Scalable multi-objective control for large scale water resources systems under uncertainty

Matteo Giuliani[1], Julianne D. Quinn[2], Jonathan D. Herman[3], Andrea Castelletti [1,4], Patrick M. Reed[2]

[1]Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy
[2]Cornell University, 211 Hollister Hall, Ithaca, NY.
[3]University of California, Davis, 3138 Ghausi Hall, Davis, CA.
[4]Institute of Environmental Engineering, ETH, Zurich, Switzerland.

Advances in modeling and control have always played an important role in supporting water resources systems planning and management. Changes in climate and society are now introducing additional challenges for controlling these systems, motivating the emergence of complex, integrated simulation models to explore key causal relationships and dependencies related to uncontrolled sources of variability. In this paper, we contribute a massively parallel implementation of the evolutionary multi-objective direct policy search method for controlling large scale water resources systems under uncertainty. The method combines direct policy search with nonlinear approximating networks and a hierarchical parallelization of the Borg Multi-Objective Evolutionary Algorithm. This computational framework successfully identifies control policies that address both the presence of multidimensional tradeoffs and of severe uncertainties in the system dynamics and policy performance. We demonstrate the approach on a challenging real-world application, represented by the optimal control of a network of four multipurpose water reservoirs in the Red River basin in Northern Vietnam, under observed and synthetically generated hydrologic conditions. Results show that the reliability of the computational framework in finding near-optimal solutions increases with the number of islands in the adopted hierarchical parallelization scheme. This setting reduces the vulnerabilities of the designed solutions to the system's uncertainty and improves the discovery of robust control policies addressing key system performance tradeoffs.

***Index Terms*—multi-objective control, water resources systems.**

I. Introduction

The optimal operation of water systems is a wide and challenging application domain for optimal control methodologies (for a review, see [1] and references therein), which has the potential for supporting regional growth and development by increasing water availability for different economic sectors, including hydropower production [2], drinking water supply [3], irrigation supply [4], and ecological conservation [5]. Yet, uncertain environmental, economic, and social conditions can diminish these benefits and carry severe risks of economic and environmental failure [6].

Traditionally, water control problems are formulated as Markov decision processes (MDP, see [7]), where the reservoir's storage is the system's state and the control decision determines the volume of water to release at discrete time instants (for a review, see [8] and references therein). The system then moves through a generally non-linear transition into a new state, influenced by the control of the dam and the stochastic disturbances (e.g., inflows), producing an immediate reward (or cost) after the transition (e.g., hydropower production). The flexibility of this formulation is particularly suitable for modeling most water resources systems as it removes the simplifying assumptions of linear models, quadratic costs, and white Gaussian disturbances required by the largely adopted Linear-Quadratic-Gaussian framework [9]. In principle, MDP can be solved via Dynamic Programming (DP, [10]) under relatively mild modeling assumptions, namely discrete domains of state, control, and disturbance variables, time-separability of objective functions and constraints, disturbance processes uncorrelated in time. In practice, beside being limited by the well known curse of dimensionality [10], the application of DP in large-scale control schemes is constrained by two additional drawbacks: (i) the curse of modeling [11] as no exogenous information (i.e., variables that are observed, but do not depend on the control and are not endogenous to the problem formulation) can be used for conditioning the control policy, thus limiting the controller's ability to handle the severe natural variability of the disturbances; (ii) the curse of multiple objectives [12], as the generation of a Pareto front to explore the tradeoffs between multiple conflicting objectives scales factorially with the growth in the number of the objectives [13]. DP is inherently single-objective and would require repeated scalarized single-objective optimizations for every Pareto optimal point.

Capitalizing on the recent improvements in terms of scientific knowledge of the natural processes, efficiency of control systems technologies, and availability of computational resources [14], in this paper we contribute a scalable implementation of the evolutionary multi-objective direct policy search (EMODPS) method [15], where we combine direct policy search with nonlinear approximating networks and hierarchically parallelized multi-objective evolutionary algorithms (MOEAs). This computational framework enhances the scalability of traditional DP-based methods by enlarging the information used for conditioning operational decisions under uncertain conditions and by increasing the number of objectives formulated in the control policy design problem. In fact, DPS allows the direct use of exogenous information through a partially data-driven controller tuning approach [16], which reduces the uncertainty about the system dynamics due to the variability of natural processes that cannot be accurately modeled and would therefore produce detrimental effects on

Corresponding author: M. Giuliani (email: matteo.giuliani@polimi.it).

the performance of model-based controllers [17]. However, this increases the complexity of the control policy as the number of parameters necessary to obtain a good approximation for the unknown optimal control policy grows with the increasing dimension of the controller's argument [18]. We define the parameterized control policies as nonlinear approximating networks to ensure the possibility of approximating the unknown optimal solution to any desired degree of accuracy [15]. Optimizing the parameters of these nonlinear approximating networks requires searching high dimensional spaces that map to noisy and multimodal objective function values. This search can be efficiently performed by using MOEAs. Although evolutionary strategies do not guarantee the optimality of the solutions, they represent a promising alternative to gradient-based methods because evolving a set of candidate solutions based on their ranking has been demonstrated to better handle performance uncertainties than methods relying on the estimation of absolute performance or performance gradient [19]. In particular, we use the Multi-Master Borg MOEA [20], a hierarchically parallelized self-adaptive version of the Borg MOEA recently proposed for static optimization problem and applied here for the first time to an optimal control problem. The possibility of handling many-objective control problems, where the number of objective functions might be in the order of three or more [21], is crucial for exploring multidimensional tradeoffs and overcoming decision biases produced by narrow or restrictive definitions of optimality that strongly limit the discovery of decision relevant alternatives [22], [23]. Finally, beyond simply being faster in terms of wall clock, the Multi-Master Borg parallelization transforms the search dynamics of the algorithm through communication across multiple parallel optimizations, increasing the algorithmic reliability for finding favorable solution sets [24].

In summary, this study has two key contributions. First, we introduce a partial model-free control scheme based on parameterization of the control laws via nonlinear networks. The control policies can be directly conditioned on non-modeled information about the reservoir inflows in order to capture variability in the hydrologic regime and attain better policy performance under uncertain conditions. Second, we couple this control scheme with a a hierarchically parallelized self-adaptive version of the Borg MOEA to optimize the large number of policy parameters and explore the tradeoffs for high dimensional multi-objective control problems through new algorithmic feedbacks. The parallelization scheme uses multiple, cooperating master-worker algorithm instances to avoid high variance (and possibly failure) in algorithmic performance, which would skew stakeholders' perceptions of key system tradeoffs.

II. Problem formulation

Water reservoir control problems require making sequential decisions $\mathbf{u}_t$ about the volume of water to be released at discrete time instants on the basis of current system conditions described by the state vector $\mathbf{x_t}$ (e.g., reservoir storage). The state of the system is then altered according to a stochastic transition function, affected by the vector of stochastic external drivers $\boldsymbol{\varepsilon}_{t+1}$ (e.g., reservoir inflows). Such system can be modeled as a discrete-time, periodic, non-linear, stochastic Markov Decision Process [1] as follows

$$\mathbf{x}_{t+1} = f_t(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\varepsilon}_{t+1}) \quad t = 0, 1, \ldots, h-1 \tag{1}$$

where the vectors $\mathbf{x}_t \in \mathbb{R}^{n_x}$, $\mathbf{u}_t \in \mathcal{U}_t(\mathbf{x}_t) \subseteq \mathbb{R}^{n_u}$, and $\boldsymbol{\varepsilon}_t \in \mathbb{R}^{n_\varepsilon}$ have dimensions n_x, n_u, and n_ε, respectively. The state is assumed to be observable and the stochastic disturbance is described by a probability density function (i.e., $\boldsymbol{\varepsilon}_{t+1} \sim \phi$ for $t = 0, 1, \ldots, h-1$). In the adopted notation, the time subscript of a variable indicates the instant when its value is deterministically known. The state $\mathbf{x}_t$ is observed at time t, whereas the disturbances vector has subscript $t+1$, denoting the realization of the stochastic process in the time interval $[t, t+1)$.

For each of the M stakeholders representing the different water-related interests in the system (e.g., hydropower production, water supply, environmental protection), we formulate a specific objective function (herein assumed to be a cost) evaluated over the time horizon $[0, h]$ as:

$$J^m = \Psi_{\boldsymbol{\varepsilon}_1,\ldots,\boldsymbol{\varepsilon}_h}\left[\Phi_{0,\ldots,h}\left(g_1^m(\mathbf{x}_0, \mathbf{u}_0, \boldsymbol{\varepsilon}_1), \ldots, g_h^m(\mathbf{x}_h)\right)\right] \tag{2}$$

where $g_{t+1}^m(\cdot)$ for $t = 0, 1, \ldots, h-1$ is the m-th immediate cost function (with $m = 1, \ldots, M$) associated with the time transition from t to $t+1$, $g_h^m(\mathbf{x}_h)$ is a penalty function over the final state, Φ is an operator for time aggregation over h (e.g., the average $\Phi = \sum / (h+1)$), and Ψ is a statistic used to filter the noise generated by the disturbances (e.g., expected value $\Psi = E$). The optimal control problem is then formulated as finding the optimal control policy p^* defined as the periodic sequence of control laws $\mathbf{u}_t = \mu_t^*(\mathbf{x}_t)$, with period T, i.e. $p^* \triangleq [\mu_0^*(\mathbf{x}_t), \ldots, \mu_{T-1}^*(\mathbf{x}_{T-1})]$, that minimizes the M-dimensional objective function vector $\mathbf{J}$, i.e.

$$\begin{aligned} p^* = \arg\min_p \mathbf{J}(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h) = \\ \left[J^1(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h), \ldots, J^M(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h)\right] \end{aligned} \tag{3}$$

subject to the dynamics of the system (eq. (1)) and given the initial state $\mathbf{x}_0$. The multi-objective formulation of Problem 3 does not yield a single solution that minimizes all the M objectives as it does not generally exist. Instead, it determines a set of Pareto optimal solutions $\mathcal{P}^*$, which maps onto the so-called Pareto front $\mathcal{F}^* = \{\mathbf{J}(p^*, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h) | p^* \in \mathcal{P}^*\}$. The set $\mathcal{P}^*$ is generally infinite and only a finite subset of $\mathcal{P}^*$ can actually be computed.

Definition 1: Policy p dominates policy p', denoted by $p \prec p'$, if: $\forall i \in \{1, \ldots, M\}, J^i(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h) \leq J^i(p', \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h) \wedge \exists i \in \{1, \ldots, M\}, J^i(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h) < J^i(p', \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h)$.

Definition 2: If there is no policy p' such that $p' \prec p$, the policy p is Pareto optimal.

The traditional approach to solve a multi-objective problem via DP-based methods is to reformulate it as a series of single-objective problems by combining the M objectives with some scalarization function $\Gamma : \mathbb{R}^M \to \mathbb{R}$, such as a convex combination of the objectives using weights $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_M] \in \Lambda^{M-1}$, where Λ^{M-1} is the unit $(M-1)$-dimensional simplex (so that $\sum_{i=1}^{M} \lambda_i = 1$ and $\lambda_i \geq 0 \quad \forall i$). Problem 3 is hence reformulated as

$$p^* = \arg\min_p \mathcal{J}(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h) = \Gamma\left([J^1(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h), \ldots, J^M(p, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h)]\right) \tag{4}$$

subject to the dynamics of the system (eq. (1)) and given the initial state $\mathbf{x}_0$. Problem 4 can be solved by Stochastic Dynamic Programming (SDP, [10]) by estimating the expected long-term cost of a policy for each state $\mathbf{x}_t$ at time t by means of the Bellman function:

$$H_t(\mathbf{x}_t) = \min_{\mathbf{u}_t} E_{\boldsymbol{\varepsilon}_{t+1}} \left[G_{t+1}(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\varepsilon}_{t+1}) + H_{t+1}(\mathbf{x}_{t+1})\right] \tag{5}$$

where $H_t(\cdot)$ is the optimal cost-to-go function defined over a discrete grid of states for the scalarized objective $\mathcal{J}$ and $G_{t+1}(\cdot)$ is the corresponding scalarized immediate cost function. The optimal control policy p^* is then derived as the sequence of control laws that minimizes eq. (5).

III. Evolutionary Multi-Objective Direct Policy Search

Evolutionary Multi-Objective Direct Policy Search (EMODPS) is an approximate dynamic programming method that replaces the traditional SDP policy design approach, based on the search of the value function in the objective space, with a functional optimization that directly searches the optimal control policy within an infinite-dimensional space of functions [9]. EMODPS is based on the parameterization of the control policy p_θ and the simulation-based exploration of the parameters' space Θ to find a Pareto approximate set of parameterized control policies $\mathcal{P}_\theta^*$ that minimizes the M-dimensional objective function vector, i.e. $\mathcal{P}_\theta^* = \arg\min_{\mathcal{P}_\theta} \mathbf{J}(\mathcal{P}_\theta, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h)$ (see Fig. 1). The advantage of this approach is transforming the functional SDP problem formulated in eq. (5) into a nonlinear optimization problem, where finding $\mathcal{P}_\theta^*$ is equivalent to finding the corresponding optimal policy parameters θ^*, with $\theta \in \Theta$. Different policy search approaches have been proposed in recent years (for a review see [25] and references therein), whose effectiveness (i.e., ability in discovering high-quality solutions) is strongly dependent on the choice of the parameterization for the control policy [26] as well as on the ability of the optimization algorithm used to search for the optimal control policy parameters [27].

A. Control policy parameterization

Since DPS can, at most, find the best possible control policy within the prescribed family of functions, a flexible nonlinear function depending on a high number of parameters has to be selected to avoid restricting the search for the optimal policy to a subspace of the decision space that does not include the optimal solution. Typically, ad-hoc policy parameterizations are designed for specific problems, using intuition and prior knowledge about the optimal policy [28].

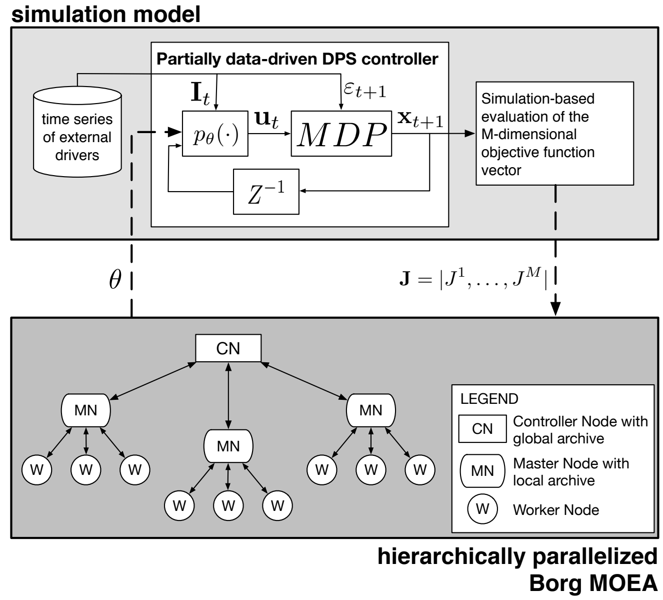

Fig. 1. Schematization of the hierarchically parallelized EMODPS approach.

To ensure the possibility of approximating the unknown optimal solution to any desired degree of accuracy, we use Gaussian radial basis functions (RBFs) to parameterize the control policy as they are capable of representing policies for a large class of MDPs [29], allow conditioning the control policy on high-dimensional policy argument vectors ($\mathcal{I}_t$), which, in addition to the state variables $\mathbf{x}_t$, may include exogenous information $\mathbf{I}_t$ in a data-driven approach [30], and have been demonstrated to be more effective than other nonlinear approximating networks [15]. In this formulation, the k-th control in the n_u-dimensional vector $\mathbf{u}_t$ is defined as:

$$u_t^k = \sum_{i=1}^{A} w_{i,k} \varphi_i(\mathcal{I}_t) \quad \text{where}$$
$$\varphi_i(\mathcal{I}_t) = \exp\left[-\sum_{j=1}^{B} \frac{((\mathcal{I}_t)_j - c_{j,i})^2}{b_{j,i}^2}\right] \tag{6}$$

where A is the number of RBFs, B the number of policy arguments, $\mathbf{c}_i$ and $\mathbf{b}_i$ the B-dimensional center and radius vectors of the i-th RBF, respectively, and $w_{i,k}$ the weight of the i-th RBF for the k-th control variable. The weights are formulated such that $[w_{1,k}, \ldots, w_{A,k}] \in \Lambda^{A-1}$, where Λ^{A-1} is the unit $(A-1)$-dimensional simplex (i.e., $\sum_{i=1}^{A} w_{i,k} = 1$ and $w_{i,k} \geq 0 \quad \forall i, k$). The policy parameters vector θ is therefore defined as $\theta = [(c_1, \ldots, c_B), (b_1, \ldots, b_B), (w_1, \ldots, w_{n_u})]_1^A$. The total number of policy parameters is equal to $A(2B+n_u)$.

B. Optimization of control policy parameters

The general MDP formulation in eq. (1) prevents expressing eq. (2) in explicit form, thus reducing the effectiveness of gradient-based methods for searching the optimal values of the control policy parameters. Gradient-based algorithms would indeed estimate $\nabla_\theta E[J(p_\theta, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h)]$ through the approximation $\nabla_\theta E[J(p_\theta, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h(k))]$ dependent on k considered realizations of the stochastic disturbances, either historical or synthetically generated. Conversely, evolutionary algorithms have been shown to better handle performance uncertainties than methods relying on the estimation of absolute performance or performance gradient [19].

To cope with the high-dimensional space of policy parameters as well as nonlinear, noisy, and multimodal objective functions, in this work we use the Multi-Master Borg MOEA [20], which also allows the exploration of multidimensional tradeoffs in a single run of the algorithm. The Borg MOEA [31] includes epsilon-dominance archiving [32], adaptive population sizing [33], a steady-state algorithm structure [34], and multiple variation operators whose probability of selection adaptively changes throughout the search based on recent success in generating new non-dominated solutions. These features overcomes the limitations of tuning the algorithm parameters to the specific fitness landscape of the problem.

The multi-master extension of the Borg MOEA implements a hierarchical parallelization scheme in which multiple populations of solutions (or "islands" [35]) evolve cooperatively. Traditional master-worker parallelization approaches only distribute the simulation-based evaluation of objectives across cores without fundamentally changing the search dynamics of the original serial algorithm. The linear speedup in terms of computing more function evaluations in a fixed wallclock period for the traditional master-worker parallelizations (i.e., four cores yields approximately 4 times as much search as the serial search in the same amount of time) is bounded by the communication costs and the serial algorithmic operations' processing time at the master node. Although a traditional master-worker parallelization can help overcome the computational barriers for accurately approximating $E[J(p_\theta, \mathbf{x}_0, \boldsymbol{\varepsilon}_1^h(k))]$ by increasing the number of realizations of the stochastic disturbances that can be used in Monte Carlo simulation evaluations, the hierarchical parallelization of the Multi-Master Borg MOEA leverages multiple, interacting copies of the algorithm to improve it's speedup as well as the reliability of attaining high quality Pareto approximations for severely challenging search problems. Each master has a dedicated set of workers for its own population while being able to interact with other master-worker populations on other cores. In addition, while classic island-based models exacerbate the difficulties of tuning both the algorithm's setting (e.g., population size, operator selection) and the islands' interactions (e.g., migration policies), the auto-adaptivity of the Multi-Master Borg MOEA means the algorithm automatically configures its setting for the local conditions encountered during search and its migration mechanism based on the search progress made by each island.

Within this configuration, a crucial role is played by a new node, called the "controller", which is responsible for maintaining a global epsilon-dominance archive, and for providing guidance to master nodes when they need help. Since each master node is running an instance of the master-worker Borg MOEA, it includes all of the mechanisms to detect search stagnation and trigger restarts, which can be supported by the controller through the injection into the local archive of good solutions derived from the global search state. The combination of these features has been found to increase the efficiency and reliability of search compared to serial and master-worker implementations for challenging real-valued optimization problems [20].

IV. Real world application

The proposed massively parallel implementation of EMODPS is tested on a large and complex real-world system: the optimal control of a network of four multipurpose water reservoirs in the Red River basin, Northern Vietnam, under historical as well as synthetically generated hydrologic conditions.

A. The Red River basin

The Red River basin is an international basin covering an area of 169,000 km^2 between Vietnam (51.3%), China (48%), and Laos (0.7%). The three main tributaries of the Red River (Da, Lo, and Thao) rise in the northern part of the basin and join before reaching a large flood plain in the Red River delta region. As part of the Vietnamese national energy and food security strategy, a number of large dams have been constructed on the Red River tributaries (i.e., Hoa Binh and Son La on the Da River, and Tach Ba and Tuyen Quang on the Lo River), for a total storage capacity of 23.5 billion m^3. These dams are operated to satisfy three main competing objectives: hydropower production in the four power plants connected to the dams, which account for a total installed power capacity of 6,366 MW (i.e., 44% of the Vietnamese national capacity); flood control, particularly in the city of Hanoi, where 6.5 million people live; water supply to support a variety of water related activities in the Red River delta, where agriculture accounts for 58% of the total water demand.

The system is modeled as a discrete-time, periodic, nonlinear, stochastic MDP with the following features (Fig. 2b): a continuous state variable vector $\mathbf{x}_t = |x_t^1, x_t^2, x_t^3, x_t^4|$ representing the water volumes stored in the four reservoirs; a continuous control vector $\mathbf{u}_t = |u_t^1, u_t^2, u_t^3, u_t^4|$ representing the release decision for each reservoir; a discrete time, nonlinear state-transition function $\mathbf{x}_{t+1} = f_t(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\varepsilon}_{t+1})$, which is affected by a vector of stochastic disturbances $\boldsymbol{\varepsilon}_{t+1} = |q_{t+1}^1, q_{t+1}^2, q_{t+1}^3, q_{t+1}^4, q_{t+1}^5|$, representing the inflow to each reservoir. The nonlinear dynamics of the system are due to the release function, which determines the actual release from the j-th reservoir r_{t+1}^j as a function of the control u_{t+1}^j, along with the minimum v_t^j and maximum releases V_t^j that can be produced in the time interval $[t, t+1)$ by keeping all the dam's gates completely closed and completely open, respectively. These two functions are computed by integrating, over the same time interval $[t, t+1)$, the continuous-time mass balance equation of the j-th reservoir $ds^j/d\varsigma = q^j(\varsigma) - r^j(\varsigma) - e^j(\varsigma)$ (with $j = 1, \ldots, 4$), where e_t^j represents the evaporation losses. In addition, the routing of the water from the reservoir outlets to Hanoi and the irrigation district in the delta is described by two data-driven artificial neural networks (ANNs). These latter allow the simulation of the water level in Hanoi z_t^{HN} and the water volume in the irrigation canals Υ_t, and are also affected by the tide τ_t. The two ANNs provide a simplified representation of the routing process in the main river and in the nearly 337 rivers and canals of 22 irrigation districts, accounting for 149 drainage culverts, 89 sluice gates, and 160 pumps (for further details, see [36]).

By direct interaction with the stakeholders[1], the objective function vector $\mathbf{J}$ is formulated as follows, where we set the penalty of the final state $g_h^m(\mathbf{x}_h) = 0$ (with $m = 1, 2, 3$):

- Hydropower production (J^{hyd}): the daily average energy production (GWh/day), to be maximized:

$$J^{hyd} = \sum_{j=1}^{3} \left(\frac{1}{h+1} \sum_{t=0}^{h-1} \eta_{t+1}^{j}(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\varepsilon}_{t+1}) \right) \quad (7)$$

  where $\eta_{t+1}^{j}(\cdot)$ is the daily energy produced by the power plant connected to the j-th reservoir, accounting for the optimal hourly operation of the turbines given the net hydraulic head (i.e., the reservoir level minus the tailwater level) and the daily reservoir release.

- Flood control (J^{flood}): the daily average flood damage in Hanoi (-), estimated by a dimensionless nonlinear cost function, defined by direct consultation with stakeholders and local experts, which depends on the water level z_t^{HN} in Hanoi, to be minimized:

$$J^{flood} = \frac{1}{h+1} \sum_{t=0}^{h-1} F(z_{t+1}^{HN}) \quad \text{where } F(z_{t+1}^{HN}) =$$

$$= \begin{cases} 0 \quad \text{if } z_{t+1}^{HN} \leq 6.0 \text{ m} \\ (z_{t+1}^{HN} - 6) \cdot 750000/5.25 \quad \text{if } 6.0 < z_{t+1}^{HN} \leq 11.25 \text{ m} \\ 1.51 \cdot 10^6 (z_{t+1}^{HN})^4 - 7.00 \cdot 10^7 (z_{t+1}^{HN})^3 + \\ +1.22 \cdot 10^9 (z_{t+1}^{HN})^2 + -9.45 \cdot 10^9 (z_{t+1}^{HN}) + \\ +2.74 \cdot 10^{10} \quad \text{otherwise} \end{cases} \quad (8)$$

- Water supply (J^{supply}): the daily average squared water deficit $(\text{m}^3/\text{s})^2$ with respect to the total water demand of the Red River delta, to be minimized:

$$J^{supply} = \frac{1}{h+1} \sum_{t=0}^{h-1} (ANN(\varrho_{t+1}, W_t, \tau_{t-1}, \Upsilon_t))^2 \quad (9)$$

  where the water deficit is estimated by a third ANN dependent on the flow in the Red River delta ($\varrho_{t+1} = r_t^2 + r_t^3 + r_t^4 + q_t^2 + q_t^4$) obtained from the previous day's releases at Hoa Binh, Thac Ba and Tuyen Quang, and natural flows in the Thao and Lo Rivers, the water demand (W_t), the previous day's tide (τ_{t-1}), and the water volume in the irrigation canals (Υ_t). The adopted quadratic formulation aims to penalize severe deficits in a single time step, while allowing for more frequent, small shortages [37].

More details about the integrated model of the system and its formulations can be found in [38]. Note that, beside the nonlinear system dynamics, the Red River system includes nonlinear and nondifferentiable objective functions (see the

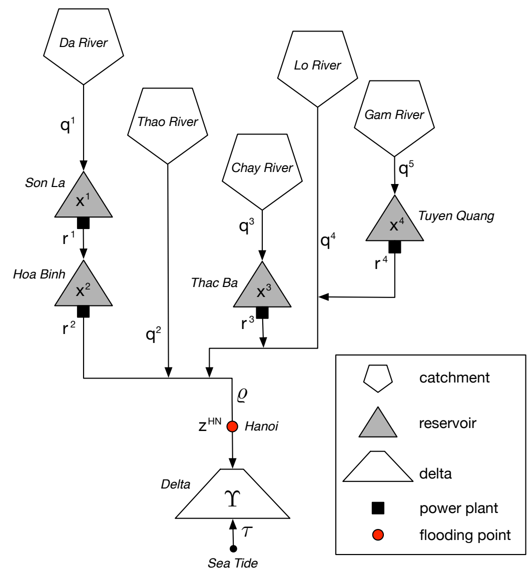

Fig. 2. Schematic representation of the components of Red River basin model.

4th order polynomial used in the flood cost function $F(z_{t+1}^{HN})$) and non-Gaussian disturbances, which are both spatially and temporally correlated.

B. Computational Experiment

The set of Pareto approximate control policies for operating the four reservoirs in the Red River basin is designed via EMODPS by parameterizing the control policies as RBFs and by optimizing the control policy parameters using the Multi-Master Borg MOEA. The RBF policy provides the 4-dimensional control vector $\mathbf{u}_t = |u_t^1, u_t^2, u_t^3, u_t^4|$ as a function of the day of the year t to embed time-variability and cyclostationarity in the control policy, the current reservoirs storages $(x_t^1, x_t^2, x_t^3, x_t^4)$, and the total previous day's inflow $Q_t = \sum_{i=1}^{5} q_t^i$, which is not modeled but employed as exogenous information in a data-driven fashion to better inform the control policy on the natural variability of the unmodeled inflow processes. The total number of RBF policy parameters is equal to 176. This level of complexity prevents benchmarking the proposed approach against a traditional approach based on Dynamic Programming.

In this work, we consider two different EMODPS problem formulations. A *deterministic formulation*:

$$\mathcal{P}_\theta^* = \arg\min_{\mathcal{P}_\theta} [-J^{hyd}, J^{flood}, J^{supply}] \quad (10)$$

where the EMODPS simulation-based optimization is performed over the time horizon May 1st, 1990 to April 30th, 2010. For this formulation, we used different parallelization schemes of the Multi-Master Borg MOEA, namely 2, 4, 8, and 16 islands. Each island was allocated 512 processors and 1 million evaluations, requiring a total of 350,000 computational hours. These experiments served to simultaneously ensure convergence to a high quality set of solutions, and to identify the parallelization strategy offering a higher probability of discovering quasi-optimal solutions for subsequent optimization

[1] *IMRR - Integrated and sustainable water Management of the Red-Thai Binh Rivers System in changing climate research* project funded by the Italian Ministry of Foreign Affair as part of its development cooperation program.

runs. For each experiment, the optimization was performed on 50 random seeds to characterize the variability in the MOEA's stochastic search process. The final Pareto approximate set is obtained as the set of non-dominated solutions identified across all 50 seeds and 4 parallelization schemes. The reliability of each parallelization scheme is quantified using empirical attainment functions that measure the probability that the algorithm can attain progressively higher levels of HV performance.

To assess the vulnerability of the deterministic history-based solutions to hydrologic uncertainties, we re-evaluated them over a 50-member stochastic ensemble Ξ composed by 20-year series of hydrologic inflows in the five sub-catchments. We generated these synthetic series on a monthly time step via Cholesky decomposition to preserve the auto-correlation andspatial correlation of the observed monthly streamflow, while also including plausible inflows outside those observed in the historical record [39]. We then disaggregated the synthetic monthly flows to daily flows using the method of [40], which proportionally scales historical daily flows at each site from a probabilistically selected month of the historical record such that the synthetic monthly total is preserved.

To guarantee the design of robust control policies with respect to the hydrologic uncertainty, the Monte Carlo simulations over Ξ were then embedded into the EMODPS policy design in the *stochastic formulation*:

$$\mathcal{P}^*_\theta = \arg\min_{\mathcal{P}_\theta} \max_{\Xi} [-J^{hyd}, J^{flood}, J^{supply}] \tag{11}$$

where the uncertainty introduced by the Monte Carlo simulations is filtered adopting the minimax approach, which minimizes the objectives in the worst-case realization to capture the high risk aversion expressed by the local stakeholders in the system. This approach identifies robust control policies able to guarantee certain performance. The minimax operator has been independently applied for each objective, thus discounting the correlations among the objectives and providing an estimated lower-bound performance for each objective. In the stochastic optimization, 5 seeds of the 16-island Multi-Master Borg MOEA were run using 512 processors and 400,000 evaluations per island, for a total of 400,000 computational hours. Again, the final stochastic Pareto approximate set is obtained as the set of non-dominated solutions identified from the results of the five optimization trials. All optimizations were performed on the Texas Advanced Computing Center (TACC) Stampede Cluster (http://www.tacc.utexas.edu/stampede). The 6400 nodes of the TACC Stampede system each contain two Intel Xeon E5 processors and one Intel Xeon Phi Coprocessor, for a total of 102,400 processing cores.

We assess the quality of the Pareto approximate sets of control policies designed in the two problem formulations by measuring the hypervolume indicator HV[41], which accounts for both convergence and diversity of an approximate set of solutions $\mathcal{P}$ with respect to the Pareto optimal set $\mathcal{P}^*$ (or its best known approximation). The hypervolume measures the volume of objective space Y dominated ($\preceq$) by the considered approximate set, with HV formally defined as:

$$\begin{aligned} HV(\mathcal{P}, \mathcal{P}^*) &= \frac{\int \alpha_{\mathcal{P}}(\mathbf{y}) dy}{\int \alpha_{\mathcal{P}^*}(\mathbf{y}) dy} \quad \text{where} \\ \alpha_{\mathcal{P}}(\mathbf{y}) &= \begin{cases} 1 & \text{if } \exists \mathbf{y}' \in \mathcal{P} \text{ such that } \mathbf{y}' \preceq \mathbf{y} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{12}$$

C. Numerical results

The results of the deterministic optimization, reported in Fig. 3a, show that a clear tradeoff emerges between flood damages (J^{flood}, plotted on the x-axis) and hydropower production (J^{hyd}, plotted on the y-axis), while most policies have a fairly low water supply deficit (J^{supply}, represented by circles' size). In fact, the best performance in terms of J^{flood} (circles in the bottom-left part of the figure) is obtained by maintaining the water level in the reservoirs as low as possible to buffer the monsoon peak of the inflows, thus negatively impacting the energy production due to the reduction in the water turbined when the reservoir is almost empty. On the contrary, the best solutions in terms of J^{hyd} (circles in the top-right part of the figure) prefer to keep the reservoirs at their maximum level and to release the maximum turbine capacity for the entire year, thus reducing the buffering capacity during the monsoon season. The conflict between J^{flood} and J^{supply} is weaker because the drawdown of the reservoirs during the monsoon season indirectly makes large volumes of water available for irrigation in the Red River delta.

Fig. 3b shows the empirical cumulative distribution function of HV (eq. (12)) across the 50 random seeds for different parallelization schemes. Nearly all the runs achieve at least 70% of the highest HV by the end of the search, confirming the effectiveness of the Multi-Master Borg MOEA in addressing the complex search of the optimal policy parameters. The results show that the probability of attaining high levels of HV (i.e., its algorithmic reliability) increases with the number of islands. The performance of the lowest quantiles improves the most, demonstrating the benefit of using multiple islands for reducing search failures. Even the worst seed of the 16-island formulation attains over 78% HV, suggesting that running 50 seeds would have been unnecessary with this parallelization scheme of the Multi-Master Borg MOEA.

The results in Fig. 3 are obtained for the deterministic problem formulation (see eq. (10)), where the control policy performance $E[J(\mathcal{P}_\theta, \mathbf{x}_0, \varepsilon^h_1(k))]$ is evaluated over a single historical realization of the stochastic disturbances (i.e., $k = 1$). To quantify the vulnerability of these solutions to hydrologic uncertainties, we re-evaluated their performance over the 50-member stochastic ensemble of synthetically generated inflows. Results are reported in Fig. 4a, where the deterministic Pareto approximate set evaluated over history is shown in dark blue, with its performance over the stochastic ensemble in light blue. This stochastic re-evaluation adopting the minimax filter to characterize the extremely high risk aversion of the stakeholders, especially with respect to flood damages, shows significant performance degradation, on average equal to 11% in J^{hyd}, 899% in J^{flood}, and 347% in J^{supply}. In order to reduce the vulnerability of the designed solutions, we solved

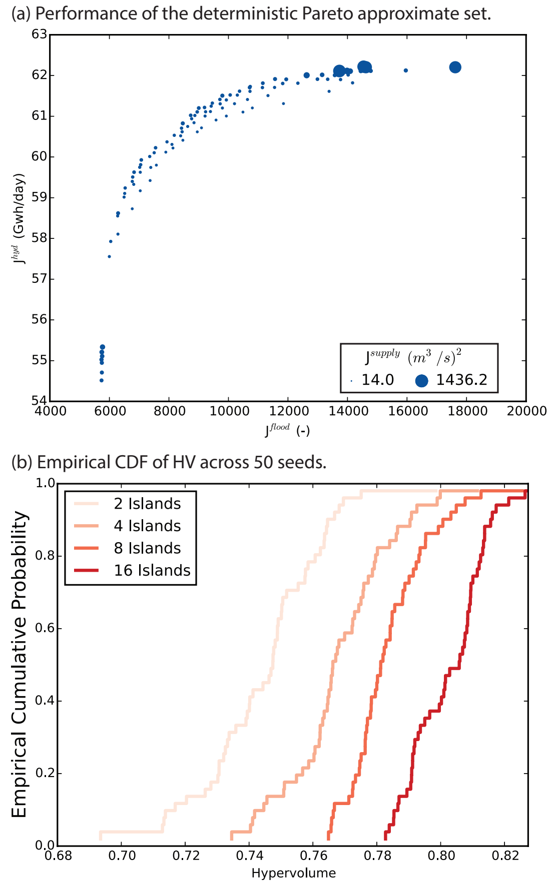

Fig. 3. Performance of the deterministic Pareto approximate policies over history (1990-2010), where hydropower production and flood damages are plotted on the primary axes, while the circles' size represents water supply deficit (panel a). Reliability of different parallelization schemes of the Multi-Master Borg MOEA for the historical optimization (panel b) .

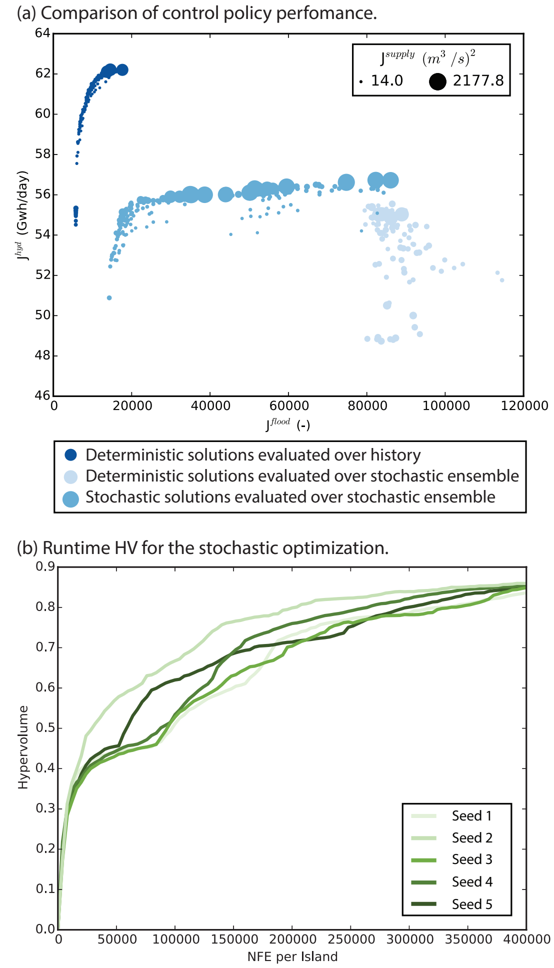

Fig. 4. Comparison of the deterministic control policies' performance evaluated over historical inflows (dark blue circles) and their re-evaluation over the stochastic ensemble (light blue circles), with the performance of the stochastic Pareto approximate set (middle tone circles), where hydropower production and flood damages are plotted on the primary axes, while the circles' size represents water supply deficit (panel a). Runtime search dynamics in terms of HV of the five stochastic optimization seeds for different Number of Function Evaluations (panel b).

the stochastic EMODPS problem formulation (see eq. (11)) to discover robust control policies with respect to the 50-member ensemble of synthetically generated inflows. Since the function evaluation now is 50 times more expensive, far fewer evaluations could be performed, and over a much smaller number of seeds. Given the results illustrated in Fig. 3b, we used the 16-island Multi-Master Borg MOEA as it should guarantee a high probability of convergence to quasi-optimal solutions. The results in Fig. 4a show that the stochastic Pareto approximate set (middle tone circles) successfully reduces the vulnerability of the deterministic solutions to the hydrologic uncertainty (light blue circles), attaining an average improvement of 2.4% in J^{hyd}, 65% in J^{flood}, and 44% in J^{supply}. Finally, the analysis of the runtime search dynamics in terms of the values of HV during the stochastic optimization (Fig. 4b) demonstrates that, despite the low number of seeds run, the adopted 16-island Multi-Master Borg MOEA is able to reliably find good solutions, as all the 5 seeds converged to greater than 80% of the highest HV within 400,000 evaluations per island.

V. Conclusions

In this paper, we contribute a massively parallel implementation of the evolutionary multi-objective direct policy search method for controlling large scale water resources systems under uncertainty. The proposed approach is shown to be scalable for large-scale real-world applications, where it overcomes some limitations of traditional approaches (i.e., the curse of modelling and the curse of multiple objectives) in solving the associated multi-objective optimal control problem, formulated as a discrete-time, periodic, nonlinear, stochastic Markov Decision Process. Numerical results show that the combination of direct policy search with nonlinear approximating networks and the Multi-Master Borg MOEA supports the identification of alternative control policies able to address multidimensional tradeoffs as well as severe hydrologic uncertainties. Ongoing research activities are focused on extensive testing of different parallelization schemes of the Multi-Master Borg MOEA and on better understanding the overall impacts of using different filtering criteria with respect to the resulting solutions. Moreover, we will explore the possibility of using additional climate information (e.g., El Nino Southern Oscillation) for forecasting the intensity of the monsoon and improving the system operations, particularly during extreme climate events.

References

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

[2] F. Delebecque and J. Quadrat, "Contribution of stochastic control singular perturbation averaging and team theories to an example of large-scale systems: Management of hydropower production," *IEEE Transactions on Automatic Control*, vol. 23, no. 2, pp. 209–222, 1978.

[3] C. Ocampo-Martinez, D. Barcelli, V. Puig, and A. Bemporad, "Hierarchical and decentralised model predictive control of drinking water networks: Application to barcelona case study," *Control Theory & Applications, IET*, vol. 6, no. 1, pp. 62–71, 2012.

[4] A. Sadowska, B. De Schutter, and P.-J. van Overloop, "Delivery-oriented hierarchical predictive control of an irrigation canal: Event-driven versus time-driven approaches," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1701–1716, 2015.

[5] M. Foo, S. K. Ooi, and E. Weyer, "System identification and control of the broken river," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 2, pp. 618–634, 2014.

[6] N. L. Poff, C. M. Brown, T. E. Grantham, J. H. Matthews, M. A. Palmer, C. M. Spence, R. L. Wilby, M. Haasnoot, G. F. Mendoza, K. C. Dominique *et al.*, "Sustainable water management under future uncertainty with eco-engineering decision scaling," *Nature Climate Change*, 2015.

[7] D. White, "Multi-objective infinite-horizon discounted markov decision processes," *Journal of Mathematical Analysis and Optimization*, vol. 89, no. 2, pp. 639–647, 1982.

[8] J. Labadie, "Optimal operation of multireservoir systems: State-of-the-art review," *Journal of Water Resources Planning and Management*, vol. 130, no. 2, pp. 93–111, 2004.

[9] M. Baglietto, C. Cervellera, M. Sanguineti, and R. Zoppoli, "Water reservoirs management under uncertainty by approximating networks and learning from data," in *Topics on System Analysis and Integrated Water Resource Management*. Elsevier Amsterdam, NL, 2006, pp. 117–139.

[10] R. Bellman, *Dynamic programming*. Princeton: Princeton University Press, 1957.

[11] J. Tsitsiklis and B. Van Roy, "Feature-Based Methods for Large Scale Dynamic Programming," *Machine Learning*, vol. 22, pp. 59–94, 1996.

[12] W. Powell, *Approximate Dynamic Programming: Solving the curses of dimensionality*. NJ: Wiley, 2007.

[13] P. M. Reed and J. B. Kollat, "Visual analytics clarify the scalability and effectiveness of massively parallel many-objective optimization: A groundwater monitoring design example," *Advances in Water Resources*, vol. 56, pp. 1–13, 2013.

[14] W. Washington, L. Buja, and A. Craig, "The computational future for climate and earth system models: on the path to petaflop and beyond," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1890, pp. 833–846, 2009.

[15] M. Giuliani, A. Castelletti, F. Pianosi, E. Mason, and P. Reed, "Curses, tradeoffs, and scalable management: advancing evolutionary multi-objective direct policy search to improve water reservoir operations," *Journal of Water Resources Planning and Management*, vol. 142, no. 2, 2016.

[16] S. Formentin, A. Karimi, and S. M. Savaresi, "Optimal input design for direct data-driven tuning of model-reference controllers," *Automatica*, vol. 49, no. 6, pp. 1874–1882, 2013.

[17] S. Formentin, K. Heusden, and A. Karimi, "A comparison of model-based and data-driven controller tuning," *International Journal of Adaptive Control and Signal Processing*, vol. 28, pp. 882–897, 2012.

[18] R. Zoppoli, M. Sanguineti, and T. Parisini, "Approximating networks and extended ritz method for the solution of functional optimization problems," *Journal of Optimization Theory and Applications*, vol. 112, no. 2, pp. 403–440, 2002.

[19] R. Busa-Fekete, B. Szörényi, P. Weng, W. Cheng, and E. Hüllermeier, "Preference-based reinforcement learning: evolutionary direct policy search using a preference-based racing algorithm," *Machine Learning*, vol. 97, no. 3, pp. 327–351, 2014.

[20] D. Hadka and P. Reed, "Large-scale parallelization of the borg multiobjective evolutionary algorithm to enhance the management of complex environmental systems," *Environmental Modelling & Software*, vol. 69, pp. 353–369, Jul. 2015.

[21] P. Fleming, R. Purshouse, and R. Lygoe, "Many-Objective optimization: an engineering design perspective," in *Proceedings of the Third international conference on Evolutionary Multi-Criterion Optimization*, Guanajuato, Mexico, 2005, pp. 14–32.

[22] E. Brill., J. Flach, L. Hopkins, and S. Ranjithan, "MGA: A Decision Support System for Complex, Incompletely Defined Problems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 4, pp. 745–757, 1990.

[23] M. Giuliani, J. Herman, A. Castelletti, and P. Reed, "Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management," *Water Resources Research*, vol. 50, pp. 3355–3377, 2014.

[24] P. M. Reed and D. Hadka, "Evolving many-objective water management to exploit exascale computing," *Water Resources Research*, vol. 50, no. 10, pp. 8367–8373, Oct. 2014.

[25] M. Deisenroth, G. Neumann, and J. Peters, "A Survey on Policy Search for Robotics," in *Foundations and Trends in Robotics*, 2011, vol. 2, pp. 1–142.

[26] M. Giuliani, E. Mason, A. Castelletti, F. Pianosi, and R. Soncini-Sessa, "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), 24-29 August 2014.

[27] J. Zatarain-Salazar, P. Reed, J. Herman, M. Giuliani, and A. Castelletti, "A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control," *Advances in Water Resources*, vol. 92, pp. 172–185, 2016.

[28] P. Marbach and J. N. Tsitsiklis, "Approximate gradient methods in policy-space optimization of Markov reward processes," *Discrete Event Dynamic Systems*, vol. 13, no. 1-2, pp. 111–148, 2003.

[29] L. Busoniu, D. Ernst, B. De Schutter, and R. Babuska, "Cross–Entropy Optimization of Control Policies With Adaptive Basis Functions," *IEEE Transactions on systems, man and cybernetics–Part B: cybernetics*, vol. 41, no. 1, pp. 196–209, 2011.

[30] M. Giuliani, F. Pianosi, and A. Castelletti, "Making the most of data: an information selection and assessment framework to improve water systems operations," *Water Resources Research*, vol. 51, no. 11, pp. 9073–9093, 2015.

[31] D. Hadka and P. Reed, "Borg: An Auto–Adaptive Many–Objective Evolutionary Computing Framework," *Evolutionary Computation*, vol. 21, no. 2, pp. 231–259, 2013.

[32] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multiobjective optimization," *Evolutionary computation*, vol. 10, no. 3, pp. 263–282, 2002.

[33] J. B. Kollat and P. M. Reed, "A computational scaling analysis of multiobjective evolutionary algorithms in long-term groundwater monitoring applications," *Advances in Water Resources*, vol. 30, no. 3, pp. 335–353, 2007.

[34] K. Deb, M. Mohan, and S. Mishra, "Evaluating the epislon-domination based multiobjective evolutionary algorithm for a quick computation of pareto-optimal solutions," *Evolutionary Computation Journal*, vol. 13, no. 4, pp. 501–525, 2005.

[35] E. Cantù-Paz, *Efficient and Accurate Parallel Genetic Algorithms*. Springer US, 2001. [Online]. Available: http://dx.doi.org/10.1007/978-1-4615-4369-5

[36] N. Q. Dinh, "Multi-objective evolutionary algorithm, dynamic and non-dynamic emulators in the desing of optimal policies for water resources management," Ph.D. dissertation, Politecnico di Milano, 2015.

[37] T. Hashimoto, J. Stedinger, and D. Loucks, "Reliability, resilience, and vulnerability criteria for water resource system performance evaluation," *Water Resources Research*, vol. 18, no. 1, pp. 14–20, 1982.

[38] M. Giuliani, D. Anghileri, P. Vu, A. Castelletti, and R. Soncini-Sessa, "Large storage operations under climate change: expanding uncertainties and evolving tradeoffs," *Environmental Research Letters*, vol. 11, no. 3, 2016.

[39] B. Kirsch, G. Characklis, and H. Zeff, "Evaluating the impact of alternative hydro-climate scenarios on transfer agreements: Practical improvement for generating synthetic streamflows," *Journal of Water Resources Planning and Management*, vol. 139, no. 4, pp. 396–406, 2013.

[40] K. Nowak, J. Prairie, B. Rajagopalan, and U. Lall, "A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow," *Water Resources Research*, vol. 46, 2010.

[41] E. Zitzler, L. Thiele, M. Laumanns, C. Fonseca, and V. da Fonseca, "Performance assessment of multiobjective optimizers: an analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, 2003.