

where the policy p is defined as a closed loop control policy (Castelletti et al., 2008b) that determines the release decision $u_t = p(d_t, h_t, \mathbf{I}_t)$ at each time step t as dependent on the day of the year d_t , the current level of the lake h_t (i.e., the modeled state), and, possibly, a vector of exogenous information \mathbf{I}_t (i.e., state surrogates or/and inflow forecast). The release decision u_t may differ from the actual release r_{t+1} in specific situations, where normative and physical constraints may restrict the ability to subjectively regulate the lake (Piccardi and Soncini-Sessa, 1991).

3. Methods and tools

We use the Information Selection and Assessment framework (Giuliani et al., 2015) to quantify the value of exogenous information for improving the operations of Lake Como. The framework is composed of three main building blocks as illustrated in Fig. 3 and detailed in the next sections.

3.1. Expected Value of Perfect Information

We define the Expected Value of Perfect Information (EVPI) as the performance improvement that could be achieved assuming to have full and perfect information on the future at the moment when operational decisions must be made. Under this assumption, Problem 3 can be solved with the entire trajectory of the disturbances (i.e., the inflow) deterministically known over the entire evaluation horizon H , producing a sequence of optimal reservoir release decisions $u_{[0,H-1]}^{POP}$. The problem is solved via Deterministic Dynamic Programming as further detailed in Section 3.4. Conceptually, this solution can be seen as a closed loop operating policy, called Perfect Operating Policy (POP), conditioned upon the current system state along with perfect information on the future disturbances.

Since the Perfect Operating Policy performance \mathbf{J}^{POP} is an absolute measure of the system performance, which, intrinsically, depends on the characteristics of the system under study, the EVPI has to be estimated by comparing \mathbf{J}^{POP} with the value of the objective functions that could be obtained by an operating policy relying on a basic set of information. In this work, we consider a Basic

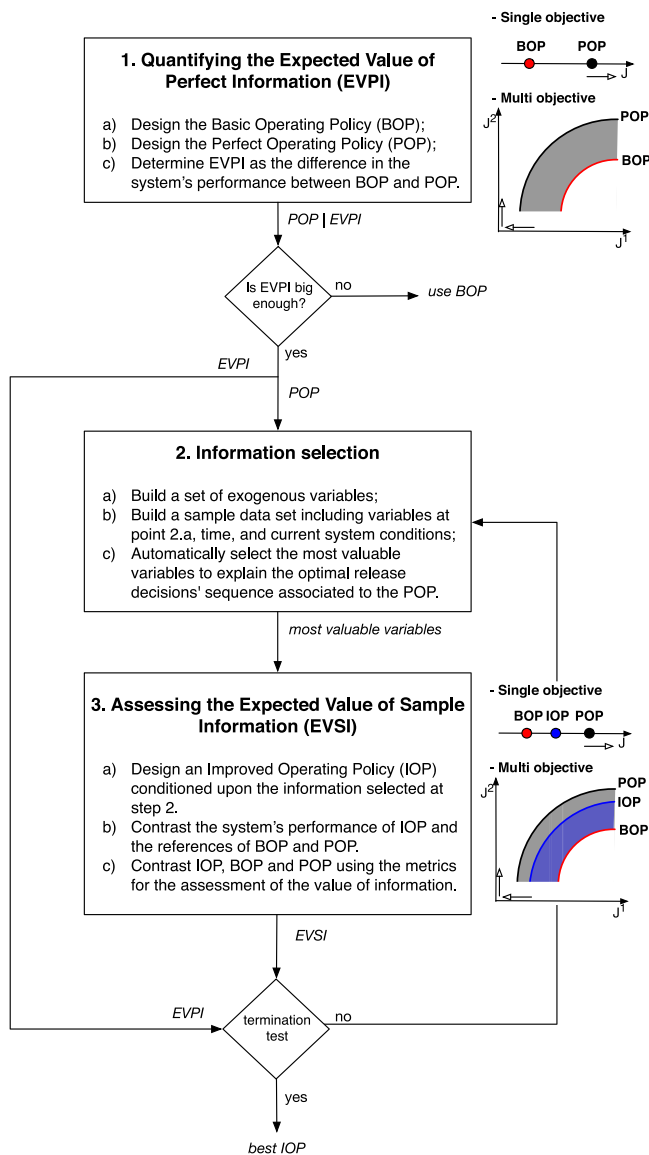


Fig. 3. Flowchart of the Information Selection and Assessment (ISA) framework (Giuliani et al., 2015).

Operating Policy (BOP) that approximates the historical lake regulation and is defined as an operating policy dependent on the day of the year and the lake level, i.e., $u_t = p(d_t, h_t)$ (see Section 4.1).

Given these two reference solutions, the EVPI can be estimated as the difference between their performance. In case of single-objective problems, the EVPI is simply given by the scalar difference $J^{POP} - J^{BOP}$. For multi-objective problems, the solution of Problem 3 is not unique but rather a Pareto optimal set, where the performance \mathbf{J}^{POP} and \mathbf{J}^{BOP} are objective function vectors. Here, the estimation of the EVPI requires to evaluate multiple metrics (Zitzler et al., 2000) accounting for (i) the convergence of the final solution to the true Pareto front that would result from an optimal-method, (ii) the coverage of the non-dominated space (diversity), (iii) the extent of the non-dominated front, i.e., the degree over which the solution space is evenly covered by the final solution. Among the commonly used metrics adopted in the literature (see Maier et al. (2014), and references therein), we use the hypervolume indicator (HV) which captures both the proximity of the Pareto front J^{BOP} to the ideal one J^{POP} as well as the distribution of the BOP solutions in the objective space (Zitzler et al., 2003). Moreover, to ac-

count for the specific advance toward a pre-specified target solution of interest, we also compute two additional metrics measuring the proximity between the target solution and the closest point of the Pareto front under exam (D_{min}) and the average Euclidean distance of the entire Pareto front under exam from the target solution (D_{avg}) (see Giuliani et al., 2015, for details).

3.2. Information selection

When the EVPI is large, it might be worth exploring whether the addition of further information to the decision process allows closing the gap between Basic and Perfect Operating Policies. To this purpose, we identify a set of information $\mathbf{I}_t \in \Xi_t$, known at time t when operational decisions are made, which can be effective in replacing the perfect information about the future realization of the disturbances. As anticipated in the Introduction, this might be inflow forecasts on different lead times, when available, or any other observed variable which might act as state surrogate, or both.

In general, the set Ξ_t comprises any exogenous variable that can be valuable when making decisions. Since the set of candidate variables Ξ_t can be rather vast, including redundant and collinear variables, we adopt a feature selection technique to select which variables are more informative. The choice of the appropriate technique, among the many available, should pursue three desirable features: (i) modeling flexibility for capturing non-linear functions, because the functional relationship between the candidate inputs Ξ_t and the output $u_{[0,H-1]}^{POP}$ is usually unknown a priori; (ii) computational efficiency for dealing with potentially large data-sets, comprising long time series and many candidate variables; (iii) scalability with respect to the number of candidate input variables for handling numerous input variables with different ranges of variability. According to the guidelines provided in Galelli et al. (2014), we used the Iterative Input Selection (IIS) algorithm (Galelli and Castelletti, 2013b) combined with Extremely Randomized Trees (Galelli and Castelletti, 2013a; Geurts et al., 2006), which holds all the three properties mentioned above. The IIS algorithm is used to select the subset $\mathbf{I}_t \in \Xi_t$ of most informative exogenous variables, which better characterizes the optimal sequence of release decisions $u_{[0,H-1]}^{POP}$.

3.3. Expected Value of Sample Information

We define the Expected Value of Sample Information (EVS) as the performance improvement that could be achieved when the selected information \mathbf{I}_t is used to inform operational decisions, i.e., to design the Improved Operating Policies (IOPs) determining the release decision as $u_t = p(d_t, h_t, \mathbf{I}_t)$. In general, we expect the Improved Operating Policies to fill the performance gap between the Basic and Perfect Operating Policies and, possibly, to produce a performance \mathbf{J}^{IOP} as close as possible to \mathbf{J}^{POP} .

Note that the ISA procedure is an iterative process (see Fig. 3). At first, we consider only the first variable selected by the IIS algorithm, assuming that it also has the highest potential to improve the operating performance. We design a set of Improved Operating Policies conditioned on this variable only, and estimate the corresponding EVSI by comparison with the BOP performance. We then iterate the procedure by incrementally adding variables to the information vector \mathbf{I}_t , designing the associated IOPs, and evaluating the corresponding EVSI. The iterative procedure stops when either the attained performance is satisfactory (i.e., we obtain performance close enough to \mathbf{J}^{POP}) or the marginal improvement in the EVSI between two consecutive iterations is negligible.

Table 1
Set of candidate variables as system state surrogates.

name	description	period
SWE_t	snow water equivalent ARPA estimates [m^3]	2006–2013
P_t	precipitation measured in the catchment [mm]	2006–2013
SM_t	snow melting derived by SWE_t [m^3]	2006–2013
z_t^{iso0}	freezing level from temperature regression [masl]	2006–2013
$S_{t, HP}$	total storage of upstream reservoirs [m^3]	2006–2013
$r_{t, HP}$	total release from upstream reservoirs [m^3/s]	2006–2013

Table 2
Set of inflow forecasts over different lead times.

name	description	period
$lead7_t$	cumulative future inflow over 1 week	2006–2013
$lead14_t$	cumulative future inflow over 2 weeks	2006–2013
$lead21_t$	cumulative future inflow over 3 weeks	2006–2013
$lead30_t$	cumulative future inflow over 1 month	2006–2013
$lead37_t$	cumulative future inflow over 1 month and 1 week	2006–2013
$lead44_t$	cumulative future inflow over 1 month and 2 weeks	2006–2013
$lead51_t$	cumulative future inflow over 1 month and 3 weeks	2006–2013
$lead60_t$	cumulative future inflow over 2 months	2006–2013

3.4. Experiment strategy and setting

Following the ISA framework, we first identify the set of Basic Operating Policies and Perfect Operating Policies, which allows quantifying the Expected Value of Perfect Information. Then, we search for the most valuable exogenous information among a set of candidate observational data, collected from the existing monitoring system, which potentially surrogate the current system state over slow and fast dynamics. Finally, we repeat the procedure replacing these observational data with a set of streamflow forecasts over different lead times to estimate the upper limit of the potential performance improvement. The resulting Expected Values of Sample Information in the two cases are compared one to another by contrasting the associated Improved Operating Policies performance and by computing the corresponding metrics HV, D_{min} , and D_{avg} (see Section 3.1). In addition, we analyze the dynamics of the system produced by the simulation of different operating policies to gain further insights on how the selected information is used for triggering specific operational strategies that allow improving the system performance, especially differentiating between the short-term objective (i.e., flood control) and the long-term one (i.e., water supply).

The experimental setting is characterized as follows:

- **Observations of state surrogates:** this set of candidate exogenous information (Table 1) collected over the period 2006–2013, includes meteorological variables (i.e., daily average precipitation over the entire watershed interpolated from punctual precipitation measures using the Thiessen method, and estimates of the zero degree isotherm, obtained interpolating levels and temperature values from ground stations), snow data (i.e., SWE estimates provided by ARPA as well as values of the weekly snow melting derived from the SWE estimates), and information on the hydropower reservoirs (i.e., total storage and total release from the upstream reservoirs of the two main companies operating in the basin, i.e., Enel and a2a).
- **Inflow forecasts:** this set of candidate exogenous information (Table 2) includes forecasts of the lake inflows for the period 2006–2013, computed over different lead times, ranging from one week to two months. In the absence of a reliable long-term streamflow forecasting system for this catchment, we assume a retrospective streamflow dataset as perfect forecast generated from a hypothetical forecasting system, which also removes

possible modeling biases in the construction of the forecasts (Zhao et al., 2011).

- **Evaluation horizon:** we use the time horizon 2006–2013 to run our analysis as the SWE monitoring data are available only since 2006. This time horizon is used for both the optimization and evaluation of the policy performance, thus ensuring that the Perfect Operating Policies represent the upper bound in terms of system performance.
- **Perfect Operating Policies:** the set of POPs was designed via Deterministic Dynamic Programming (Bellman, 1957) over the evaluation horizon 2006–2013. The weighting method (Gass and Saaty, 1955) is used to convert the 2-objective problem into a single-objective one via convex combinations. The exploration of the trade-off is performed by varying the weights used in the objectives' aggregation.
- **Basic and Improved Operating Policies:** we designed both the BOP and IOP sets by solving Problem 3 with the Evolutionary Multi-Objective Direct Policy Search (EMODPS) method (Giuliani et al., 2016), an approximate dynamic programming approach that combines direct policy search, nonlinear approximating networks, and multi-objective evolutionary algorithms. In particular, we parameterized the operating policy as Gaussian radial basis functions, because they have been demonstrated to be effective in solving this type of multi-objective policy design problems (Giuliani et al., 2014a; 2014b), particularly when exogenous information is directly used for conditioning the operations (Giuliani et al., 2015). To perform the optimization, we use the self-adaptive Borg Multi-Objective Evolutionary Algorithm (MOEA) (Hadka and Reed, 2013), which has been shown to be highly robust in solving multi-objective optimal control problems, where it met or exceeded the performance of other state-of-the-art MOEAs (Zatarain-Salazar et al., 2016). Each optimization was run for 2 millions function evaluations over the evaluation horizon 2006–2013. To improve solution diversity and filter out the randomness of both the initial population set and the optimization operators, the solution set for each optimization is the result of 40 random optimization trials. The final set of Pareto optimal policies for each experiment is defined as the set of non-dominated solutions from the results of these optimization trials.

4. Numerical results

4.1. Quantifying the Expected Value of Perfect Information

The Expected Value of Perfect Information (EVPI) represents the potential space for improvement generated by the knowledge of perfect information of the future inflows trajectories. The computation of the EVPI requires contrasting the performance of the Perfect Operating Policies (POPs) and the Basic Operating Policies (BOPs). The former are defined as ideal operating policies dependent on the current system conditions along with perfect information on future disturbances (i.e., inflows). The latter are defined as operating policies conditioned upon the day of the year (d_t) and lake level (h_t), assuming that these two variables are the ones the real operator mostly considers in his daily operations of the lake.

We validated this assumption by modeling the historical release time series on the evaluation horizon 2006–2013 by means of Extremely Randomized Trees (Galelli and Castelletti, 2013a): numerical results reveal that nearly 65% of the variance of the historical releases can be explained by a tree-based model using as input the day of the year d_t , which informs about the seasonality of the downstream water demand, and lake inflow patterns. An additional 20% of the variance in the historical releases is explained by a model that, in addition to the day of the year, also considers as input the reservoir level h_t . This latter has the twofold role of in-

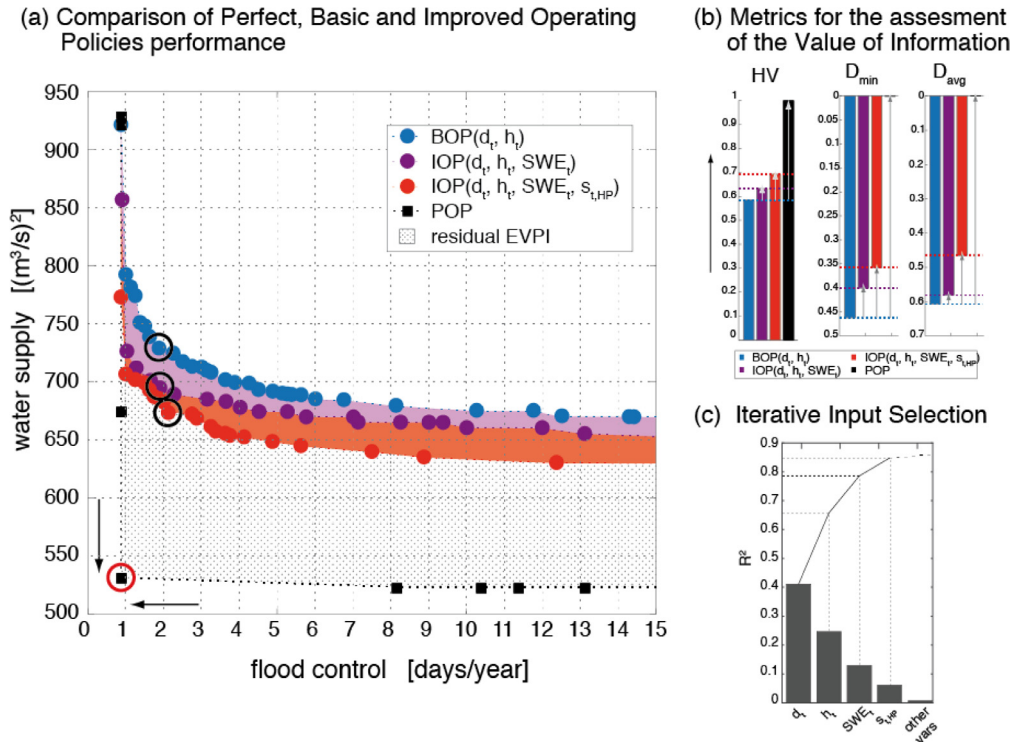


Fig. 4. Performance obtained by the Perfect Operating Policies, Basic Operating Policies, and two Improved Operating Policies relying on SWE (SWE_t) and alpine hydropower storage ($s_{t,HP}$) as surrogates of the actual system state. Panel (a) shows the policies' performance in the objective space; panel (b) shows the value of information quantified by the metrics described in Section 3.1; panel (c) shows the IIS results as average cumulative performance of the model describing the optimal release trajectory $u_{[0,H-1]}^{POP}$. The red circled square in panel (a) corresponds to the target POP, while the black circled points are the BOP and IOP analyzed in Fig. 5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

forming about the flood buffer potential and the water availability useful for water supply. The resulting model explain around 85% of the historical lake release time series, with the remaining 15% that can be arguably attributed to the sporadic use of flood forecasts and other contingencies (Todini, 2014).

Fig. 4(a) compares the performance of the POPs (black squares) and the one of the BOPs (blue circles), where the arrows indicate the direction of increasing preference, with the best solution located in the bottom-left corner of the figure. In the same figure, the purple and red circles represent the performance of the Improved Operating Policies, which will be discussed in the next section. Visual comparison of the black and blue Pareto fronts shows that the potential space for improvement generated by the knowledge of perfect information of the future inflows is relevant, especially for reducing the water supply deficit reported on the y-axis of the figure.

It is worth noting that the shape of the POP Pareto front is sharp-cornered. This means that under the assumption of perfect inflow foresight over the entire evaluation horizon, the existing conflict between the two operating objectives would disappear. In fact, under this assumption, the operator can successfully store in the lake any excess of water to satisfy the downstream demand during dry periods, always guaranteeing an adequate flood pool to buffer the inflow peaks. In the absence of such perfect knowledge of the future, the use of additional information in the system operation might have the potential to both reduce the uncertainty about the future system's conditions and mitigate the existing conflict. Given this sharp-cornered Pareto front, we select the almost no-conflict solution (red circled square in Fig. 4(a)) as a target solution representing a fair balance between the objectives.

The values of the three metrics introduced in Section 3.1 confirm the visual comparison of the POP and BOP solutions and provide a quantitative assessment of the EVPI (Fig. 4(b)). The differ-

ence in the hypervolume indicator, which measures the gap between the POPs and the BOPs sets, is equal to 0.42, indicating a potential improvement of 70% produced by the ideal perfect information on future inflows. Also the large values of D_{min} and D_{avg} point out a limited exploration of the objective space around the selected target solution.

4.2. Assessing the Expected Value of Sample Information of state surrogates

The values of EVPI discussed in the previous section show a great potential for improving the system operations by enlarging the basic information system of the BOPs. In this section, we focus on searching the most valuable information among the set of state surrogates listed in Table 1.

Fig. 4(c) illustrates the results of 100 runs of the IIS algorithm in terms of the average performance attained by the underlying regression model in describing the optimal release decisions sequence $u_{[0,H-1]}^{POP}$, measured in terms of cumulative coefficient of variation (R^2). The repetition of the experiments aims at filtering the randomness associated to the construction of the extra-trees models used by the IIS algorithm (Galelli and Castelletti, 2013b). We tentatively stopped the input selection at the first four most informative variables, which allows explaining cumulatively nearly 85% of the target releases. Not surprisingly, the day of the year and the lake level are selected as the most relevant drivers, explaining up to the 65% of the optimal release sequence. The third selected variable is the SWE estimate, accounting for a R^2 contribution of about 12%. The fourth selected variable is the total storage of the upstream hydropower reservoirs ($s_{t,HP}$), which further adds another 7%.

As expected, two low frequency variables are selected to describe the slow-dynamics in the basin: snow information is a key

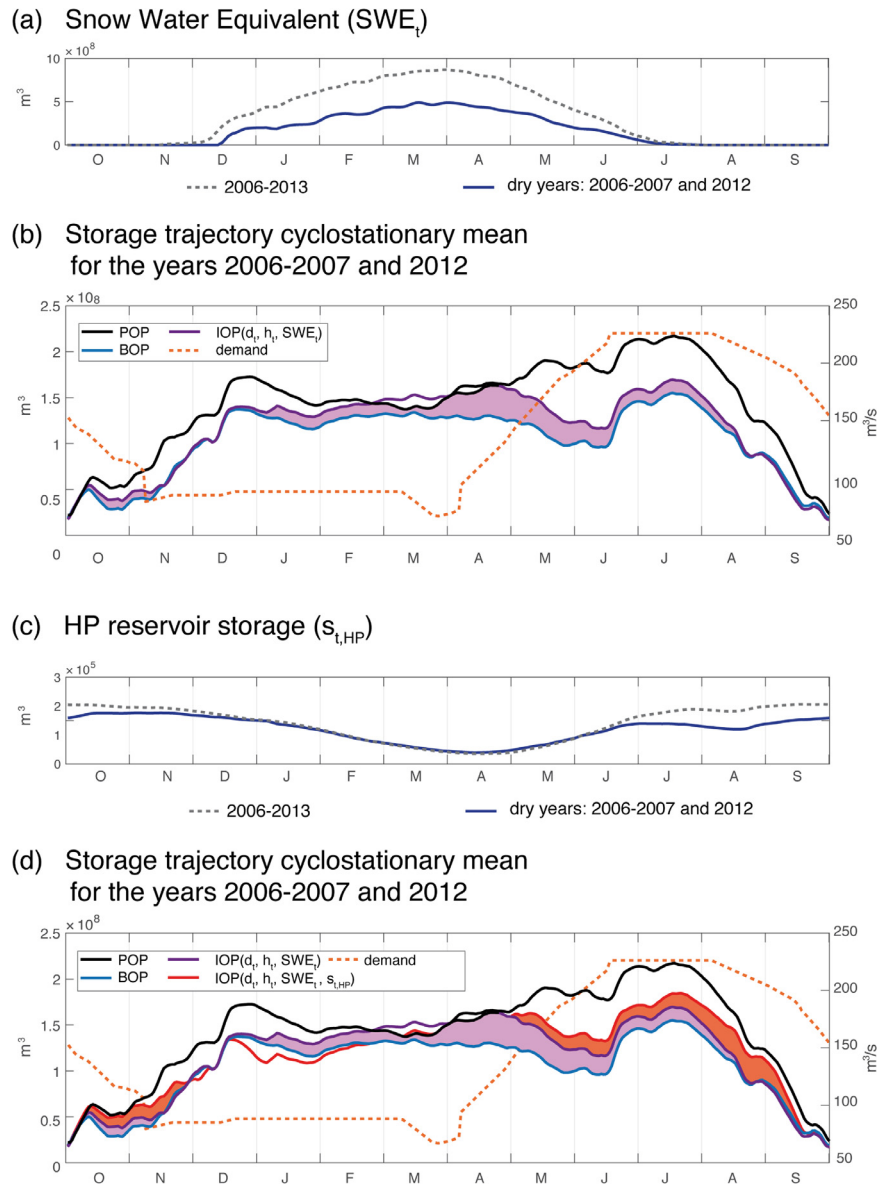


Fig. 5. Analysis of the Lake Como storage trajectory in dry years under different Improved Operating Policies (see Fig. 4) conditioned on observations of state surrogates (i.e., SWE estimate and hydropower storage). As a reference, the trajectory of the target POP and the selected BOP are also illustrated.

element for the lake operations as it approximates the amount of water stored in the upper watershed that will be made available in the next melting season, when the yearly inflow peak is reached. It is therefore a good proxy for discriminating a wet from a dry summer, triggering effective hedging strategy. The information on the total storage of upstream reservoirs is also relevant and impacts on both the operating goals: if the upstream reservoirs happen to be full in summer, the operation is informed that this retained water is potentially available in the basin and expected to reach the lake in the coming season. On the other hand, if the upstream reservoirs happen to be empty in spring or in autumn, the lake operator can rely on a bigger buffer capacity when a flood event kicks in.

Following the ISA procedure, the two selected exogenous variables (i.e., SWE_t and $s_{t,HP}$) are incrementally included as arguments of the lake operating policies for designing two different sets of Improved Operating Policies. The IOPs conditioned upon SWE information are represented by the purple Pareto front in Fig. 4(a), where the corresponding EVSI (i.e., the purple area between this BOP and IOP solutions) shows a performance improvement that is

more emphasized in terms of water supply rather than flood control. This can be explained by the seasonal slow dynamics of snow related processes, which are more informative for hedging rules but do not contribute in describing the fast dynamics associated with flood events. The EVSI of SWE is quantitatively verified by the improvement registered in all the metrics reported in Fig. 4(b): when moving from $BOP(d_t, h_t)$ to $IOP(d_t, h_t, SWE_t)$, HV increases by 8.4%, while D_{min} and D_{avg} decrease by 13.3% and 4.2%, respectively.

The introduction of the storage of the upstream hydropower reservoirs ($s_{t,HP}$) produces a further improvement in the performance of the associated set of IOPs (red circles in Fig. 4(a)). This improvement may reflect the effect of integrating $s_{t,HP}$ with SWE information thus allowing the lake operation to distinguish that part of snow melt streamflow that does not naturally flow through the river network, but is diverted and stored into the reservoirs to be released in later periods, and, dually, the amount that can freely reach the lake inlet. Again, the improvement is registered mainly in terms of water supply, as quantified by the 9.6% increase in the hypervolume indicator and by the 10.7% and 19.9% decrease of D_{min} and D_{avg} , respectively.

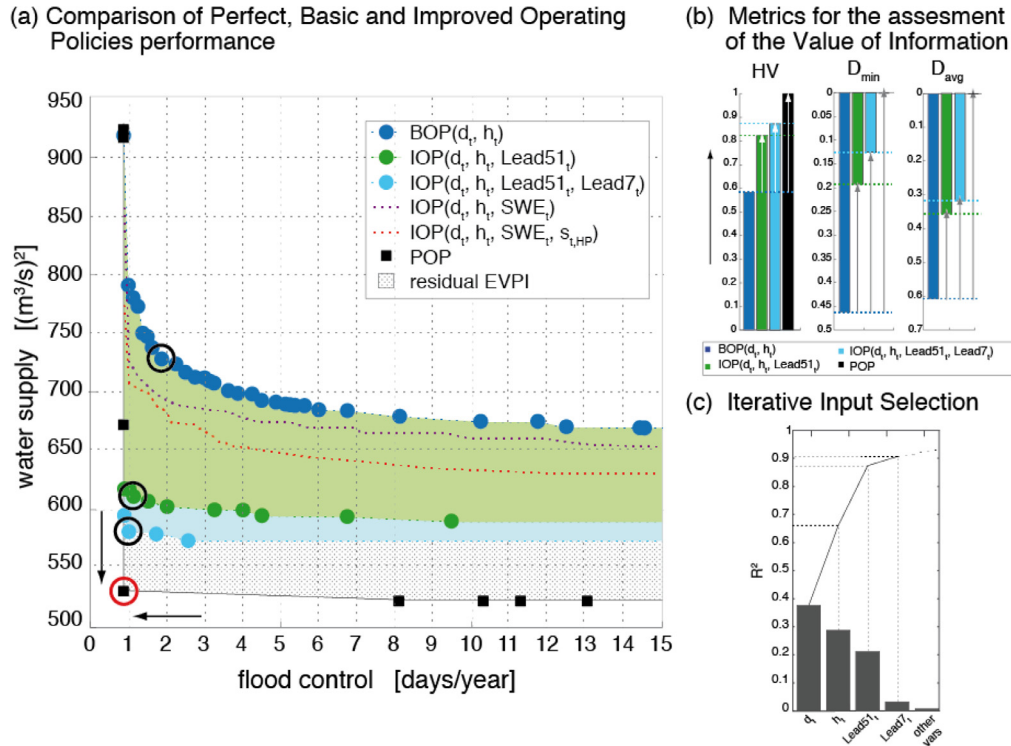


Fig. 6. Performance obtained by the Perfect Operating Policies, Basic Operating Policies, and two Improved Operating Policies relying on streamflow forecasts on different lead times (*Lead51_t* and *Lead7_t*). Panel (a) shows the policies' performance in the objective space; panel (b) shows the value of information quantified by the metrics described in Section 3.1; panel (c) shows the IIS results as average cumulative performance of the model describing the optimal release trajectory $u_{[0,H-1]}^{POP}$. The red circled square in panel (a) corresponds to the target POP, while the black circled points are the BOP and IOP analyzed in Fig. 7. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To investigate the individual operational contribution by each exogenous information, we analyze the trajectory of lake storage obtained via simulation of the different Improved Operating Policies contrasted with the ones produced by the Basic and the Perfect Operating Policies (Fig. 5). In each Pareto optimal set, we select the solution having the smallest Euclidean distance from the target POP, which are identified by the large black circles in Fig. 4(a). This analysis of the system dynamics aims to better understand which management strategy is triggered by the use of a specific information and at what time of the hydrological year (which goes from October to September) the selected information is mostly useful to approximate the ideal optimal operations of the target POP. Because most of the operational improvement concerns the water supply objective, we focus on the drier years in the considered horizon, namely 2006, 2007 and 2012.

The Improved Operating Policy informed by the SWE estimate (Fig. 5(a)) produces a trajectory of lake storage that, especially from January to August, is closer to the target trajectory than the one obtained with the BOP. This difference is highlighted by the purple area in Fig. 5(b), and corresponds to the period where the SWE information is available and snow melt takes place (Fig. 5(a)). The informative contribution of SWE actually lasts a bit longer because of the storing capacity of the lake, which allows to store the snow melting contribution for few months and to cover most of the summer water demand peak (see dashed red line in Fig. 5(b)). When also the information of hydropower storage is considered (Fig. 5(c)), the operational contribution of the enlarged information system spans from May to December, with an early small contribution in March, as shown by the red area in Fig. 5(d). The two pieces of information can be seen as complementing each other until the end of the melting season, when they provide knowledge on the snow melting that is not detained in the upstream reservoirs and is, therefore, free to immediately reach the lake. In late

summer and autumn, in the absence of SWE, the information on hydropower storage is still valuable as it informs the lake operations about the potential capacity of the watershed to buffer the coming abundant rains. This allows for risking higher water levels in the lake to support the water supply in winter to satisfy the demand of winter crops and run-of-the-river power plants. Similar conclusions can be drawn from the analysis of the wet years (not shown), yet the observed behavior is less pronounced.

4.3. Assessing the Expected Value of Sample Information of inflow forecasts

The results discussed in the previous section shows that there is still a large gap between the best Improved Operating Policies and the Perfect Operating Policies (residual EVPI in Fig. 4(a)). This is possibly due to the lack of information on the timing of the inflow to the lake, as well as on the amount of water that will be available on lead times longer than the anticipation capacity provided by SWE and hydropower storage. To validate this hypothesis, we repeated the ISA procedure a second time, replacing the set of candidate state surrogates with a set of perfect inflow forecasts on different lead times (see Table 2).

As in the previous case, we perform 100 runs of the IIS algorithm for selecting the most valuable lead times in informing the operating policy as the ones that contribute most in explaining the optimal trajectory $u_{[0,H-1]}^{POP}$. The results illustrated in Fig. 6(c) show that, after the day of the year and the lake level, the IIS algorithm consistently selects a combination of long (i.e., *Lead51_t*) and short (i.e., *Lead7_t*) lead time information. These variables contribute an average explained variance of 21% and 4%, with the final set of information explaining more than 90% of the optimal trajectory $u_{[0,H-1]}^{POP}$. These results reflect the double temporal dynamics of the

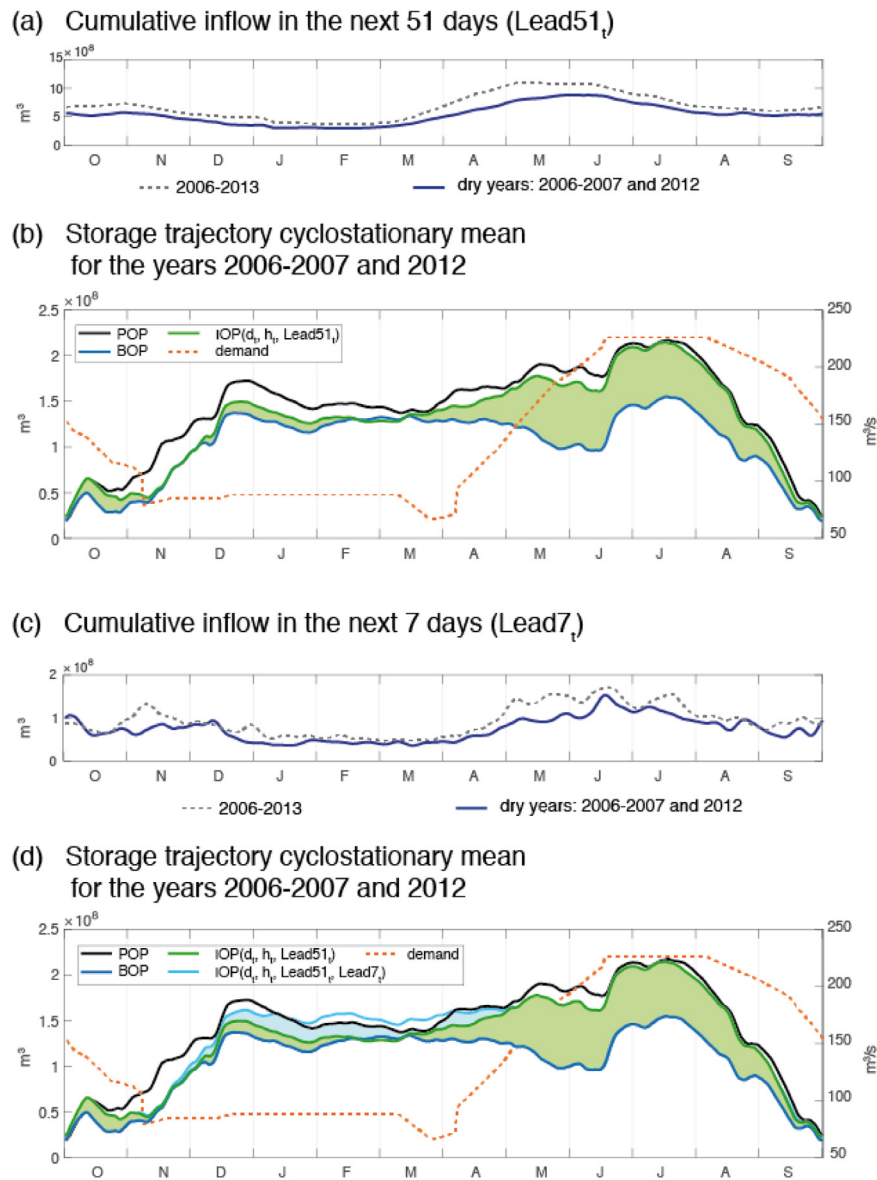


Fig. 7. Analysis of the Lake Como storage trajectory in dry years under different Improved Operating Policies (see Fig. 6) conditioned on inflow forecasts on different lead times (i.e., cumulative inflow in the next 51 and 7 days). As a reference, the trajectory of the target POP and the selected BOP are also illustrated.

lake regulation objectives: insights on the total volume entering the system over some weeks are highly beneficial for water supply, but do not inform on the timing of the flood peak. This latter is better captured by the short lead time information as the time of concentration of the catchment is less than 24 hours, while the time needed to drawdown the lake level and to buffer the flood peak is around 3 days.

To quantify the Expected Value of Sample Information of the two selected forecasts, we incrementally include them as arguments of the Improved Operating Policies. The comparison of the Basic Operating Policy performance with these new IOPs is reported in Fig. 6(a). The set of IOPs informed with the long lead time forecast $Lead51_t$ (green circles in the figure) is largely better than the BOP set (blue circles) and dominates also the Improved Operating Policies informed by the selected observations of the current system conditions (i.e., SWE estimate and hydropower storage, represented by the dashed purple and red Pareto fronts in Fig. 6(a)). This improvement is confirmed by the values of the three metrics (Fig. 6(b)), which improve from 40% to 58%. When the short lead time information $Lead7_t$ is also considered, the cor-

responding IOPs (cyan circles in Fig. 6(a)) further improve, even though the marginal improvement in this case is lower than the one provided by the long lead time forecasts. This difference in terms of EVSI is verified by the values of the metrics reported in Fig. 6(b), which show a marginal improvement of 10% to 15% when moving from $IOP(d_t, h_t, Lead51_t)$ to $IOP(d_t, h_t, Lead51_t, Lead7_t)$.

Finally, it is interesting to analyze the trajectories of the lake storage under these Improved Operating Policies (Fig. 7), because this provides a better understanding of the role of the selected inflow forecasts in informing the system operations. Again, we focus on the drier years in the considered horizon (i.e., 2006, 2007, and 2012), and, for each Pareto front, we consider the closest solution to the target POP (black circled solutions in Fig. 6(a)). As expected, the long lead time forecast (Fig. 7(a)) gives valuable information for hedging: under this Improved Operating Policy, the lake storage is generally higher than under the Basic Operating Policy throughout the entire year (see the green area in Fig. 7(b)), thus saving water towards a dry forecasted period, and is almost reproducing the target Perfect Operating Policy during the peak irrigation season (see the dashed red line in the figure). On the other hand, the contri-

bution of the short lead time forecast (Fig. 7(c)) is mainly beneficial to buffering flooding events and, in particular, the associated $IOP(d_t, h_t, Lead51_t, Lead7_t)$ reproduces the target POP during winter and early spring (see the cyan area in Fig. 7(d)), when, historically, floods are more likely to occur. Actually, the contribution of short lead time to flood protection over the entire horizon is more relevant than what is directly visible from this figure because the reported trajectory is averaged over dry years only.

5. Conclusions

In the present paper, we explored to which extent different hydrological exogenous variables are valuable in complementing basic information systems for improving the operations of multi-purpose water reservoirs and regulated lakes with short- and long-term targets. We considered the regulated Lake Como (Italy) as a study site. We used the Information Selection and Assessment (ISA) framework (Giuliani et al., 2015) to select the most valuable information to improve the lake operations with respect to both short-term (i.e., flood control) and long-term (i.e., water supply) objectives.

Results show that the available observations from the existing monitoring system in the forms of accumulated snow and hydropower reservoir storage successfully surrogate part of the true water system state and contribute to improving the system performance and to mitigating the conflict between the two considered operational objectives. Among the variables considered, Snow Water Equivalent is the most informative one, allowing for almost 10% improvement with respect to a basic information set, composed of day of the year and lake level only. The improvements are most notable when considering the water supply objective. In fact, SWE informs the operating policies about the amount of water stored at high altitudes, which will be available during spring and summer, and allows for hedging. Numerical results demonstrate that using this information, which is available from December to June, allows storing more water in the lake (higher lake levels) to be used few months later in early summer, but not enough to fully cover the demand in late summer and fall period.

Hedging can be further extended when including also the water stored in hydropower reservoirs in the information set. In this case, the lake tends to store more water also in the remaining months (from August to December), because the information is available on the whole year and not only on a seasonal basis, as for SWE. The higher lake levels in autumn increase the risk of flood events, though. The improvements in the system performance are in fact skewed in favor of water supply at the cost of flood control. Overall, the contribution of this information is smaller because the hydropower reservoirs are controlled and the timing of the releases is uncertain, because highly dependent on the electricity market conditions.

Even when considering both SWE and hydropower reservoir storage, there is still a large distance to the performance achievable with a perfectly informed operating policy. This is probably due to the inability of the considered state variables to precisely anticipate the amount and timing of incoming water. In fact, when considering perfect forecasts from one to two months in advance, the performance is largely improved with respect to both the objectives. This information allows increasing the lake storage in late spring and summer (from May to September, approximately), when the water demand is the highest and the largest part of the snow pack is already melted, and thus SWE cannot act as a skillful streamflow forecasts.

Overall, our results suggest that Lake Como operations would benefit from the use of hydro-meteorological data. Although limited, the demonstrated improvement is already attainable with little or no effort from the operator as opposed to reliable long-term

streamflow forecasts. At the time of writing, ARPA is able to produce accurate streamflow forecasts up to only 10 days. Extending this lead-time to the identified monthly/seasonal scale would require considerable research effort. As a future strategy, the water authority operating Lake Como may consider improving the existing monitoring system to collect observational data better capturing the snow dynamics. For example via SAR images (e.g., Denaro et al., 2015) or new satellite products (e.g., Malenovsky et al., 2012). Yet, to attain significantly larger improvements, the water authority should invest in obtaining and using long-term forecasts (e.g., Kim et al., 2012; Baehr et al., 2015), which allow extending the lead-time of SWE information and implementing effective hedging policies.

A further analysis could involve the use of stochastic ensembles of hydro-meteorological variables to assess the robustness of results out of the observed realizations domain. However, such application would entail the synthetic generation of coherent spatially distributed (SWE) and anthropogenic variables (the Alpine storages dynamics) which would require additional investigation and relevant modeling effort.

Acknowledgment

This work was partially supported by the IMPREX project funded by the European Commission under the Horizon 2020 framework programme (grant n. 641811) and the SOWATCH project funded by Fondazione Cariplo (grant number 2015-0220). Daniela Anghileri was supported by the Swiss Competence Center for Energy Research – Supply of Electricity (SCCER-SoE). The authors would like to thank Agenzia Regionale per la Protezione dell'Ambiente, especially Dario Bellingeri and Enrico Zini, and Luigi Bertoli from Consorzio dell'Adda for providing the data used in this study.

References

- Amigoni, F., Castelletti, A., Gazzotti, P., Giuliani, M., Mason, E., 2016. Water resources systems operations via multiagent negotiation. In: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. International Foundation for Autonomous Agents and Multiagent Systems, pp. 1379–1380.
- Anghileri, D., Castelletti, A., Pianosi, F., Soncini-Sessa, R., Weber, E., 2012. Optimizing watershed management by coordinated operation of storing facilities. *J. Water Resour. Plann. Manage.* 139 (5), 492–500.
- Anghileri, D., Pianosi, F., Soncini-Sessa, R., 2011. A framework for the quantitative assessment of climate change impacts on water-related activities at the basin scale. *Hydrol. Earth Syst. Sci.* 15 (6), 2025–2038.
- Anghileri, D., Voisin, N., Castelletti, A., Pianosi, F., Nijssen, B., Lettenmaier, D., 2016. Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments. *Water Resour. Res.*
- Baehr, J., Fröhlich, K., Botzet, M., Domeisen, D.I., Kornblueh, L., Notz, D., Piontek, R., Pohlmann, H., Tietsche, S., Müller, W.A., 2015. The prediction of surface temperature in the new seasonal prediction system based on the mpi-esm coupled climate model. *Clim. Dyn.* 44 (9–10), 2723–2735.
- Bellingeri, D., Zini, E., Serra, R., Peretti, G., 2006. Stima della disponibilità idrica sotto forma di neve nei bacini idrografici alpini lombardi mediante integrazione di dati satellitari modis e misure nivometriche a terra. ASITA Conference Acts.
- Bellman, R., 1957. *Dynamic Programming*. Princeton University Press, Princeton.
- Bertsekas, D.P., 1976. *Dynamic Programming and Stochastic Control*. Academic Press, New York.
- Bertsekas, D.P., Tsitsiklis, J.N., 1995. Neuro-dynamic programming: an overview. In: *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, 1. IEEE, pp. 560–564.
- Block, P., Rajagopalan, B., 2007. Interannual variability and ensemble forecast of upper blue Nile basin kiremt season precipitation. *J. Hydrometeorol.* 8 (3), 327–343.
- Castelletti, A., Galelli, S., Restelli, M., Soncini-Sessa, R., 2010. Tree-based reinforcement learning for optimal water reservoir operation. *Water Resour. Res.* 46 (9), W09507. <http://dx.doi.org/10.1029/2009WR008898>.
- Castelletti, A., Pianosi, F., Soncini-Sessa, R., 2008a. Receding horizon control for water resources management. *Appl. Math. Comput.* 204 (2), 621–631.
- Castelletti, A., Pianosi, F., Soncini-Sessa, R., 2008b. Water reservoir control under economic, social and environmental constraints. *Automatica* 44 (6), 1595–1607.

- Chang, L.-C., Chang, F.-J., 2001. Intelligent control for modelling of real-time reservoir operation. *Hydrol. Process.* 15 (9), 1621–1634.
- Cherubini, T., Ghelli, A., Lalaurette, F., 2002. Verification of precipitation forecasts over the alpine region using a high-density observing network. *Weather Forecasting* 17 (2), 238–249.
- Christensen, N.S., Wood, A.W., Voisin, N., Lettenmaier, D.P., Palmer, R.N., 2004. The effects of climate change on the hydrology and water resources of the colorado river basin. *Clim. Change* 62 (1–3), 337–363.
- Culley, S., Noble, S., Yates, A., Timbs, M., Westra, S., Maier, H., Giuliani, M., Castelletti, A., 2016. A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate. *Water Resour. Res.* 52. <http://dx.doi.org/10.1002/2015WR018253>.
- Denaro, S., Del Gobbo, U., Castelletti, A., Tebaldini, S., Guarnieri, A.M., 2015. Informing water management by direct use of satellite retrieved snow information in snow-rainfall dominated watersheds. In: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, pp. 1531–1534.
- Faber, B., Stedinger, J., 2001. Reservoir optimization using sampling sdp with ensemble streamflow prediction (esp) forecasts. *J. Hydrol.* 249, 113–133. [http://dx.doi.org/10.1016/S0022-1694\(01\)00419-X](http://dx.doi.org/10.1016/S0022-1694(01)00419-X).
- Formentin, S., Heusden, K., Karimi, A., 2012. A comparison of model-based and data-driven controller tuning. *Int. J. Adapt. Control Signal Process.* 28, 882–897.
- Forzieri, G., Feyen, L., Rojas, R., Floerke, M., Wimmer, F., Bianchi, A., 2014. Ensemble projections of future streamflow droughts in Europe. *Hydrol. Earth Syst. Sci.* 18 (1), 85–108.
- Fowler, H., Kilsby, C., O'Connell, P., 2003. Modeling the impacts of climatic change and variability on the reliability, resilience, and vulnerability of a water resource system. *Water Resour. Res.* 39 (8).
- Gal, S., 1979. Optimal management of a multireservoir water supply system. *Water Resour. Res.* 15 (4), 737–749.
- Galelli, S., Castelletti, A., 2013a. Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling. *Hydrol. Earth Syst. Sci.* 17, 2669–2684.
- Galelli, S., Castelletti, A., 2013b. Tree-based iterative input variable selection for hydrological modeling. *Water Resour. Res.* 49 (7), 4295–4310.
- Galelli, S., Humphrey, G.B., Maier, H.R., Castelletti, A., Dandy, G.C., Gibbs, M.S., 2014. An evaluation framework for input variable selection algorithms for environmental data-driven models. *Environ. Modell. Software* 62, 33–51. <http://dx.doi.org/10.1016/j.envsoft.2014.08.015>.
- Galelli, S., Soncini-Sessa, R., 2010. Combining metamodeling and stochastic dynamic programming for the design of reservoir release policies. *Environ. Modell. Software* 25 (2), 209–222. <http://dx.doi.org/10.1016/j.envsoft.2009.08.001>.
- García-Herrera, R., Díaz, J., Trigo, R., Luterbacher, J., Fischer, E., 2010. A review of the european summer heat wave of 2003. *Crit. Rev. Environ. Sci. Technol.* 40 (4), 267–306.
- Gass, S., Saaty, T., 1955. Parametric objective function - Part II. *Oper. Res.* 3, 316–319.
- Gelati, E., Madsen, H., Rosbjerg, D., 2011. Stochastic reservoir optimization using el niño information: case study of daule peripa, ecuador. *Hydrol. Res.* 42 (5), 413–431. <http://dx.doi.org/10.2166/nh.2011.009>. arXiv: <http://hr.iwaponline.com/content/42/5/413.full.pdf>.
- Georgakakos, A., Yao, H., Kistenmacher, M., Georgakakos, K., Graham, N., Cheng, F.-Y., Spencer, C., Shamir, E., 2012. Value of adaptive water resources management in northern california under climatic variability and change: reservoir management. *J. Hydrol.* 412–413, 34–46. Hydrology Conference 2010. <http://dx.doi.org/10.1016/j.jhydrol.2011.04.038>.
- Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. *Mach. Learn.* 63 (1), 3–42. <http://dx.doi.org/10.1007/s10994-006-6226-1>.
- Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., Reed, P., 2016. Curses, tradeoffs, and scalable management: advancing evolutionary multi-objective direct policy search to improve water reservoir operations. *J. Water Resour. Plann. Manage.* 142 (2). [http://dx.doi.org/10.1061/\(ASCE\)WR.1943-5452.0000570](http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000570).
- Giuliani, M., Herman, J., Castelletti, A., Reed, P., 2014a. Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management. *Water Resour. Res.* 50, 3355–3377. <http://dx.doi.org/10.1002/2013WR014700>. doi: 10.1002/2013WR014700.
- Giuliani, M., Mason, E., Castelletti, A., Pianosi, F., Soncini-Sessa, R., 2014b. Universal approximators for direct policy search in multi-purpose water reservoir management: A comparative analysis. In: Proceedings of the 19th IFAC World Congress. Cape Town (South Africa).
- Giuliani, M., Pianosi, F., Castelletti, A., 2015. Making the most of data: an information selection and assessment framework to improve water systems operations. *Water Resour. Res.* 51 (11), 9073–9093. <http://dx.doi.org/10.1002/2015WR017044>.
- Guariso, G., Rinaldi, S., Soncini-Sessa, R., 1986. The management of lake como: a multiobjective analysis. *Water Resour. Res.* 22 (2), 109–120. <http://dx.doi.org/10.1029/WR022i002p0109>.
- Guariso, G., Rinaldi, S., Zielinski, P., 1984. The value of information in reservoir management. *Appl. Math. Comput.* 15 (2), 165–184. [http://dx.doi.org/10.1016/0096-3003\(84\)90015-8](http://dx.doi.org/10.1016/0096-3003(84)90015-8).
- Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. *J. Mach. Learn. Res.* 3, 1157–1182.
- Hadka, D., Reed, P., 2013. Borg: an auto-adaptive many-objective evolutionary computing framework. *Evol. Comput.* 21 (2), 231–259.
- Hamlet, A.F., Lettenmaier, D.P., 1999. Columbia river streamflow forecasting based on enso and pdo climate signals. *J. Water Resour. Plann. Manage.* 125 (6), 333–341.
- Hashimoto, T., Stedinger, J.R., Loucks, D.P., 1982. Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation. *Water Resour. Res.* 18 (1), 14–20.
- Hejazi, M.I., Cai, X., 2011. Building more realistic reservoir optimization models using data mining—a case study of shelbyville reservoir. *Adv. Water Resour.* 34 (6), 701–717.
- Hejazi, M.I., Cai, X., Ruddell, B.L., 2008. The role of hydrologic information in reservoir operation learning from historical releases. *Adv. Water Resour.* 31 (12), 1636–1650. <http://dx.doi.org/10.1016/j.advwatres.2008.07.013>.
- Hou, Z.-S., Wang, Z., 2013. From model-based control to data-driven control: survey, classification and perspective. *Inf. Sci.* 235, 3–35. Data-based Control, Decision, Scheduling and Fault Diagnostics. <http://dx.doi.org/10.1016/j.ins.2012.07.014>.
- Kalra, A., Ahmad, S., Nayak, A., 2013. Increasing streamflow forecast lead time for snowmelt-driven catchment based on large-scale climate patterns. *Adv. Water Resour.* 53, 150–162. <http://dx.doi.org/10.1016/j.advwatres.2012.11.003>.
- Karamouz, M., Houck, M. H., 1987. Comparison of stochastic and deterministic dynamic programming for reservoir operating rule generation.
- Kelman, J., Stedinger, J., Cooper, L., Hsu, E., Yuan, S., 1990. Sampling stochastic dynamic programming applied to reservoir operation. *Water Resour. Res.* 26 (3), 447–454.
- Kim, H.-M., Webster, P.J., Curry, J.A., 2012. Seasonal prediction skill of ecwf system 4 and ncep cfsv2 retrospective forecast for the northern hemisphere winter. *Clim. Dyn.* 39 (12), 2957–2973.
- Kim, Y., Palmer, R., 1997. Value of seasonal flow forecasts in bayesian stochastic programming. *J. Water Resour. Plann. Manage.* 123 (6), 327–335. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(1997\)123:6\(327\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(1997)123:6(327)).
- Kim, Y.-O., Eum, H.-I., Lee, E.-G., Ko, I.H., 2007. Optimizing operational policies of a korean multireservoir system using sampling stochastic dynamic programming with ensemble streamflow prediction. *J. Water Resour. Plann. Manage.* 133 (1), 4–14.
- Koster, R.D., Mahanama, S.P., Livneh, B., Lettenmaier, D.P., Reichle, R.H., 2010. Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nat. Geosci.* 3 (9), 613–616.
- Lehner, B., Döll, P., Alcamo, J., Henrichs, T., Kaspar, F., 2006. Estimating the impact of global change on flood and drought risks in europe: a continental, integrated analysis. *Clim. Change* 75 (3), 273–299.
- Liu, P., Guo, S., Xiong, L., Li, W., Zhang, H., 2006. Deriving reservoir refill operating rules by using the proposed dpns model. *Water Resour. Manage.* 20 (3), 337–357.
- Loucks, D., Stedinger, J., Haith, D., 1981. Water resource systems planning and analysis.
- Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., Reichle, R., 2012. Soil moisture, snow, and seasonal streamflow forecasts in the united states. *J. Hydrometeorol.* 13 (1), 189–203.
- Maidment, D.R., Chow, V.T., 1981. Stochastic state variable dynamic programming for reservoir systems analysis. *Water Resour. Res.* 17 (6), 1578–1584.
- Maier, H., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L., Cunha, M., Dandy, G., Gibbs, M., Keedwell, E., Marchi, A., Ostfeld, A., Savić, D., Solomatine, D., Vrugt, J., Zecchin, A., Minsker, B., Barbour, E., Kuczera, G., Pasha, F., Castelletti, A., Giuliani, M., Reed, P., 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. *Environ. Modell. Software* 62 (0), 271–299. <http://dx.doi.org/10.1016/j.envsoft.2014.09.013>.
- Malenkovský, Z., Rott, H., Cihlar, J., Schaeppman, M.E., García-Santos, G., Fernandes, R., Berger, M., 2012. Sentinels for science: potential of sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sens. Environ.* 120, 91–101.
- Piccardi, C., Soncini-Sessa, R., 1991. Stochastic dynamic programming for reservoir optimal control: dense discretization and inflow correlation assumption made possible by parallel computing. *Water Resour. Res.* 27 (5), 729–741.
- Raso, L., Schwanenberg, D., van de Giesen, N., van Overloop, P., 2014. Short-term optimal operation of water systems using ensemble forecasts. *Adv. Water Resour.* 71, 200–208. <http://dx.doi.org/10.1016/j.advwatres.2014.06.009>.
- Sharma, A., Luk, K., Cordero, I., Lall, U., 2000. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: part 2 predictor identification of quarterly rainfall using ocean-atmosphere information. *J. Hydrol.* 239 (1), 240–248.
- Sreekanth, J., Datta, B., Mohapatra, P.K., 2012. Optimal short-term reservoir operation with integrated long-term goals. *Water Resour. Manage.* 26 (10). <http://dx.doi.org/10.1007/s11269-012-0051-z>.
- Stedinger, J., Sule, B., Loucks, D., 1984. Stochastic dynamic programming models for reservoir operation optimization. *Water Resour. Res.* 20 (11), 1499–1505.
- Tejada-Guibert, J.A., Johnson, S.A., Stedinger, J.R., 1995. The value of hydrologic information in stochastic dynamic programming models of a multireservoir system. *Water Resour. Res.* 31 (10), 2571–2579. <http://dx.doi.org/10.1029/95WR02172>.
- Todini, E., 2014. The role of predictive uncertainty in the operational management of reservoirs. In: Proceedings of the ICWRS2014 Evolving Water Resources Systems: Understanding, Predicting and Managing Water-Society Interactions, Bologna, Italy, pp. 4–6.
- Turner, S.W.D., Galelli, S., 2016. Regime-shifting streamflow processes: implications for water supply reservoir operations. *Water Resour. Res.* 52 (5), 3984–4002. <http://dx.doi.org/10.1002/2015WR017913>.
- Valeriano, S., Oliver, C., Koike, T., Yang, K., Graf, T., Li, X., Wang, L., Han, X., 2010. Decision support for dam release during floods using a distributed biosphere hydrological model driven by quantitative precipitation forecasts. *Water Resour. Res.* 46 (10).

- Xu, B., Zhong, P.-A., Stanko, Z., Zhao, Y., Yeh, W.W.-G., 2015. A multiobjective short-term optimal operation model for a cascade system of reservoirs considering the impact on long-term energy production. *Water Resour. Res.* 51 (5), 3353–3369.
- Zatarain-Salazar, J., Reed, P., Herman, J., Giuliani, M., Castelletti, A., 2016. A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. *Adv. Water Resour.* 92, 172–185. <http://dx.doi.org/10.1016/j.advwatres.2016.04.006>.
- Zhao, T., Cai, X., Yang, D., 2011. Effect of streamflow forecast uncertainty on real-time reservoir operation. *Adv. Water Resour.* 34 (4), 495–504.
- 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 Resour. Res.* 48 (1).
- Zhao, T., Zhao, J., 2014. Joint and respective effects of long-and short-term forecast uncertainties on reservoir operations. *J. Hydrol.* 517, 83–94.
- Zimmerman, B.G., Vimont, D.J., Block, P.J., 2016. Utilizing the state of enso as a means for season-ahead predictor selection. *Water Resour. Res.* 52 (5), 3761–3774. <http://dx.doi.org/10.1002/2015WR017644>.
- Zitzler, E., Deb, K., Thiele, L., 2000. Comparison of multiobjective evolutionary algorithms: empirical results. *Evol. Comput.* 8 (2), 173–195. <http://dx.doi.org/10.1162/106365600568202>.
- Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C., da Fonseca, V., 2003. Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Trans. Evol. Comput.* 7 (2), 117–132.