Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Improving the protection of aquatic ecosystems by dynamically constraining reservoir operation via direct policy conditioning

Matteo Giuliani * **Andrea Castelletti** * **Patrick M. Reed** **

* *Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milan, Italy (e-mail: giuliani@elet.polimi.it; andrea.castelletti@polimi.it).*
** *Department of Civil and Environmental Engineering, Cornell University, Ithaca NY, USA (e-mail: patrick.reed@cornell.edu).*

Abstract:
Water management problems generally involve conflicting and non-commensurable objectives. Assuming a centralized perspective at the system-level, the set of Pareto-optimal alternatives represents the ideal solution of most of the problems. Yet, in typical real-world applications, only a few primary objectives are explicitly considered, taking precedence over all other concerns. These remaining concerns are then internalized as static constraints within the problem's formulation. This approach yields to solutions that fail to explore the full set of objectives tradeoffs. In this paper, we propose a novel method, called direct policy conditioning (DPC), that combines direct policy search, multi-objective evolutionary algorithms, and input variable selection to design dynamic constraints that change according to the current system conditions. The method is demonstrated for the management problem of the Conowingo Dam, located within the Lower Susquehanna River, USA. The DPC method is used to identify environmental protection mechanisms and is contrasted with traditional static constraints defining minimum environmental flow requirements. Results show that the DPC method identifies a set of dynamically constrained control policies that overcome the current alternatives based on the minimum environmental flow constraint, in terms of environmental protection but also of the primary objectives.

Keywords: direct policy search; multi-objective optimization; environmental engineering

1. INTRODUCTION

The multi-objective nature of decision-making problems in the water management domain has long been recognized and studied (e.g., Cohon and Marks, 1975). The advantage of a centralized perspective at the system-level to generate a Pareto front explicitly considering all the objectives involved has been widely extolled (e.g., Soncini-Sessa et al., 2007). Yet, in the real world, the actual shift toward a truly multi-purpose, multi-stakeholder, integrated water management approach is not easily achievable. Most major reservoirs have had their rule curves defined in prior decades (U.S. Army Corps of Engineers, 1977), considering one of four primary uses: generation of power, flood risk reduction, irrigation or drinking water supply. In more recent years, bilateral negotiation processes have been adopted to include new operational targets, mostly related to environmental protection, recreation and transportation. Traditionally, the four primary uses take precedent over all other concerns and these secondary objectives are optimized assuming the four primary uses form constraints. However, the resulting reservoir operations are often inefficient as they fail to explore the full set of tradeoffs between evolving multi-sector objectives and preferences in the water system.

In this paper, we propose a novel method, called direct policy conditioning (DPC), which looks at the centralized Pareto-optimal solutions of the multi-objective reservoir operation problem to dynamically condition the design of control policies targeted only to the primary objectives. The DPC method relies on the direct policy search approach (DPS, see Rosenstein and Barto (2001)), which consists in the parameterization of the control policies and the subsequent search, in the parameter space, of the optimal parameter vector with respect to the operating objectives. The DPS approach can be extended to the identification of the set of Pareto-optimal solutions of the problem, designed explicitly considering a suite of conflicting operating objectives. This set, therefore, represents the best achievable performance at the system-level. The DPC method, instead, maintains the original preference structure and designs the control policy considering only the primary objectives. The Pareto-optimal set is used to formulate a set of constraints for the DPC policy design problem in order to account for the secondary objectives. These constraints are defined in the space of the policy

⋆ Matteo Giuliani was partially supported by *Fondazione Fratelli Confalonieri.*

parameters and, therefore, directly influence the functional control policy space. The result is the definition of dynamic constraints that, like the control policy, are able to exploit the feedback provided by the system conditions. Input variable selection techniques are used to support the identification of a "signature" in the decision space with respect to the desired output, namely a subset of control policy parameters that are more related to the secondary objectives. The hybrid model-based/model-free iterative input selection (IIS) algorithm (Galelli and Castelletti, 2013) is used in this work. The conditioning is then applied only to the selected variables, while no-constraints are defined on the remaining parameters to guarantee flexibility on the resulting control policies.

We apply DPC on a common problem in the water resources field, namely the protection of the aquatic ecosystems downstream of a dam, threatened by the alteration of the natural flow regime due to hydropower-driven reservoir operations (e.g., Richter et al., 1996). In most of the cases, environmental concerns are not considered as operating objectives, probably because the definition of general principles for flow regime management is challenging (Arthington et al., 2006) and the estimation of economic damages associated with flow alternations is highly uncertain (Brander et al., 2006). Minimum environmental flow (MEF) constraints on the release decisions are instead imposed on the reservoir operations, possibly with seasonal or monthly varying values. However, MEF is a constraint defined with respect to the release decisions and does not directly act on the real decision space, which is the functional space of the reservoir control policy.

The method is demonstrated on the management problem of the Conowingo Dam, located within the Lower Susquehanna River, USA. The Lower Susquehanna River is an interstate water body that has been subject to intensive water management efforts due to competing demands from urban water supply, atomic power plant cooling, hydropower production, federally regulated environmental flows, and recreational interests (Giuliani et al., 2014). We first compute the Pareto-optimal control policies explicitly considering all the water uses involved. Then, the IIS algorithm identifies the subset of policy parameters that are more related to the environment objective. A set of dynamic constraints on the selected parameters are then defined to condition the control policy, exploiting the feedback provided by the system conditions. Conversely, the MEF constraint does not consider the system state (e.g., the current reservoir storage) and provide a static constraint.

The paper is organized as follows. In the next section, the methodological aspects of the proposed approach are discussed, while Section 3 presents a short description of the case study. In Section 4 the solutions obtained with the proposed approach are compared with the one based on traditional MEF constraints. Final remarks and issues for further research are presented in the last section.

2. METHODOLOGY

The DPC method is based on a three-step procedure: *i*) multi-objective problem formulation (including all the operating objectives, both primary $\mathbf{J}^p$ and secondary $\mathbf{J}^s$) and solution via direct policy search to design a set of Pareto-optimal policies representing the best achievable performance and the associated policy parameterizations; *ii*) identification of the subset of optimal control policy parameters that are more related to the secondary objectives, generally excluded in the design of the actual reservoir regulation; *iii*) design of a control policy considering only the primary objectives, subject to a set of dynamic constraints accounting for the secondary ones.

2.1 Multi-objective problem formulation

Water reservoir management problems can be formulated as q-objective, stochastic, periodic, non-linear, closed-loop optimal control problems (e.g., Castelletti et al., 2008, and references therein) and solved via dynamic programming (DP) family methods (e.g., Powell, 2007). However, the practicability of these methods in large-scale water systems is very limited due to the curse of dimensionality (Bellman, 1957) and the curse of modeling (Tsitsiklis and Van Roy, 1996). Moreover, DP-based approaches are single-objective methods. They can be thus used to solve multi-objective problems only by reformulating them as a family of parametric single-objective problems and reiteratively running a single-objective optimization for different values of the parameter to explore the Pareto front. This remarkably affects the computational requirements, as the number of single-objective problems to solve grows exponentially with the objectives number.

Direct policy search (DPS, see Rosenstein and Barto (2001)) approaches overcome these limitations. DPS is a simulation-based method which directly operates in the policy space by parameterizing the control policy π_θ, where $\theta \in \Theta$ is a vector of unknown time-varying parameters. The water management problem can be therefore formulated as

$$\pi_\theta^* = \arg\min_{\pi_\theta} \mathbf{J}(\pi_\theta) \tag{1}$$

where $\mathbf{J}(\pi_\theta) = \left[J^1, \ldots, J^q\right]$ is the vector objective function (including both $\mathbf{J^P}$ and $\mathbf{J}^s$) computed by simulating the system over the time horizon H under the parameterized policy π_θ. The state of the system $\mathbf{x}_t \in \mathbb{R}^m$ evolves accordingly to a state transition function $\mathbf{x}_{t+1} = f_t\left(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\varepsilon}_{t+1}\right)$, where $\mathbf{u}_t \in \mathcal{U}_t(\mathbf{x}_t) \subseteq \mathbb{R}^n$ is the control vector which is determined by the policy π_θ, and $\boldsymbol{\varepsilon}_{t+1} \in \mathbb{R}^l$ the stochastic disturbance vector. Finding π_θ^* is equivalent to find the corresponding optimal policy parameters θ^*.

In addition to overcome the dual DP curse of dimensionality and modeling, DPS problems can be solved using multi-objective evolutionary algorithms (MOEAs) to obtain an approximation of the Pareto front in a single run of the algorithm. The effectiveness of this approach depends on the flexibility of the selected class of functions used to define the policy (Tikk et al., 2003) and on the ability of the optimization algorithm to deal with a large number of objectives. In this work, we use gaussian Radial Basis Functions (RBFs) to parameterize the policies as they are capable of representing functions for a large class of problems (Busoniu et al., 2011). The k-th control variable in the vector $\mathbf{u}_t$ (with $k = 1, \ldots, n$) is defined as:

$$u_t^k = \sum_{i=1}^{N} w_{i,k} \varphi_i(\chi_t) \tag{2a}$$

with

$$\varphi_i(\chi_t) = \exp\left[-\sum_{j=1}^{M} \frac{((\chi_t)_j - c_{j,i})^2}{b_{j,i}^2}\right] \quad (2b)$$

where N is the number of RBFs $\varphi(\cdot)$, $w_{i,k}$ the weight of the i-th RBF, M the number of input variables χ_t, and $\mathbf{c}_i, \mathbf{b}_i$ are the M-dimensional center and radius vectors of the i-th RBF, respectively.
To perform the optimization, we use the self-adaptive Borg MOEA (Hadka and Reed, 2013), which has been shown to be highly robust across a diverse suite of challenging multi-objective problems, where it met or exceeded the performance of other state-of-the-art MOEAs (Hadka and Reed, 2012).

2.2 Iterative Input Selection algorithm

Input variable selection (IVS) problems arise every time a variable of interest has to be modeled as a function of a subset of potential explanatory variables, or predictors, but there is uncertainty about which subset to use among a number, usually large, of candidate sets available (George, 2000). In this work, IVS techniques are used to characterize the complex relationships and feedbacks between operating objectives, management decisions, and the system dynamic evolution. The aim is to identify the subset of control policy parameters that are more related to the policy performance in terms of the secondary objectives of the problem.
The hybrid model-based/model-free, forward-selection, iterative input selection (IIS) algorithm (Galelli and Castelletti, 2013) is used in this work. Given the output variables v^o and the set of candidate inputs $\mathbf{v}^i$, the IIS algorithm first ranks these latter with respect to a statistical measure of significance and adds only the best performing input v^* to the set of selected variables $\mathcal{V}$. This operation aims to avoid the inclusion of redundant variables, as one an input is selected, all the inputs highly correlated with it may become useless in the next iterations. Then, the algorithm estimates a model of v^o with input $\mathcal{V}$, such as $v^0 = \hat{m}(\mathcal{V})$, and computes the model performance with a suitable distance metric D (e.g., the coefficient of determination) as well as the model residuals, which become the new output at the next iteration. The algorithm stops when the best variable returned by the rank is already in the set $\mathcal{V}$, or when over-fitting conditions are reached. Among the many alternative model classes, IIS relies on extremely randomize trees (Extra-Trees), a tree-based method proposed by Geurts et al. (2006) that was empirically demonstrated to outperform other models in terms of modeling flexibility, computational efficiency, and scalability with respect to the input dimensionality.

2.3 Direct policy conditioning

The aim of DPC is to effectively condition the control policies designed considering the primary objectives in order to take into account the secondary objectives. This conditioning is defined by exploiting the reference provided by the Pareto-optimal solutions of Problem (1), namely the set of optimal control policy parameters θ^*. To guarantee the flexibility of the resulting policy, the conditioning is applied to a subvector of parameters $\tilde{\theta} \subseteq \theta$, which is identified through the IIS algorithm (see the previous section) on a dataset where the candidate inputs are policy parameters, while the output variables are the performance in terms the secondary objectives. This dataset can be generated via random sampling of the policy parameters' space or, alternatively, the Pareto-optimal set can be directly used. In the first case (i.e., random sampling), a set of control policy parameters vectors is randomly sampled in Θ. Each vector defines a randomly generated control policy. The system is hence simulated following the randomly generated policy to compute the operating objectives $\mathbf{J}^p$ and $\mathbf{J}^s$. The combination of the randomly sampled policy parameters and their performance in terms of the secondary objectives forms the IIS dataset. In the second case (i.e., Pareto-optimal set), the dataset used is composed by the set of Pareto-optimal policy parameterizations, obtained solving the multi-objective problem with respect to all the operating objectives, and their corresponding performance in the secondary objectives.
Given the Pareto-optimal set θ^* and the extracted subvector $\tilde{\theta}$, the DPC problem can be formulated as:

$$\pi_{\theta}^* = \arg\min_{\pi_\theta} \mathbf{J}^p(\pi_\theta) \quad (3a)$$

s.t.

$$\theta \in \Theta' \subseteq \Theta \quad (3b)$$

$$\Theta' = \{\theta \in \Theta : \theta_i = \theta_i^* \pm \gamma, \forall \theta_i \in \tilde{\theta}\} \quad (3c)$$

where the objective function vector $\mathbf{J}^p$ includes only the primary objectives and the subvector $\tilde{\theta}$ is constrained to take values in a neighborhood of size γ of the optimal values θ^*. These constraints impose the optimal values of the policy parameters according to the signature identified by the IIS algorithm. The aim of this conditioning is to partially consider the secondary objectives in the design of the control policy targeted only to the primary objectives.

3. CASE STUDY DESCRIPTION

The potential of DPC for the identification of environmental protection measures is demonstrated on the Lower Susquehanna system, USA (Figure 1). The Susquehanna River is the longest river on the eastern United States, draining a catchment area of about 71,000 km^2 through New York, Pennsylvania, and Maryland, ultimately contributing 50% of the freshwater flowing into the Chesapeake Bay. The Conowingo reservoir is an interstate water body shared by Pennsylvania and Maryland in the Lower Susquehanna, about 16 km from the Susquehanna River mouth. The dam, which was completed in 1928 for hydropower generation purposes, is the largest non–federal dam in the U.S. regulating a large share of the flow in the Lower Susquehanna with substantial impacts on multiple stakeholders. The Conowingo reservoir contributes to the water supply of Chester (PA) and Baltimore (MD). Conowingo releases are also critical for cooling the Peach Bottom atomic power plant and downstream releases are subject to minimum flow requirements defined by the Federal Energy Regulatory Commission (FERC) to protect fishery resources. Moreover, in 1968 the reservoir was connected to the Muddy Run Pumped Storage Hydroelectric Facility, which cycles water back and forth from Conowingo for additional power generation. Finally, the Conowingo reservoir provides valuable recreational and ecosystem services. In this paper, we focus on the operation of the Conowingo dam, while a fixed weakly rule

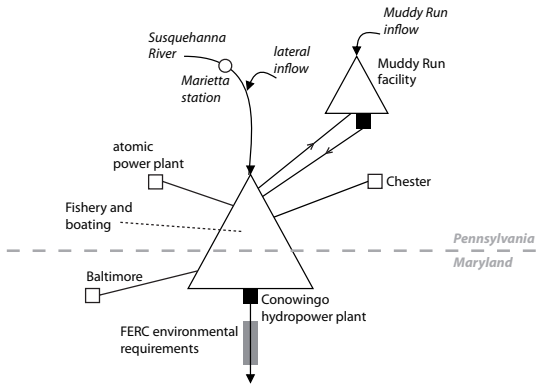

Fig. 1. Schematic representation of the main components of the Lower Susquehanna model.

is assumed for the operation of Muddy Run according to the hydropeaking strategy reported in Swartz (2006).
The multi-objective control problem requires to design a set of Pareto-optimal policies. They are defined as multi-input multi-output RBFs (accounting for 32 parameters), which provide the release decision vector $\mathbf{u}_t$ (i.e., the downstream release as well as the ones for the public water supply) as a function of time t and reservoir level h_t. Six operating objectives measures the multi-stakeholder interests affected by the Conowingo reservoir operation (for further details about the problem formulation see Giuliani et al., 2014):

- *Hydropower revenue* (to be maximized);
- *Atomic Power Plant*, *Baltimore*, and *Chester* water supply volumetric reliability (to be maximized);
- *Recreation*, modeled by the storage reliability (to be maximized);
- *Environment*, modeled by the shortage index with respect to FERC flow requirements (to be minimized).

The FERC minimum flow requirements introduced in 1988 protect fishery resources threatened by the hydropower management of the dam (Federal Energy Regulatory Commission, 1989). In average flow conditions, water availability is generally sufficient to maintain hydroelectric operations, water supply, meet environmental flow requirements, and sustain recreational activities. Yet, in low flow conditions challenging tradeoffs emerge for Conowingo operations to supply water to Baltimore, Chester, and the Peach Bottom atomic power plant, while seeking to minimize negative impacts on the recreational and touristic interests. Growing regional water demands and climate change are additional concerns for the SRBC. As a recent example, the SRBC coordinated a regional planning effort assessing a set of alternative modifications to the FERC requirements to mitigate the negative impacts of the low reservoir levels (Swartz, 2006). Given the results of the modeled alternatives, the possibility of including the leakages from the closed dam gates (i.e., 22.65 m^3/s) toward meeting the downstream minimum flow requirements has been selected as the most critical action in managing the Conowingo dam during drought periods. The result of this intensive planning effort is the identification of four alternatives management strategies for implementing the credit for leakages and specifying the hydrologic conditions under which this credit is warranted.

4. APPLICATION RESULTS

The decision analytic framework proposed by Giuliani et al. (2014) combines reservoir policy identification and many-objective optimization under uncertainty to characterize the current baseline operations of the Conowingo reservoir and discover key tradeoffs between alternative operating policies. In Giuliani et al. (2014), two formulations of the MODPS problem are contrasted: the historical case, as formulated in eq. (1), which evaluates the operating objectives over historical hydroclimatic conditions; its stochastic extension, where the same objectives are instead evaluated over an ensemble Ξ of stochastic inflows and evaporation rates realizations. The uncertainty is then filtered adopting a minimax approach, which minimizes the objectives in the worst-case realization. Results demonstrate that the system's current history-based operations are negatively biased to overestimate the reliability of the reservoir's multi-sector services. Moreover, the framework identified a set of stochastic Pareto-optimal operating policies that are more robust to hydroclimatic uncertainties, while being capable of better addressing the tradeoffs across the Conowingo Dam's multi-sector services. Yet, the recommended policies are designed by assuming the SRBC effectively controls the Lower Susquehanna system and takes into account all the stakeholders involved. Under this assumption, the stochastic Pareto-optimal policies represents the best achievable performance at the system-level, but they do not reflect the structure of preferences driving the actual system operation. According to Giuliani et al. (2014), hydropower revenue and water supply are demonstrated to strongly shape the identified current operating policy of the Conowingo reservoir, with the FERC flow requirements guaranteeing adequate environmental protection. However, the performance in terms of environmental shortage index under stochastic hydroclimatic conditions is a significant concern under the baseline policy as well as the alternatives negotiated by the SRBC, which are all based on the FERC minimum flow requirements. The adoption of direct policy conditioning (DPC, see Section 2) provides alternative mechanisms to better constrain the reservoir operation in order to protect the environment.

4.1 Iterative input selection

In order to identify the most significant decision variables with respect to the Environment objective, the IIS algorithm is run working on two different datasets.
In the first experiment, which will be termed DPC on random samples, the dataset was generated with a design of experiments that samples the RBFs parameter space using the quasi-random Sobol sequence (Sobol, 2001) coupled with the cross-sampling method proposed by Saltelli et al. (2008). In this paper, 10,000 parameter sets are generated from the Sobol sequence, with the cross-sampling method creating a total of 660,000 parameterizations of the RBFs policies. Each policy is then simulated to obtain the corresponding performance. The Environment objective is selected as output to explain, while the candidate input for IIS are the parameters of the RBFs policies. Figure 2a shows the 18 parameters selected by the IIS algorithm and their contribution in explaining the output (i.e., Environment objective) variance. Not surprisingly, IIS selects

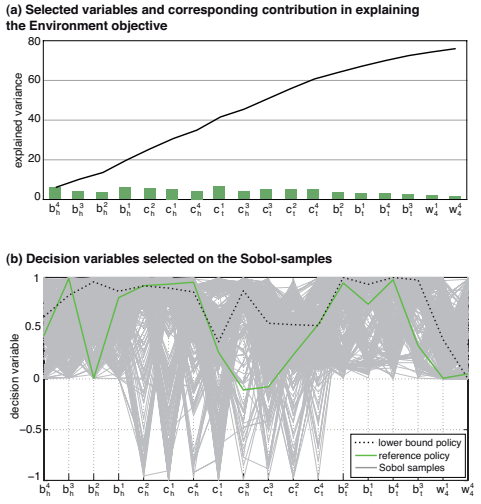

Fig. 2. Representation of the IIS results on random samples. Panel (a) shows the set of selected variables and their contributions in explaining the Environment objective. Panel (b) represents the values of the selected decision variables for different policies

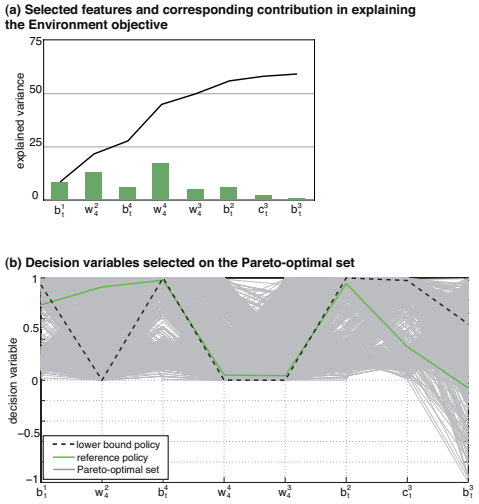

Fig. 3. Representation of the IIS results on the Pareto-optimal set. Panel (a) shows the selected variables and their contributions in explaining the Environment objective. Panel (b) represents the selected decision variables for different policies.

all centers and radii of the RBFs, which are strongly affecting the policy shape and, consequently, the release decisions impacting on the Environment objective. Figure 2b shows the values of the selected parameters in the reference policy (i.e., the one with the best performance in terms of environmental shortage index) with respect to the sampled parameter sets. Moreover, the RBFs parameters in the lower bound policy with respect to the Environment objective(i.e., the policy obtained removing the FERC flow requirements) are shown to highlight the main differences with respect to the reference policy.

In the second experiment, which will be termed DPC on the Pareto-optimal set, the IIS algorithm is run directly on the parameterizations of the Pareto-optimal policies. The underlying idea is that the information content of the Pareto-optimal set might be more informative in conditioning the control policies than a randomly sampled dataset. Figure 3a shows the parameters selected and their contribution in explaining the Environment objective. In this case, less parameters are selected (i.e., only 8), with a lower explained variance (i.e., 60% with respect to 75%). Figure 3b compares the values of the selected parameters in the reference and lower bound policies with respect to the Pareto-optimal set. Note that most of the selected parameters are RBFs radii related to the time variable (i.e., b_t^i), while no parameters related to the Conowingo level are selected. These results suggest that the time of the releases is the most significant factor affecting the Environment objective, while the level importance decreases when moving from random parameterizations to the Pareto-optimal set.

4.2 Direct policy conditioning policies

The two formulations of the direct policy conditioning problem (i.e., DPC on random samples and DPC on the Pareto-optimal set) replace the historical formulation of the Lower Susquehanna problem and the resulting policies are compared with the baseline policy and the four alternatives negotiated by the SRBC. In the DPC problem, the decision variables (i.e., the RBFs parameters) selected by the IIS algorithm are constrained to be close to the values assumed in the reference policy, namely the one with the best performance in terms of environmental shortage index under stochastic hydroclimatic conditions. Figure 4 reports the comparison of DPC policies with respect to the baseline and the SRBC negotiated alternatives. The performance of the stochastic Pareto-optimal policies and the one of the lower bound policy are reported as references. All the DPC policies overcome the performance of the lower bound policy with respect to the Environment objective. Moreover, most of these policies meet or improve the performance of the baseline and the SRBC negotiated alternatives in terms of environmental shortage index, demonstrating that the constraints imposed in the RBFs parameters decision space represents an effective alternative to the classical minimum environmental flow constraints. Finally, the flexibility of this dynamic conditioning mechanism, which exploits the feedback represented by the system condition by means of the corresponding operating policies, allows the identification of solutions attaining higher performance than the baseline in terms of Hydropower revenue and Atomic Power Plant.
Two solutions (identified by the thick lines) obtained with DPC on random samples and DPC on the Pareto-optimal set, respectively, were selected for further analysis. Results show that adopting the DPC on the Pareto-optimal set (random samples) solution instead of maintaining the current reservoir regulation, the SRBC would potentially attain an increase in the hydropower revenue equal to 18.6 (17.6) million US$/year, 0.29 (0.33) in the Atomic Power Plant reliability, and a reduction of 0.038 (0.032) in terms of environmental shortage index. These solutions therefore exhibits the potential to outperform the baseline regulation, producing a robust policy that reflects the actual structure of preference in the Lower Susquehanna problem and successfully protects the environment by means of effective conditioning of the Conowingo reservoir operating policy.

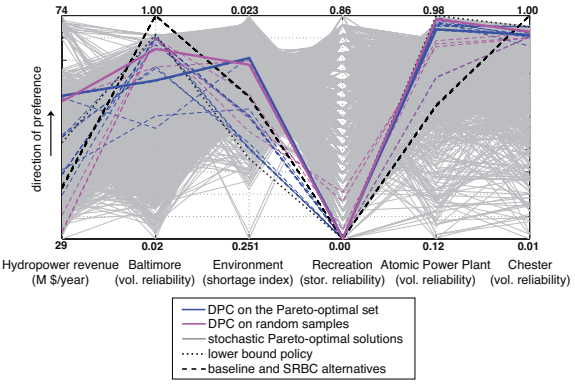

Fig. 4. Comparison of the performance over stochastic hydroclimatic conditions of the DPC policies, the baseline policy, and the alternatives negotiated by the SRBC.

5. CONCLUSIONS

The paper presents a novel method, called Direct Policy Conditioning, for designing effective environmental protection mechanisms in complex, stochastic, multi-objective water reservoir systems. This method provides encouraging results on the Lower Susquehanna management problem, which requires to deal with many competing demands and hydroclimatic uncertainty.

The proposed method first identifies an environmental signature in the stochastic Pareto-optimal policies through input variable selection techniques. This signature is then used to condition the decision space in order to design improved operating policies for the Conowingo Dam, which reflect the actual preference structure driving current system operation. Results show that the DPC policies exhibits the potential to outperform the baseline alternative in terms of Environment and also Hydropower revenue, Atomic Power Plant reliability, and Recreation.

Future research will assess the applicability of DPC in designing policies for different objective tradeoffs. Furthermore, DPC seems a promising method to support the identification of coordination mechanisms in multireservoir systems.

REFERENCES

Arthington, A., Bunn, S., Poff, N.L., and Naiman, R. (2006). The challenge of providing environmental flow rules to sustain river ecosystems. *Ecological Applications*, 16(4), 1311–1318.

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

Brander, L., Florax, R., and Vermaat, J. (2006). The empirics and wetlands valuation: A comprehensive summary and a meta-analysis of the literature. *Environmental & Resource Economics*, 33, 223–250.

Busoniu, L., Ernst, D., De Schutter, B., and Babuska, R. (2011). Cross–Entropy Optimization of Control Policies With Adaptive Basis Functions. *IEEE Transactions on systems, man and cybernetics–Part B: cybernetics*, 41(1), 196–209. doi:10.1109/TSMCB.2010.2050586.

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

Cohon, J.L. and Marks, D. (1975). A review and evaluation of multiobjective programing techniques. *Water Resources Research*, 11(2), 208–220.

Federal Energy Regulatory Commission (1989). Order Approving Settlement Agreement for the Conowingo Hydroelectric Project. Technical Report Project No. 405-009, Docket No. EL80-38-000.

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

George, E. (2000). The variable selection problem. *Journal of the American Statistical Association*, 95(452), 1304–1308.

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3–42.

Giuliani, M., Herman, J., Castelletti, A., and Reed, P. (2014). Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management. *Water Resources Research*. (In Press).

Hadka, D. and Reed, P. (2012). Diagnostic assessment of search controls and failure modes in many–objective evolutionary optimization. *Evolutionary Computation*, 20(3), 423–452.

Hadka, D. and Reed, P. (2013). Borg: An Auto–Adaptive Many–Objective Evolutionary Computing Framework. *Evolutionary Computation*, 21(2), 231–259.

Powell, W. (2007). *Approximate Dynamic Programming: Solving the curses of dimensionality*. Wiley, NJ.

Richter, B., Baumgartner, J., Powell, J., and Braun, D. (1996). A method for assessing hydrologic alteration within ecosystems. *Conservation biology*, 10(4), 1163–1174.

Rosenstein, M. and Barto, A. (2001). Robot weightlifting by direct policy search. In *International Joint Conference on Artificial Intelligence*, volume 17, 839–846. Citeseer.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and S., T. (2008). *Global sensitivity analysis: the primer*. Wiley.

Sobol, I. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1-3), 271–280.

Soncini-Sessa, R., , Cellina, F., Pianosi, F., and Weber, E. (2007). *Integrated and participatory water resources management: Practice*. Elsevier, Amsterdam, NL.

Swartz, P. (2006). Conowingo pond management plan. Technical Report 242, Susquehanna River Basin Commission.

Tikk, D., Kóczy, L., and Gedeon, T. (2003). A survey on universal approximation and its limits in soft computing techniques. *International Journal of Approximate Reasoning*, 33(2), 185–202.

Tsitsiklis, J. and Van Roy, B. (1996). Feature-Based Methods for Large Scale Dynamic Programming. *Machine Learning*, 22, 59–94.

U.S. Army Corps of Engineers (1977). *Reservoir System Analysis for Conservation, Hydrologic Engineering Methods for Water Resources Development*. Hydrologic Engineering Center, Davis, CA.