AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR020524

Key Points:
- We advance a rival framings framework for designing and evaluating alternative water systems management policies
- Testing multiple problem framings reduces the probability of formulating policies with unintended consequences
- Minimizing expected flood damages may be ineffective in preventing severe flooding

Supporting Information:
- Supporting Information S1

Correspondence to:
J. Quinn,
jdq8@cornell.edu

Citation:
Quinn, J. D., P. M. Reed, M. Giuliani, and A. Castelletti (2017), Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems, *Water Resour. Res.*, *53*, doi:10.1002/2017WR020524.

Received 1 FEB 2017
Accepted 17 JUL 2017
Accepted article online 21 JUL 2017

© 2017. American Geophysical Union. All Rights Reserved.

Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems

J. D. Quinn[1], P. M. Reed[1], M. Giuliani[2], and A. Castelletti[2,3]

[1]Department of Civil and Environmental Engineering, Cornell University, Ithaca, New York, USA, [2]Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy, [3]Institute of Environmental Engineering, ETH Zurich, Zurich, Switzerland

Abstract Managing water resources systems requires coordinated operation of system infrastructure to mitigate the impacts of hydrologic extremes while balancing conflicting multisectoral demands. Traditionally, recommended management strategies are derived by optimizing system operations under a single problem framing that is assumed to accurately represent the system objectives, tacitly ignoring the myriad of effects that could arise from simplifications and mathematical assumptions made when formulating the problem. This study illustrates the benefits of a rival framings framework in which analysts instead interrogate multiple competing hypotheses of how complex water management problems should be formulated. Analyzing rival framings helps discover unintended consequences resulting from inherent biases of alternative problem formulations. We illustrate this on the monsoonal Red River basin in Vietnam by optimizing operations of the system's four largest reservoirs under several different multiobjective problem framings. In each rival framing, we specify different quantitative representations of the system's objectives related to hydropower production, agricultural water supply, and flood protection of the capital city of Hanoi. We find that some formulations result in counterintuitive behavior. In particular, policies designed to minimize expected flood damages inadvertently increase the risk of catastrophic flood events in favor of hydropower production, while min-max objectives commonly used in robust optimization provide poor representations of system tradeoffs due to their instability. This study highlights the importance of carefully formulating and evaluating alternative mathematical abstractions of stakeholder objectives describing the multisectoral water demands and risks associated with hydrologic extremes.

1. Introduction

Managing both intraannual and interannual hydrologic variability has posed a continual challenge to human societies. This challenge is especially difficult for low income countries whose economies depend largely on agriculture, but lack the institutional and infrastructure capacity to adapt to variable hydrologic conditions [*Hall et al*., 2014]. Climate change is only expected to exacerbate this issue, as greater warming should increase both evaporation and precipitable water, paradoxically leading to both longer, more severe droughts and more intense flooding [*Trenberth*, 2011]. Recent observations indicate intensification of the hydrologic cycle has already begun [*Huntington*, 2006], with more frequent heat and precipitation extremes observed over the last half century [*Coumou and Rahmstorf*, 2012]. Again, these impacts are felt most by the disadvantaged, deepening poverty in low income, climate-dependent economies [*Olsson et al*., 2014; *Hallegatte et al*., 2015; *World Bank*, 2016]. Furthermore, as these economies grow and diversify out of agriculture, competition for water resources across their developing sectors will increase. In order to reduce, and if possible overcome, the negative impacts and water conflicts associated with hydrologic extremes, it is of paramount importance that innovative water management policies be discovered [*Tanaka et al*., 2006; *Giuliani et al*., 2016a; *World Bank*, 2016].

Conventionally, water resources managers have attempted to reduce the multisectoral impacts of hydrologic variability through optimized reservoir operations. Given that most river basins now contain multiple reservoirs, optimizing operations is mathematically challenging just considering the competing objectives and the high-dimensional and stochastic nature of the multireservoir control problem [*Giuliani et al*., 2016b;

Zatarain Salazar et al., 2016]. While addressing these challenges, this study also confronts the often-ignored epistemic uncertainties surrounding how to formulate the control problem itself. In classical decision theory, the problem formulation is designed to conform to the chosen modeling approach [*Tsoukias*, 2008], disregarding the fact that the chosen procedure will affect the predictions of the consequences of alternative solutions [*Majone and Quade*, 1980], and consequently which solutions are considered "optimal" [*Roy*, 1990]. *Kasprzyk et al*. [2009] and *Zeff et al*. [2014] illustrate this on separate multiobjective water supply portfolio planning problems in which the attainable system performance depends heavily on which objectives, constraints, and decisions are included in the optimization. More specifically, *Kasprzyk et al*. [2009] find that different families of solutions emerge from different formulations, with some formulations missing entire regions of decision relevant tradeoff solutions. As such, how one frames a problem can wield an inadvertent influence of power on the outcome [*Stirling*, 2008]. This has led natural resources managers to advocate for the exploration of alternative problem structures [*Hoppe*, 2011] in participatory planning processes in order to discover tensions between competing framings formulated under different perceptions of stakeholder values [*Bosomworth et al*., 2017].

Acknowledging that the most appropriate problem formulation is itself uncertain, in this study we explore alternative problem structures using what *Tsoukias* [2008] dubs a "constructive" decision aiding approach in which the problem formulation itself is constructed, not just the optimal solutions. Within the water resources literature, similar methods were developed in the 1960s under the Harvard Water Program through which *Maass et al*. [1962] proposed a four-step process for designing water resources systems: (1) identifying the objectives, (2) translating the objectives into design criteria, (3) using these criteria to design water resources development and management plans, and (4) evaluating the consequences of the plans that have been developed, in particular, by quantifying regrets associated with using one objective over another. Emerging from the early origins of behavioral economics, this approach has inspired new decision theories such as the version concept of *Roy* [2010], the rival problem framings concept of *Walker et al*. [2003], and de novo programming of *Zeleny* [1981], which *Kasprzyk et al*. [2012] expand on to explicitly capture multiple objectives. This approach is perhaps best described by *Zeleny* [1989]:

Making decisions does not mean finding our ways through a fixed maze (problem solving)—decision making refers to the very construction of that maze—ordering of nature so that we ourselves can find our way through it.

In this study, we highlight the importance of evaluating alternative constructions of the maze that is the multiobjective, multireservoir control problem through Vietnam's Red River basin, where operations of the four largest reservoirs must balance agricultural water demands for food and energy production, while also reducing flood risks to the capital city of Hanoi. Similar in concept to the approaches of *Kasprzyk et al*. [2009] and *Zeff et al*. [2014], we build and evaluate four rival problem framings of the Red River control problem; however, we not only vary the objectives and constraints included in each formulation, but also the mathematical quantification of those objectives. Building on insights from *Giuliani and Castelletti* [2016] in highlighting the importance of capturing multiple risk attitudes in problem framings for water management applications, we construct several alternative management objectives representing a range of stakeholder risk preferences from highly risk-averse (e.g., min-max objectives) to risk-neutral (e.g., expectation objectives). By optimizing Red River reservoir operations to alternative formulations encompassing a gradient of risk attitudes and reevaluating the resulting solutions from each formulation on the objectives from each of the other competing formulations, we seek to mitigate the unintended systematic biases that *Majone and Quade* [1980] caution analysts to guard against. We also perform visual diagnostics of the Pareto-approximate operating policies discovered under each problem formulation to better understand the effects of preference and framing on the resulting system behavior.

One of the difficulties of applying this constructive decision aiding approach is that traditional optimization methods may be limited in the type and scale of problems they can solve. For example, linear programming methods can only solve problems with linear objectives and constraints, while linear quadratic programming methods can only solve problems for linear systems with quadratic objective functions [*Yeh*, 1985]. In terms of scale, *Giuliani et al*. [2016b] note that commonly used stochastic dynamic programming methods are limited by several dimensional curses that confine the number of reservoirs whose operations can feasibly be optimized simultaneously, the number of exogenous variables, such as streamflow and precipitation, that can be used to condition reservoir release decisions, and the number of Pareto-optimal solutions that

can be discovered. Additionally, the system states must be discretized, the objective functions and constraints must be time-separable and the disturbance process must be uncorrelated in time [*Castelletti et al*., 2012a]. These latter constraints severely limit the type of objectives that can be formulated by traditional methods. For example, time-separability constraints make it impossible to reflect different risk attitudes with respect to different objectives within a single problem formulation, e.g., by calculating some objectives in expectation and others using a min-max formulation.

Fortunately, *Giuliani et al*. [2016b] show that these restrictions can be overcome with Evolutionary Multi-Objective Direct Policy Search (EMODPS), a simulation-optimization approach in which multiobjective evolutionary algorithms (MOEAs) are used to optimize the performance of multireservoir operating policies simulated over stochastic streamflows. Importantly, the objectives calculated over this simulation need not be time-separable, and the operating policies can be flexibly formulated to approximate any mathematical form. For example, nonlinear approximators can be used to describe the operating policies, allowing for adaptable, state-dependent operating rules. Consequently, EMODPS allows us to formulate complex problem formulations, improving our ability to accurately assess system performance. This added layer of problem complexity further motivates the need to test competing formulations of water resources optimization problems: first, because different combinations of performance measures may more effectively capture the stakeholders' objectives, and second, because multiple nonlinear objectives may interact in unpredictable ways, increasing the risk of unintended consequences. Fortunately, recent computational advancements in our ability to solve complex, multiobjective control problems [*Reed and Hadka*, 2014] have enabled a formal implementation of a rival framings approach to better account for problem formulation uncertainty.

Building off of foundational work by the Harvard Water Program [*Maass et al*., 1962] and others in highlighting the importance of utilizing multiple performance measures to evaluate system performance, in this study we exploit the computational power of EMODPS to optimize multireservoir operating policies for the Red River basin, described in section 2, under multiple problem formulations outlined in section 3. While many uncertainties surround reservoir operations, such as model, demand, and climate uncertainty, as well as nonstationarity in risk-preferences, we focus our analyses on stationary problem formulation uncertainty to isolate its effects. In section 4 we use visual diagnostics to assess the importance of this uncertainty by illustrating how policy operations designed under different formulations impact system performance. Through this analysis, we find several unintended consequences and unforeseen benefits of particular framings that have important implications for how the system can better manage extremes and conflicting multisectoral demands. Unfortunately, it is not always possible to know a priori what the effects of alternative objectives will be. For this reason, we conclude in section 5 with a discussion of the importance of applying constructive decision aiding processes to effectively manage the negative impacts and water conflicts associated with hydroclimatic variability and change. Future work will explore how problem formulation uncertainty compares with other sources of uncertainty in influencing overall system performance.

2. Red River Context

2.1. Basin Description

From its source in southern China to its mouth in the South China Sea, the Red River basin spans 169,000 km^2, 51% of which lies in Vietnam. As the second largest river basin in Vietnam, the Red River serves as a vital agricultural and economic resource to the developing nation. Recent reservoir construction in the system has significantly contributed to Vietnam's energy growth, with hydropower currently representing 46% of the country's total installed electric power capacity [*Asian Development Bank*, 2016]. These reservoirs have also enabled more secure and stable food production through irrigable agriculture, a key component in poverty alleviation, as 70% of the Vietnamese population is employed in agriculture, 76% of which is irrigated [*Nguyen et al*., 2002]. With cultivation and fisheries representing 58% and 29% of average water demand in the delta, respectively (Figure 1b), managing droughts is vital for Vietnam's food security.

Yet, while drought concerns during the dry season threaten the region's ability to provide sufficient water supply for agriculture and hydropower, large-scale floods during the monsoon season endanger the basin's infrastructure. The rapidly urbanizing Vietnamese capital of Hanoi lies in the Red River delta, where average annual flood damages have been estimated at 130 million USD [*Hansson and Ekenberg*, 2002]. Seeking to reduce the impacts of severe and frequent flooding, the Red River's second largest reservoir, Hoa Binh, was

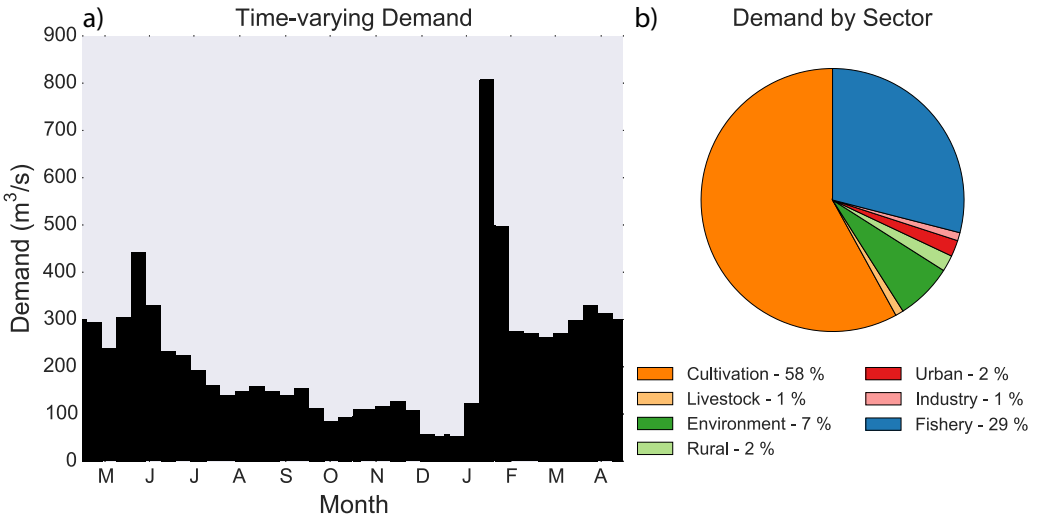

Figure 1. (a) Average water demand over time in the Red River delta and (b) its distribution across sectors, obtained from the Vietnamese Institute of Water Resources Planning (IWRP). There is a large spike in demand at the beginning of February for field flooding at the time of planting, illustrating why agriculture represents the largest source of demand (58%). The next most important sector is fisheries (29%), further highlighting the importance of water supply for the region's food security.

specifically designed to reduce the maximum observed flood peak at Hanoi from 14.8 to 13.3 m, just below the 13.4 m dike height [*Le Ngo et al*., 2007]. Flood protection requires maintaining low storage at Hoa Binh and the other system reservoirs during monsoonal months to ensure that there is sufficient storage capacity to capture large flood events. However, maintaining low storage in the reservoirs reduces hydropower production and the ability to supply water for irrigation. In this study, we investigate if improved multireservoir operations in Vietnam's Red River basin can better balance the multisectoral demands of agricultural water supply, energy production, and flood protection.

2.2. Model Description

Figure 2a shows the locations of the four largest reservoirs within the Red River basin whose operations we optimize. Figure 2b, reproduced from *Giuliani et al*. [2017], provides a more detailed schematic of how flows are simulated through the system. The two largest reservoirs, Son La (SL) and Hoa Binh (HB), are located in series along the Da River, which provides roughly half of the total system flow. Hoa Binh is the most important reservoir for flood protection since it is the last reservoir before Hanoi along the largest tributary. Parallel to Son La and Hoa Binh are the Thac Ba (TB) reservoir on the Chay River and the Tuyen Quang (TQ) reservoir on the Gam River. These reservoirs are much smaller in terms of storage and power capacity. Altogether, the four modeled reservoirs have a storage capacity of 22.67 billion m^3 and power capacity of 4782 MW. Table 1 lists the storage and power capacities of each reservoir individually.

We simulate flows through the Red River system using two submodels: (1) flows through the reservoirs and power plants and (2) flows through the delta. All data used to build the model are from the Ministry of Agriculture and Rural Development (MARD) of Vietnam and were collected during the Integrated and sustainable water Management of Red Thai Binh Rivers system in changing climate (IMRR) project (http://xake.elet.polimi.it/imrr/). In the first submodel, we estimate the volume of storage, s_t^k, in the k-th reservoir at time t using simple mass balance equations:

$$s_t^{SL} = s_{t-1}^{SL} + q_t^{Da} - r_t^{SL} - e_t^{SL} S(s_{t-1}^{SL}) \quad (1)$$

$$s_t^{HB} = s_{t-1}^{HB} + q_t^{Da,lat} + r_t^{SL} - r_t^{HB} - e_t^{HB} S(s_{t-1}^{HB}) \quad (2)$$

$$s_t^{TB} = s_{t-1}^{TB} + q_t^{Chay} - r_t^{TB} - e_t^{TB} S(s_{t-1}^{TB}) \quad (3)$$

$$s_t^{TQ} = s_{t-1}^{TQ} + q_t^{Gam} - r_t^{TQ} - e_t^{TQ} S(s_{t-1}^{TQ}) \quad (4)$$

where r_t^k is the actual release from the k-th reservoir in the time interval $[t-1, t)$; q_t^{Da}, q_t^{Chay}, and q_t^{Gam} are the water volumes from the Da, Chay, and Gam rivers flowing into the Son La, Thac Ba, and Tuyen Quang

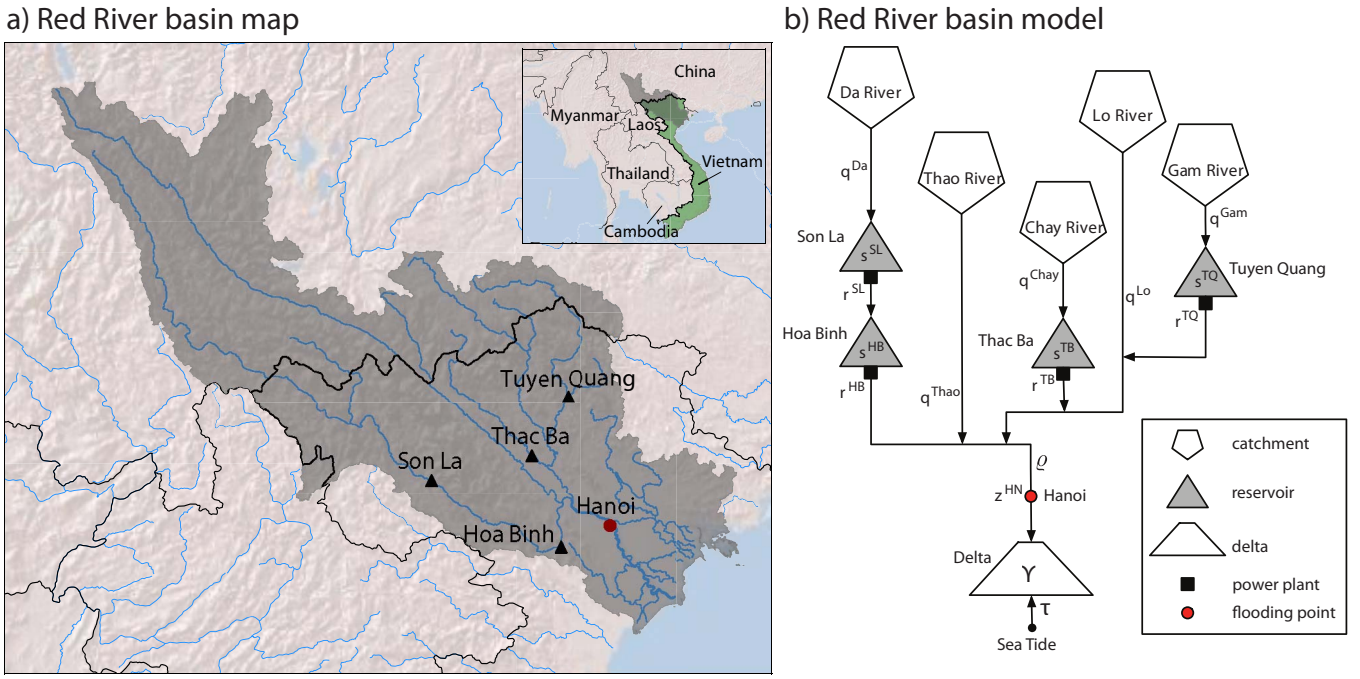

Figure 2. (a) Map of the Red River basin and (b) schematization of the main components of the Red River basin model (reproduced from *Giuliani et al.* [2017]). The inflows shown in Figure 2b are generated synthetically, the releases at each of the reservoirs are determined by the optimized operating policies, and subsequent flows through the delta are modeled by a dynamic emulator of a MIKE 11 simulation of the downstream hydraulics.

reservoirs, respectively, in this time interval; $q_t^{Da,lat}$ is the lateral inflow to the Da River between Son La and Hoa Binh during this time interval; e_t^k is the average evaporation rate from the k-th reservoir during this time interval and $S(s_{t-1}^k)$ is the surface area of k-th reservoir at time $t-1$ as a function of its storage level at time $t-1$. For each reservoir k, deterministic rates of e_t^k are assumed for each calendar day based on a 10 day moving average of the average historical evaporation rates from the nearest meteorological station over the period 1959–2011 [see *Bernardi et al.*, 2014 for more details]. In our notation, the value of the subscript indicates the time step at which each variable's value is deterministically known. The time step at which the release decision is made is 1 day; however, this volume of water is allocated hourly in the model assuming the operator optimally engages the turbines to maximize daily energy production. Following this assumption, we estimate the daily hydropower produced by the k-th reservoir's hydropower plant using the function $\eta_t^k = f(s_{t-1}^k, r_t^k)$, which is described by an artificial neural network (ANN) [see *Giuliani et al.*, 2016a for more details].

Because it is unrealistic and unsafe to assume no future streamflows will lie outside of those which have already been observed [*Thomas and Fiering*, 1962], we run the first submodel with synthetically generated hydrology. Compared to the limited 51 year historical record (1960–2010), these synthetic streamflows expand the range of hydrologic scenarios to which reservoir operations are optimized. In this study, we assume hydrologic stationarity in generating synthetic flows for the model simulations such that optimized operating policies represent baseline tradeoffs under our best perception of the current state of the world. Consequently, this study focuses solely on problem formulation uncertainty. In future work, we will explore the effects of uncertainty in the distribution of future hydrologic flows on the performance of these optimized policies.

Table 1. Storage and Power Capacities of Reservoirs in the Red River Basin Whose Operations Are Optimized

Reservoir	Storage in Bm^3 (% of Total)	Maximum Power Capacity (MW)
Son La	9.58 (42.3%)	2400
Hoa Binh	8.38 (37.0%)	1920
Thac Ba	2.81 (12.4%)	120
Tuyen Quang	1.90 (8.4%)	342

Here we synthetically generate correlated monthly streamflows on the five tributaries, q_t^{Da}, q_t^{Thao}, q_t^{Chay}, q_t^{Lo}, and q_t^{Gam} using the method of *Kirsch et al.* [2013]. This method uses Cholesky decomposition to preserve autocorrelation, and a simultaneous resampling of historical flows at

each site to preserve spatial correlation. We then disaggregate the synthetic monthly flows to daily flows using the method of *Nowak et al.* [2010], 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. Finally, we scale the lateral inflow between the Son La and Hoa Binh reservoirs, $q_t^{Da,lat}$, from q_t^{Da} assuming a constant flow per unit drainage area. Readers interested in a more detailed discussion and statistical validation of the synthetic streamflows can reference the supporting information.

For the delta submodel, we use a meta-model developed by *Dinh* [2015] to approximate a 1-D hydrodynamic model (MIKE 11) of the flow routing from the reservoirs to Hanoi and the irrigation districts. The meta-model employs an ANN to approximate the water volume in the irrigation canals, Υ_t, the water level at Hanoi, z_t^{HN}, and the supply deficit, D_t:

$$\Upsilon_t = f(\Upsilon_{t-1}, \varrho_t, W_t, \tau_{t-1}) \tag{5}$$

$$z_t^{HN} = f(z_{t-1}^{HN}, \varrho_t, \tau_{t-1}) \tag{6}$$

$$D_t = f(\varrho_t, W_t, \tau_{t-1}, \Upsilon_t) \tag{7}$$

where $\varrho_t = r_{t-1}^{HB} + r_{t-1}^{TB} + r_{t-1}^{TQ} + q_{t-1}^{Thao} + q_{t-1}^{Lo}$ is the total inflow to the canals assuming a 1 day travel time from the reservoirs and streamflow gauges of the Thao and Lo rivers to the delta; W_t is the time-dependent water demand; and τ_{t-1} is the previous day's tide. Using this meta-model reduces the computational demands of simulating 20 years of operations from a few days to a few seconds, making optimization computationally feasible. See *Dinh* [2015] for more details.

3. Methods

In this study, we evaluate four competing problem formulations of the Red River control problem. Because many of the problem formulations we explore are mathematically complex, we need a flexible optimization approach that does not require a specific problem structure. The Evolutionary Multi-Objective Direct Policy Search (EMODPS) framework [*Giuliani et al.*, 2016b] provides this flexibility. EMODPS is a parameterization-simulation-optimization approach [*Koutsoyiannis and Economou*, 2003] in which reservoir operating policies are parameterized within a given family of functions (e.g., piecewise linear functions, radial basis functions, etc.), simulated over a series of stochastic inputs, and then optimized to improve performance over multiple system objectives computed in the simulation. EMODPS utilizes nonlinear universal approximators to parameterize candidate operating policies and multiobjective evolutionary algorithms (MOEAs) to optimize their performance over the problem's conflicting objectives. Earlier work in the Red River basin by *Castelletti et al.* [2012b] found that EMODPS was able to discover operating policies for Hoa Binh that outperformed historical operations on every objective, and more recent work by *Giuliani et al.* [2016a] indicated it was capable of converging on a more challenging three-reservoir version of the model.

Here we advance the EMODPS framework with an additional diagnostic verification step. This step includes reevaluating the optimized policies on an out-of-sample set of stochastic inputs to ensure that they generalize well, and analyzing the policies themselves to understand how they operate on the system to achieve the given objectives. In summary, the primary steps in the EMODPS framework presented in this study are: (1) formulation of the system objectives, (2) formulation of reservoir operating policies as functions whose parameters are to be optimized, (3) multiobjective optimization of the policies, and (4) diagnostic verification of the optimized policies. We describe each of these steps in detail below.

3.1. Formulation of Objectives

A core goal and contribution of this work is to better understand the nature of the operational tradeoffs across the Red River system's three primary functions: flood protection, hydropower production, and agricultural water supply. However, as noted in the introduction, translating each of the system objectives into quantitative performance measures is not straightforward. Consequently, we explore four rival problem framings that capture a range of stakeholder attitudes toward risk from highly risk-averse to risk-neutral. We take a multiobjective optimization approach since collapsing these objectives into a single economic performance measure weighting the different objectives may lead to a single objective unexpectedly dominating the system performance [*Arrow*, 1950; *Kasprzyk et al.*, 2015]. This could be especially concerning if

the estimates of the costs and benefits associated with flood damages, hydropower revenue, and agricultural losses during drought are highly uncertain and nonstationary [*Dittrich et al.*, 2016; *Smith et al.*, 2017]. For these reasons, in each formulation, we quantify the three key stakeholder objectives using unmonetized measures of performance, but recommend that economic estimates be incorporated a posteriori to aid decision makers in choosing among alternative operating policies.

The four candidate formulations we explore in this study are as follows: (1) Worst Case (WC), (2) Worst First Percentile (WP1), (3) Expected Value (EV), and (4) Expected Value & Standard Deviation of Hydropower (EV&SD_H). In each formulation, operating policies are simulated over N ensemble members of T years of synthetically generated streamflows, with N and T varying by formulation. In all formulations, each T-year simulation begins on 1 May, the first day of the monsoon season, and initial conditions for 31 April must be specified for the storages at the four reservoirs, $\{s_0^{SL}, s_0^{HB}, s_0^{TB}, s_0^{TQ}\}$, the water level at Hanoi, z_0^{HN}, the water volume in the canals, Υ_0, the flow to the delta, ϱ_0, and the total system inflow, q_0^{TOT}. In the WC formulation, objectives are calculated over $N = 50$ ensemble members of length $T = 20$ years (i.e., 50 unique 20 year streamflow records). Since this is a fairly long simulation length, performance is relatively insensitive to initial conditions, so constant, typical values for 31 April are assumed. In the WP1, EV and EV&SD_H formulations, however, objectives are calculated over $N = 1000$ ensemble members of length $T = 1$ year (i.e., 1000 unique 1 year streamflow records). In order to better sample interannual variability under these shorter simulations, initial conditions are randomized for each ensemble member by sampling joint conditions on 31 April from 10,000 year simulations of the optimal policies from the WC formulation. Table 2 summarizes the characteristics of the simulations over which policies are optimized for each formulation.

When formulating candidate objective functions for optimization, a single performance statistic across the N ensembles of T-year simulations must be quantified mathematically. We calculate the d-th objective, J_d, according to equation (8):

$$J_d = \Psi_{i\in(1,\ldots,N)}[\Phi_{t\in(1,\ldots,365T)}[g_d(t,i)]] \tag{8}$$

where $g_d(t, i)$ is the value of the d-th objective on day t of the i-th ensemble member, Φ is an operator for the aggregation of $g_d(t, i)$ over time, such as the sum ($\sum$), and Ψ is a statistic used to filter the noise across ensemble members, such as the expected value ($\mathbb{E}$). It is through these key variables, $g_d(t, i)$, Φ, Ψ, N, and T, that the problem formulation can vary to reflect different risk attitudes, e.g., by changing how objectives are aggregated over time and filtered across noise, as well as the time horizon over which they are calculated [*Soncini-Sessa et al.*, 2007]. While these variables change across the four formulations we explore, the general form of the highest dimensional multiobjective optimization problem can be summarized by equations (9)–(11) below:

$$\theta^* = \text{argmin}_\theta J(\theta) \tag{9}$$

where

$$J = \begin{bmatrix} -J_{Hydro}(\theta) \\ J_{Deficit^2}(\theta) \\ J_{Flood}(\theta) \\ J_{Recovery}(\theta) \\ J_{Hydro\ Std}(\theta) \end{bmatrix} \tag{10}$$

Table 2. Characteristics of Simulations Over Which Policies Are Optimized for Each Formulation

Formulation	Initial Conditions	Ensemble Size (N)	Years/ Ensemble (T)
Worst Case (WC)	Constant, average conditions	50	20
Worst First Percentile (WP1)	Randomly sampled from simulation of WC policies over 10,000 years	1000	1
Expected Value (EV)	Randomly sampled from simulation of WC policies over 10,000 years	1000	1
Expected Value & Standard Deviation of Hydropower (EV&SD_H)	Randomly sampled from simulation of WC policies over 10,000 years	1000	1

Table 3. Objectives and Constraints Included in Optimization Under Each Problem Formulation[a]

Formulation	Objectives	Constraints
Worst Case (WC)	J^{WC}_{Hydro}, $J^{WC}_{Deficit^2}$, J^{WC}_{Flood}	–
Worst First Percentile (WP1)	J^{WP1}_{Hydro}, $J^{WP1}_{Deficit^2}$, J^{WP1}_{Flood}, $J^{WP1}_{Recovery}$	$J^{WP1}_{Flood} \leq 2.15$ m
Expected Value (EV)	J^{EV}_{Hydro}, $J^{EV}_{Deficit^2}$, J^{WP1}_{Flood}, $J^{EV}_{Recovery}$	$J^{WP1}_{Flood} \leq 2.15$ m
Expected Value & Standard Deviation of Hydropower (EV&SD_H)	J^{EV}_{Hydro}, $J^{EV}_{Deficit^2}$, J^{WP1}_{Flood}, $J^{EV}_{Recovery}$, $J^{EV\,\&\,SD_H}_{Hydro\,Std}$	$J^{WP1}_{Flood} \leq 2.15$ m

[a]See sections 3.1.1–3.1.4 for further explanation.

subject to

$$J_{Flood}(\theta) \leq C \tag{11}$$

where θ is a vector of decision variables describing the operating policies defined in section 3.2 and C is a constraint on the flooding objective defined for each formulation in sections 3.1.1–3.1.4. Table 3 summarizes which objectives and constraints are included in each formulation. As indicated by the superscripts, some of the objectives are the same across formulations, while others are not. Mathematical descriptions of each of the objectives under each formulation, including $g_d(t, i)$, Φ and Ψ, are provided in Appendix A, while summary text descriptions and our rationale for each candidate formulation are provided in sections 3.1.1–3.1.4 below.

3.1.1. Worst Case (WC) Formulation

The WC formulation assumes a highly risk-averse operator concerned with minimizing the worst case performance of the hydropower, flooding, and water supply objectives across an ensemble of potential conditions, similar to prior published studies by *Orlovski et al*. [1984] and *Soncini-Sessa et al*. [1990]. This formulation was designed based on the desire of the MARD to formulate conservative operating policies, particularly with respect to flooding.

In the WC formulation, the first objective, J^{WC}_{Hydro}, seeks to maximize hydropower production based on the desire of the Vietnamese Ministry of Industry and Trade (MOIT) to generate as much hydropower as possible in order to minimize costs of production from thermal plants and import. Here, we calculate average hydropower production over the synthetically generated 20 year streamflow sequences ($\Phi = \mathbb{E}_{365T}$) and minimize the worst case average production across the 50 member ensemble ($\Psi = \min_N$). Simulations of 20 years are chosen to provide estimates of production over a typical planning period, while the worst case across 50 simulations of 20 years is minimized to ensure reasonable performance in even the worst case potential planning period. In the Red River system, we optimize hydropower production rather than revenue because the Vietnamese electricity market is regulated by the Government and energy is sold at a fixed rate. Since the price is fixed, maximizing production is equivalent to maximizing the revenue from production [*Castelletti et al*., 2012b]. While unexplored here, uncertainty in how best to formulate this objective could be considered in an additional rival framing.

The second objective, $J^{WC}_{Deficit^2}$, seeks to minimize the squared water supply deficit. As with hydropower, the average daily squared deficit is calculated over every 20 year ensemble member ($\Phi = \mathbb{E}_{365T}$) and we minimize the maximum of these averages across the 50 ensemble members ($\Psi = \max_N$). The daily deficit is squared to numerically favor several small deficits over a small number of large deficits. This objective was accepted by the MARD through the IMRR project in 2013.

The final objective of this formulation, J^{WC}_{Flood}, seeks to minimize flood damages. In each of the 20 year simulations, we approximate expected flood damages by the penalty function shown in Figure 3 and minimize the maximum expected damages across the 50 member ensemble ($\Phi = \mathbb{E}_{365T}$ and $\Psi = \max_N$). Based on alarm levels elicited from stakeholders, water levels between 6 and 11.25 m are penalized minimally by a linearly increasing function of depth, while water levels above 11.25 m are more harshly penalized by a fourth-order polynomial to reduce the probability of overtopping the dikes at 13.4 m [*Giuliani et al*., 2016a]. This damage function was suggested by the Vietnamese Central Committee for Flood and Storm Control (CCFSC).

Minimizing flooding damages is a common approach to optimizing reservoir operations for flood control [*Windsor*, 1973; *Needham et al*., 2000; *Lund*, 2002; *Malekmohammadi et al*., 2009]. In past studies of flooding in the Red River basin, *Vinh Hung et al*. [2007] estimated damages using a 2-D hydrodynamic model of the delta mapping water levels to inundated area, while *De Kort and Booij* [2007] used past flood recovery costs

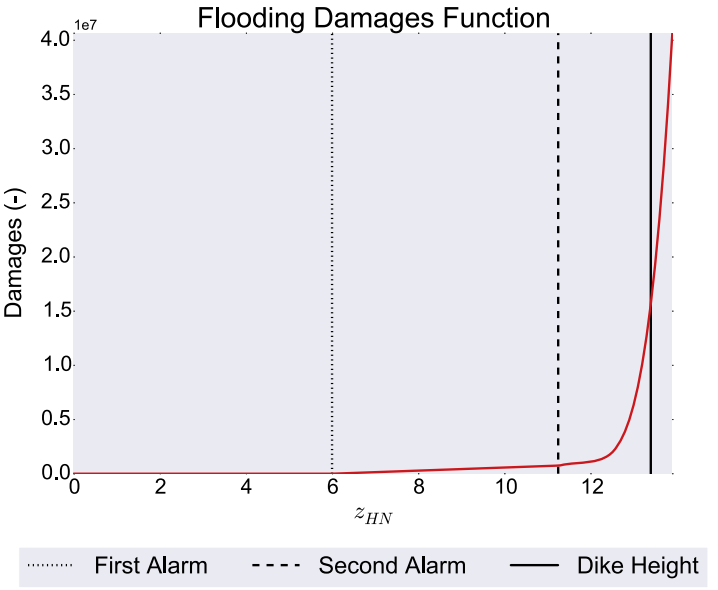

Figure 3. Flood penalty function used to approximate damages at Hanoi. Below 6 m (dotted black line) there are assumed to be no damages. Between 6 and 11.25 m (dashed black line), damages are assumed to be minor and linearly increasing with depth, but above 11.25 m they become severe and are modeled by a fourth-order polynomial. This shape is intended to keep water levels from breaching the dikes at 13.4 m (solid black line). The shape of the damage function and the alarm levels were elicited from stakeholders.

and flow rates to estimate a damage curve. In contrast, *Castelletti et al.* [2012b] concluded that minimizing estimated damages was not appropriate in the Red River because the delta is constantly changing as a result of dike breaching events and urban development.

When estimating actual damages is difficult or they are nonstationary as is the case here, it is common to instead create a flood penalty function that harshly penalizes high water levels [see e.g., *Orlovski et al.*, 1984; *Needham et al.*, 2000]. This is the intent of the above function. *Castelletti et al.* [2012b] take a similar approach by minimizing the average squared excess of 9.5 m at Hanoi, an alarm level chosen from *Hansson and Ekenberg* [2002]. The excesses are squared to reduce the total force on the levee, the driver of collapse, which increases with the square of the water level. Similarly, *Le Ngo et al.* [2007] minimize a weighted sum of squared maximum water levels at Hanoi and squared deviations of the Hoa Binh reservoir level from its maximum each flood season. The piecewise fourth-order polynomial used here is intended to be extremely risk-averse.

3.1.2. Worst First Percentile (WP1) Formulation

In prior work in the Red River Basin, *Giuliani et al.* [2017] solved the WC formulation of the problem as both a challenging computational benchmark application and to provide an initial understanding of the system's multisectoral tradeoffs. Subsequent to this effort, the authors reevaluated the policies derived from the WC formulation on a larger set of out-of-sample streamflows and observed that the policies did not generalize well. For this reason, we explore an alternative risk-averse formulation here in which we minimize the worst first percentile across a 1000 member ensemble of 1 year simulations rather than the absolute worst across a 50 member ensemble of 20 year simulations. The motivations for this are twofold: (1) the worst first percentile should be more stable than the worst case, as the worst case has a higher sampling variance and may be unbounded [*Stedinger et al.*, 1993], and (2) aggregating objectives over 1 year simulations and minimizing the worst first percentile across a 1000 member ensemble may better capture interannual variability. Aggregating objectives over 20 year simulations and minimizing the worst case across a 50 member ensemble as in the WC formulation may mask particularly bad years if several good years are also included.

In the WP1 formulation, we compute hydropower production and the squared deficit in expectation within each ensemble member's 1 year simulation ($\Phi = \mathbb{E}_{365T}$), just as in the 20 year simulations of the WC formulation. However, we then calculate J^{WP1}_{Hydro} and $J^{WP1}_{Deficit^2}$ as the worst first percentile of these averages across the 1000 member ensemble rather than the absolute worst ($\Psi = \text{quantile}_N\{\Phi, 0.01\}$ for J^{WP1}_{Hydro} and $\Psi = \text{quantile}_N\{\Phi, 0.99\}$ for $J^{WP1}_{Deficit^2}$). In the WP1 formulation, we also reframe how the flooding objective is calculated within each 1 year simulation so that the objective values are more semantically meaningful. While the fourth-order polynomial is intended to be very conservative with respect to flooding, the value of the damage function is hard to comprehend since it maps water levels to a dimensionless number, not a monetary value. If stakeholders are trying to weigh the trade-off between two solutions, it is unclear how much better or worse one solution does with respect to the other in terms of flooding based on their objective values.

Motivated by *Hashimoto et al*. [1982], we partition the flooding objective into two components: resilience, $J^{WP1}_{Recovery}$, and vulnerability, J^{WP1}_{Flood}.

In this study, we quantify flood resilience within each 1 year simulation using its inverse, measured as the average time to recovery after the water level at Hanoi exceeds 6 m ($\Phi = \mathbb{E}_{365T}$). We quantify flood vulnerability as the maximum annual water level in excess of 11.25 m ($\Phi = \max_{365T}$). These two thresholds are based on the same cutoffs used to define the piecewise polynomial function used to estimate flood damages in the worst case formulation (see Figure 3). Since the WP1 formulation minimizes the worst first percentile across a 1000 member ensemble of 1 year simulations ($\Psi = \text{quantile}_N\{\Phi, 0.99\}$ for both J^{WP1}_{Flood} and $J^{WP1}_{Recovery}$), the flood vulnerability objective, J^{WP1}_{Flood}, is equivalent to minimizing the amount by which the 100 year flood exceeds 11.25 m. Unlike the WC flood damages objective, which we do not constrain because it is unclear what is an acceptable level of dimensionless damages, we constrain the WP1 flood vulnerability objective to be less than 2.15 m (the difference between the second alarm level and the dike height) under the assumption that stakeholders would like to be protected to at least the 100 year flood level. The resilience objective, $J^{WP1}_{Recovery}$, is intended to keep water levels at Hanoi persistently low in order to reduce the sustained pressure on the dikes. It is expected that these two flood objectives will conflict; maintaining low water levels at Hanoi may require higher storages in the reservoirs, reducing their capacity to capture large floods, putting Hanoi at risk of higher maximum water levels.

3.1.3. Expected Value (EV) Formulation

The WC and WP1 formulations both assume a risk-averse operator who is concerned with the tails of the distribution of each objective. However, optimizing to the tails often requires sacrifices in the mean [*Beyer and Sendhoff*, 2007]. In the EV formulation, we assume a risk neutral operator who is primarily concerned with average performance, representing common practice in water resources optimization problems (for examples, see reviews by *Yakowitz* [1982]; *Yeh* [1985]; *Labadie* [2004], and sources cited therein). Under the EV formulation, we quantify J^{EV}_{Hydro}, $J^{EV}_{Deficit^2}$, and $J^{EV}_{Recovery}$ as the expected annual hydropower production, squared deficit, and recovery time for water levels over 6 m at Hanoi, respectively, calculating their averages across a 1000 member ensemble of 1 year simulations ($\Phi = \mathbb{E}_{365T}$ and $\Psi = \mathbb{E}_N$). We do not change the flood vulnerability objective and constraint from the WP1 formulation, though, as flooding is not a concern in the average year; it is only the extremes that put the city of Hanoi at risk and consequently need to be minimized.

3.1.4. Expected Value and Standard Deviation of Hydropower (EV&SD$_H$) Formulation

The final formulation we test can be viewed as a compromise between the risk-averse WP1 formulation and the risk-neutral EV formulation, with a specific focus on the interannual variability of hydropower production. In the classical robust optimization literature, it has long been recognized that there is often a direct conflict between the mean and variance of stochastic performance measures [*Taguchi*, 1986]. Knowing this, operators may be willing to trade off exceptionally high years of hydropower production if operations can be discovered that reduce their exposure to drought-driven losses in production. In the water resources literature, these concerns have been addressed by including measures of variability in addition to expectation, either as an additional objective in a multiobjective optimization problem [*Kawachi and Maeda*, 2004; *Reed and Kasprzyk*, 2009], as a constraint [*Kasprzyk et al*., 2012], or as part of a weighted single objective function [*Watkins and McKinney*, 1997; *Ray et al*., 2013]. Here we take the first approach and explicitly quantify the trade-off between maximizing mean hydropower performance and minimizing the variability about that mean by adding an objective to the EV formulation, $J^{EV\,\&\,SD_H}_{Hydro\,Std}$, to minimize the standard deviation in average annual hydropower production ($\Phi = \mathbb{E}_{365T}$ and $\Psi = \text{std}_N$). All other objectives and constraints are the same as in the EV formulation.

Table 4 provides a summary of the objective calculations from each formulation. For a more detailed, mathematical description of the objectives from each formulation, see Appendix A.

3.2. Formulation of Operating Policies

In order to optimize the complex objective functions defined for each of the four rival framings described in sections 3.1.1–3.1.4, we need to specify an operating policy for each of the reservoirs. In this study, we apply flexible, nonlinear functions that do not require an assumed mathematical form but can universally approximate a variety of functional shapes. Two such common functions are artificial neural networks (ANNs) and radial basis functions (RBFs). *Giuliani et al*. [2016b] compare operating policies optimized for Hoa Binh with EMODPS using ANNs and RBFs and find that ANNs tend to overfit to the stochastic simulations they are

Table 4. Within-Ensemble Aggregators and Across-Ensemble Noise Filters for Objective Calculations Under Each Formulation

Objective	Φ (Within-Ensemble Aggregator)	Ψ (Across-Ensemble Noise Filter)
J^{WC}_{Hydro}	$\mathbb{E}_{365T}$	$\min_N$
J^{WP1}_{Hydro}	$\mathbb{E}_{365T}$	quantile{Φ, 0.01}
J^{EV}_{Hydro}	$\mathbb{E}_{365T}$	$\mathbb{E}_N$
$J^{WC}_{Deficit^2}$	$\mathbb{E}_{365T}$	$\max_N$
$J^{WP1}_{Deficit^2}$	$\mathbb{E}_{365T}$	quantile{Φ, 0.99}
$J^{EV}_{Deficit^2}$	$\mathbb{E}_{365T}$	$\mathbb{E}_N$
J^{WC}_{Flood}	$\mathbb{E}_{365T}$	$\max_N$
J^{WP1}_{Flood}	$\max_{365T}$	quantile{Φ, 0.99}
$J^{WP1}_{Recovery}$	$\mathbb{E}_{365T}$	quantile{Φ, 0.99}
$J^{EV}_{Recovery}$	$\mathbb{E}_{365T}$	$\mathbb{E}_N$
$J^{EV\ \&\ SD_H}_{Hydro\ Std}$	$\mathbb{E}_{365T}$	std_N

trained on, generalizing less well when reevaluated out-of-sample. For this reason, we parameterize the operating policies of the four reservoirs with RBFs.

As shown in equation (12), the RBF-based representation of operational policies prescribe releases, u^k_t (normalized on [0,1]), from the k-th reservoir at time t as a function of B time-varying inputs, x_t (normalized on [0,1]):

$$u^k_t = \sum_{i=1}^{A} w^k_i \exp\left(-\sum_{j=1}^{B} \frac{((x_t)_j - c_{j,i})^2}{b^2_{j,i}}\right) \quad (12)$$

where $(x_t)_j$ is the normalized value of the j-th input at time t, A is the number of RBFs, w^k_i is the weight of the i-th RBF associated with the k-th reservoir, and $c_{j,i}$ and $b_{j,i}$ are the centers and radii, respectively, of the i-th RBF associated with the j-th input. Due to physical constraints, the actual release from reservoir k at the end of the time interval $[t,t+1)$, r^k_{t+1}, is not always the same as the unnormalized policy-prescribed release. If there is insufficient water to meet the unnormalized value of u^k_t, only the available water is released, and if there is insufficient storage capacity, s^k_{cap}, to allow only releasing the unnormalized value of u^k_t, the excess is spilled.

The centers, radii, and weights of the RBF policies compose the decision variables, θ, optimized by the MOEA (see equation (9)):

$$\theta = \begin{bmatrix} c_{i,j} \\ b_{i,j} \\ w^k_i \end{bmatrix} \text{with } i=\{1,\ldots,A\},\ j=\{1,\ldots,B\} \text{ and } k=\{1,\ldots,M\} \quad (13)$$

where $c_{i,j} \in [-1,1]$, $b_{i,j} \in [0,1]$, and $w^k_i \in [0,1]$ with $\sum_{i=1}^{A} w^k_i = 1\ \forall\ k$. For M outputs (reservoirs), this corresponds to $A(M+2B)$ decision variables. We model the releases at $M=4$ reservoirs using $A=11$ RBFs and $B=6$ inputs, where the inputs are the storages at each reservoir, the total system inflow, and the day of the year: $x_t = \{s^{SL}_t, s^{HB}_t, s^{TB}_t, s^{TQ}_t, q^{TOT}_t, t\}$ where $q^{TOT}_t = q^{Da}_t + q^{Da,lat}_t + q^{Thao}_t + q^{Chay}_t + q^{Gam}_t$. This represents a total of 176 decision variables.

3.3. Multiobjective Optimization

Since we perform multiobjective optimization in this study, we do not seek a single optimal solution for each problem formulation but a set of nondominated solutions, also called the Pareto optimal set [*Pareto*, 1896]. Within this set, performance in any component objective can only be improved by degrading performance in one or more of the remaining objectives. Over the last decade, multiobjective optimization problems of increasing complexity have been explored using multiobjective evolutionary algorithms (MOEAs). MOEAs are heuristic algorithms that evolve approximations to the Pareto optimal set through search processes that exploit global probabilistic search operators for mating, mutation, and selection [*Reed et al.*, 2013]. We use the Multi-Master Borg MOEA [*Hadka and Reed*, 2015] to optimize the operating policies of the four reservoirs in the Red River basin. Multi-Master Borg is a hierarchical parallelization of the Borg MOEA [*Hadka and Reed*, 2013], which has been shown to improve the reliability of attaining high-quality approximations to the Pareto optimal set for challenging real-world problems [*Hadka and Reed*, 2015; *Giuliani et al.*, 2017]. The Multi-Master Borg consists of multiple master-worker implementations of the Borg MOEA, called islands, which coevolve through the aid of a controller that keeps a global archive of the best solutions across all of the islands.

We use Multi-Master Borg with 16 islands and run 5 random algorithm trials, or seeds, of 400,000 function evaluations per island. Visual inspection of search progress indicated that this was sufficient, as progress had reached an asymptote of diminishing returns with little variability across the five random seeds. The epsilon dominance archiving used in Borg requires that users specify levels of precision for each objective

Table 5. Epsilons Used for Multiobjective Optimization Under Each Problem Formulation

Objective	WC Formulation	WP1 Formulation	EV Formulation	EV&SD_H Formulation
J_{Hydro}	0.1	0.5	0.5	0.5
$J_{Deficit^2}$	5.0	25.0	25.0	25.0
J_{Flood}	275.0	0.05	0.05	0.05
$J_{Recovery}$		0.5	0.5	0.5
$J_{Hydro\ Std}$				0.05

below which they are indifferent to differences in performance. Table 5 shows the values of epsilon (or significant precisions) used for each objective in this study. *Giuliani et al.* [2017] optimized the worst case formulation on the Texas Advanced Computing Center (TACC) Stampede Cluster (https://www.tacc.utexas.edu/stampede/) using 512 cores per island and 400,000 computational hours, and we optimized the remaining formulations on the Blue Waters supercomputer (http://www.ncsa.illinois.edu/enabling/bluewaters) using 1024 cores per island and a total of 1.7 million computational hours. For each formulation, we obtained approximate Pareto sets by combining and re-sorting the best solution sets found by each seed.

3.4. Diagnostic Verification of Optimized Policies

The final step we have added to the EMODPS framework in this paper is the diagnostic verification of the optimized policies. We evaluate the performance of the control policies in two ways: (1) by reevaluating their performance over out-of-sample streamflow ensembles, and (2) by analyzing their multireservoir management behavior and the system dynamics that result from operating with policies that emphasize different objective preferences from the rival problem formulations. In the first step, we reevaluate all of the policies over a second ensemble of stochastic streamflows that is 100 times larger than the ensemble used during optimization. This corresponds to a 5000 member ensemble of 20 year simulations for the WC formulation, and a 100,000 member ensemble of 1 year simulations for all other formulations. If the solutions achieve similar objective values in the reevaluation as in optimization, then both the objectives and policies are stable, so we can trust our representation of policy performance.

In this study, we also reevaluate the solutions from each problem formulation on the objectives from each of the other formulations using the sets of streamflows from both the optimization and validation. This allows us to visualize the regrets associated with risk-averse versus risk-neutral objectives to determine the costs in expectation of optimizing to the worst case and vice versa. We can also see if the stability of a particular objective depends on whether or not that objective was included in the optimization. This can be used to diagnose whether poor performances in reevaluation are due to the policies being overfit, or to the objectives themselves being inherently unstable.

Finally, the second step of the diagnostic verification is to analyze the behavior and consequences of operating with the optimized policies. In this study, we select a few solutions from different formulations to see how operations vary as a function of both preference and formulation. Analyzing the operations and state behavior of the system opens the black box of the policy function, providing insights into how the policies are able to achieve the objective values that they do.

4. Results and Discussion

Here we present the results of the EMODPS policy optimization and diagnostic verification. In section 4.1, we show the multiobjective trade-offs that emerge from the Pareto-approximate solutions discovered under each of the four candidate problem formulations (see Tables (2–4)). In section 4.2, we assess how well these solutions generalize on out of sample streamflows, and reevaluate the competing formulations in each other's spaces (i.e., we resimulate the control policies found under each problem formulation to calculate their performance on the objectives from all of the other formulations). Lastly, in section 4.3 we illustrate how these control policies affect downstream flood dynamics at Hanoi.

4.1. Rival Representations of Red River Trade-Offs

The best known approximations of the Pareto optimal sets discovered for each of the four candidate problem formulations of the Red River test case are shown using parallel axes plots in Figures 4a–4d. In these plots, each shaded line corresponds to an operating policy for the system's four reservoirs that intersects each vertical axis at the value it achieves for the objective that axis represents. Solutions from the WC

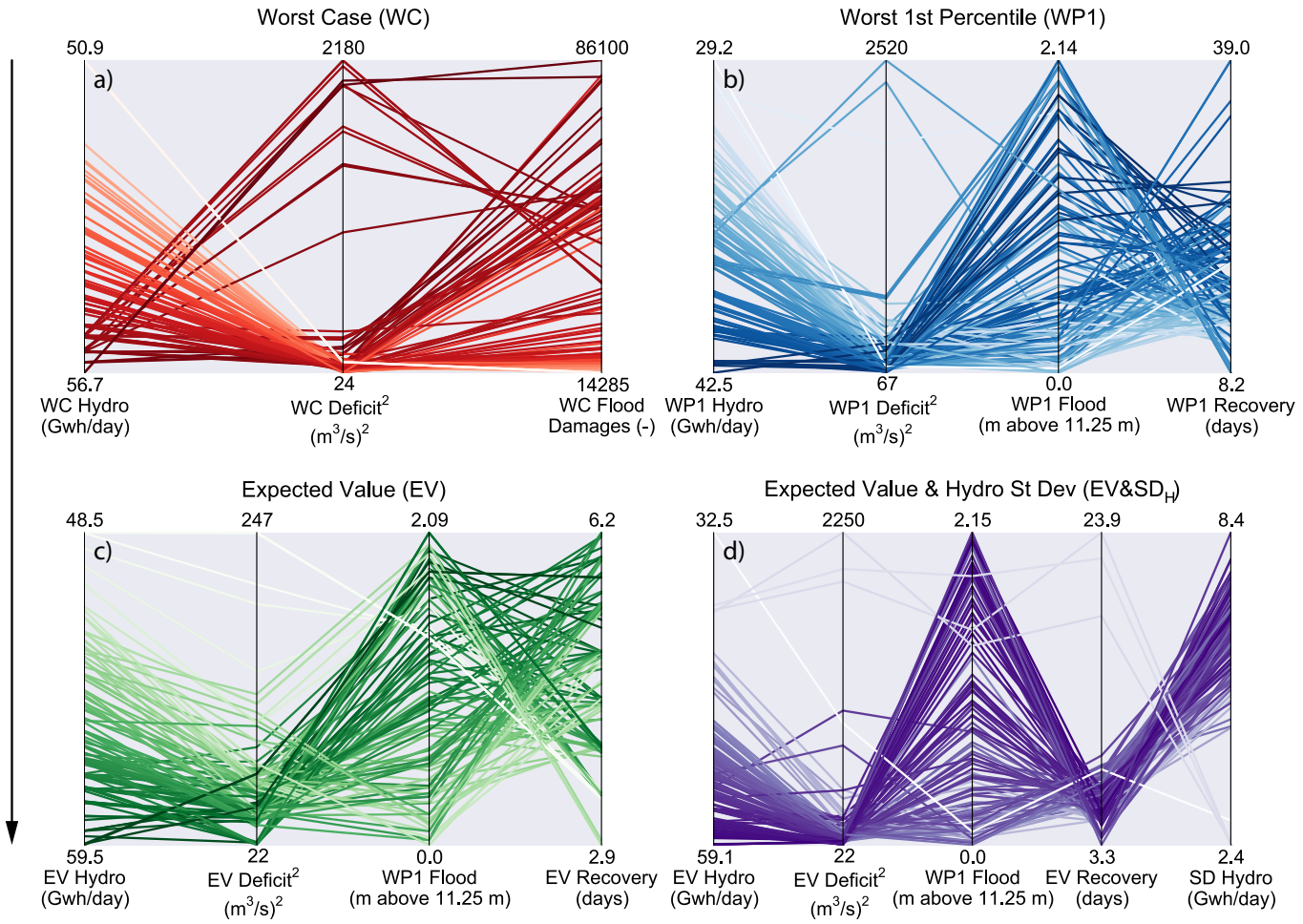

Figure 4. Approximate Pareto sets from the (a) WC formulation, (b) WP1 formulation, (c) EV formulation, and (d) EV&SD$_H$ formulation. Each axis represents a different objective from that formulation and each shaded line a solution in the approximate Pareto set. All lines are shaded by their performance on the hydropower objective, with darker shades representing better performance, and all axes are oriented such that the optimal direction is down. An ideal solution would therefore be a dark horizontal line across the bottom of the axes.

formulation are shown in red (plot a), from the WP1 formulation in blue (plot b), the EV formulation in green (plot c), and the EV&SD$_H$ formulation in purple (plot d). We use this color scheme to distinguish the candidate Red River problem formulations in all subsequent figures. In Figures 4a–4d, all the vertical axes have been oriented such that the optimal direction is downward. All lines have been shaded according to their performance on the hydropower objective, with darker shades representing greater production. Consequently, theoretical ideal solutions in each of the spaces plotted in Figure 4 would be dark shaded horizontal lines intersecting the bottom of each axis.

In parallel axes plots, intersecting lines between pairs of vertical axes designate trade-offs between those two objectives, as superior performance in one objective comes at the expense of inferior performance in another. In Figures 4a–4d, one can also observe trade-offs between the hydropower objective and objectives oriented on nonadjacent axes through shading. For visual clarity, we have thinned the four Pareto approximation sets illustrated in Figure 4 by re-sorting them with larger epsilons to attain representative sets of approximately 100 solutions in each plot that fully span the trade-offs discovered in this study.

Across Figures 4a–4d, the major trade-off of note is between hydropower and flooding, which can be seen by the inversion of the color gradients along these two axes. This conflict results because high storages favor hydropower production, while low storages favor flood protection. There is also a weak, but nonlinear trade-off between the squared water supply deficit and flooding, as well as between the squared water supply deficit and hydropower, which can be seen by the crossing diagonal lines between these adjacent axes across all formulations. For the formulations that include both the flood vulnerability and resilience

objectives (WP1 in plot b, EV in plot c and EV&SD$_H$ in plot d), crossing diagonal lines between these axes indicate that there is a strong trade-off between these two objectives. This suggests that in order to reduce maximum flood levels, moderately high flood levels must be maintained, resulting in more sustained pressure on the dikes.

Interestingly, the shapes of the trade-offs that emerge for the WP1 (plot b) and EV formulations (plot c) are similar, suggesting that these conflicts are not quantile-dependent. However, this does not imply that there is not a trade-off between average performance and the stability of performance. As can be seen by the inversion of colors along the axes for expected hydropower and standard deviation of hydropower in the EV&SD$_H$ formulation (plot d), these objectives strongly conflict. In particular, the solutions with the lowest standard deviation in annual hydropower production have similar average hydropower production to the worst first percentile hydropower production observed in the WP1 formulation. This severe degradation in average performance occurs with the squared deficit and recovery time objectives as well.

4.2. Verification of Control Policies

As summarized in section 3.4, our first diagnostic verification step is to reevaluate the solutions from each of the problem formulations in the objective spaces of all of the other formulations using both the synthetic streamflow ensembles to which they were optimized and an out-of-sample validation set with 100 times as many ensemble members. The results of the reevaluation are shown in Figure 5, where each plot represents a different objective. The first row of Figure 5 shows the WC objectives, the second row the WP1 objectives, and the bottom row the EV and EV&SD$_H$ objectives. Within each plot, points representing the multireservoir control policies are positioned along the *x* axis at their objective values achieved in optimization, and along the *y* axis at their objective values in validation. Each plot is oriented such that the lower left corner represents the most favorable direction. Solutions that achieve similar values in optimization and validation will fall near the 1:1 line, shown by a black, dashed line. If solutions lie

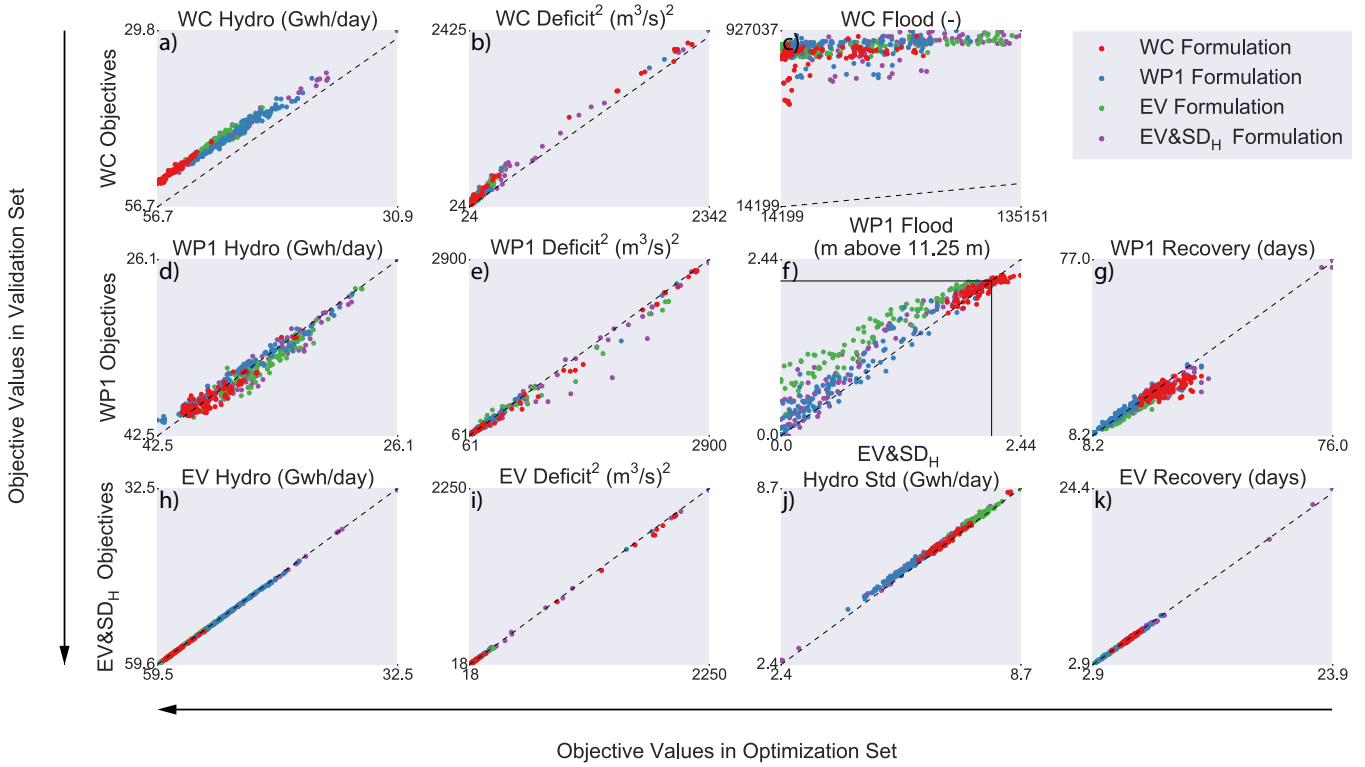

Figure 5. Validation of Pareto-approximate solutions from each formulation. (a–c) Objectives from WC formulation, (d–g) WP1 formulation, (h–k) EV and EV&SD_H formulation. In each plot, each dot represents a different control policy positioned along the *x* axis at its objective value over the optimization set of streamflows, and along the *y* axis at its value over an out-of-sample validation set with 100 times as many ensemble members. Solutions with stable performance between optimization and validation lie near the black dashed 1:1 line. All plots are arranged such that the optimal direction is toward the lower left corner. The black solid line in Figure 5f represents the dike height at Hanoi.

above this line, their performance degraded in reevaluation, and if they lie below, their performance improved.

Figure 5 provides several insights into the stability of each of the problem formulations, as well as their inherent biases, some of which were intended and others of which were not. Beginning with the WC objectives in the first row (plots a–c), the most prominent observation is the instability of these objectives, as nearly all of the solutions lie above the dashed 1:1 line, indicating degrading performance in reevaluation. The degradation is particularly bad on the flood damages objective (plot c) due to the fourth-order polynomial used to approximate damages when the water level at Hanoi exceeds 11.25 m. The fact that all of the solutions degrade similarly whether or not they were optimized under this problem formulation indicates that the degradation is not due to overfitting the radial basis functions that define the control policies but instead due to the worst case formulation of the flood objective itself. This is not surprising, as the worst case is often unbounded and therefore likely to worsen as the sample size increases. This may not be problematic, for example, in the case of the hydropower objective (plot a) where the ordering of the solutions from most favorable to least favorable does not considerably change in reevaluation. However, Figure 5c indicates that this is not the case on the WC Flood objective, as some solutions that do relatively poorly over the optimization set do relatively well over the validation set, suggesting that the original representation of the trade-offs from optimization may not be accurate.

This difference in the magnitude of degradation across objectives is likely due to their distributions. Hydropower has a bounded minimum production, and the squared deficit a bounded maximum. Flooding, however, is unbounded, and the fourth order polynomial used to estimate damages in this study results in an extremely fat tail. Optimizing the worst case of an unbounded, nonlinear objective is particularly difficult, as its value degrades severely with increasing sample sizes, resulting in greater noise than signal. Compounding this difficulty is the nonuniqueness of calculating damages in expectation; a solution with frequent small floods may have similar expected damages to a solution with infrequent large floods. As a result, the performance of the optimized operating policies is highly sensitive to the streamflows they are optimized to, making it difficult to reliably compare alternative solutions. These results call into question the effectiveness of using min-max objectives for robust optimization, as is often recommended in the literature [*Wald*, 1992; *Beyer and Sendhoff*, 2007], at least when the objective's performance is noisy and unbounded.

Fortunately, the second row of Figure 5 (plots d–g) shows that the WP1 objectives are much more stable, as the points cluster around the 1:1 line, with some solutions degrading in reevaluation and others improving. Interestingly, though, unlike on the WC objectives for which instability is independent of formulation, there is evidence of formulation-dependent instability on the WP1 Flood objective (plot f), as the WC and WP1 solutions do not systematically degrade, while the EV and EV&SD_H solutions do. Since the WC formulation is the only formulation that does not include the WP1 Flood objective, the stability of these solutions on this objective suggests that the formulation of the objective itself does not cause instability. The degradation of the EV and EV&SD_H solutions therefore must be due to overfitting of the control policies to the streamflows over which they were optimized. While the degradation of the EV&SD_H solutions on the WP1 Flood objective is less than for the EV solutions, it is still greater than for the WP1 solutions. This suggests that optimizing to the worst first percentile across all objectives results in fairly stable policies from year to year, while including expectation in the formulation results in more variable interannual performance on all objectives, not just the objectives optimized in expectation. Adding an objective related to interannual variability such as the standard deviation in hydropower production enables more stable performance on WP1 objectives, but not as stable as when optimizing to the worst first percentile across all objectives.

Another noteworthy observation from the second row of Figure 5 (plots d–g) is that the WC solutions lie far from the ideal point on both the WP1 flood resilience (WP1 Recovery, plot g) and flood vulnerability (WP1 Flood, plot c) objectives. Performance is particularly bad on the WP1 Flood objective, as several of the WC solutions lie above the black line drawn at 2.15 m, indicating that these solutions do not provide protection to the 100 year flood. While the WC formulation of the flooding objective was intended to be especially risk averse by modeling damages above 11.25 m with a fourth-order polynomial and minimizing the worst case performance across the ensemble, the calculation of expected damages over 20 years enables severe flood events to be masked by drier years with little to no damages. Instead of forcing the discovery of more conservative flood policies, the fourth-order polynomial only serves to make the objective values unstable, as shown in the row above (plot c). Ironically, the WC solutions perform well on the hydropower objective

from every formulation (plots a, d, and h), though, indicating that despite the harsh flood penalty this formulation actually favors optimizing hydropower production over flood protection.

On the contrary, the WP1 formulation of the flood objective, which focuses solely on large events, allows for the discovery of policies that provide protection at the 100 year level under both the optimization and validation streamflow ensembles. Unlike the nonunique damage function, the flood resilience and vulnerability objectives are able to distinguish between flooding caused by small, frequent events (captured by the resilience objective) and large, infrequent events (captured by the vulnerability objective). Additionally, despite the WP1 Flood objective only having the equivalent of a linear penalty on the maximum water level at Hanoi as opposed to the fourth-order polynomial on damages, the WP1 solutions still obtain low flood damages according to the WC Flood objective (plot c). This linear penalty is able to reduce the noise in the tails of the flooding objective while the worst first percentile bounds its performance, resulting in a stable objective that is able to simultaneously minimize expected damages. This greater flood protection does come at a cost, however, as the WP1 solutions do not do well on the WC Hydro objective (plot a).

These results suggest that minimizing flood damages in expectation may be ill-advised because it is difficult to know a priori whether or not doing so will be effective in reducing severe floods, especially when damages are uncertain or nonstationary and need to be approximated by a nonlinear penalty function. However, damage functions may still be useful for comparing optimized solutions a posteriori. For example, one could resimulate alternative nondominated policies over a larger ensemble of streamflows to estimate the maximum water level and corresponding damages of more extreme events like the 500 year flood, which one may not be able to estimate precisely over computationally tractable ensemble sizes for optimization. This can also provide stakeholders with a more realistic representation of the nonlinear mapping of stage to damages without suffering the negative consequences of optimizing to a noisy, nonlinear objective function.

The final row of Figure 5 (plots h–k) shows the performance of all of the solutions on the EV and EV&SD_H objectives. With most of the solutions lying nearly on the 1:1 line, these objectives are the most stable due to the smaller sampling variability of the mean and standard deviation than quantiles in the tails [*Stedinger et al*., 1993]. This row also highlights the regret associated with optimizing to the worst first percentile, as the solutions from the WP1 formulation do poorly on the EV Hydro objective (plot h). Regret in the opposite direction is not as severe, as the EV solutions do fairly well on the WP1 Hydro objective in the row above (plot d). However, while the WP1 solutions sacrifice EV Hydro performance, they do fairly well on the EV&SD_H Hydro Std objective (plot j) despite not explicitly including it in optimization. Additionally, while the greater stability in interannual hydropower production enabled by the WP1 formulation does degrade performance in EV Hydro, the degradation is not as severe as for the best Hydro Std solutions from the EV&SD_H formulation (plot h), suggesting that optimizing to the worst first percentile is a more effective way to reduce variability in performance without excessively sacrificing average performance. As shown in supporting information Figure S1, the operating behavior of the WP1 solutions is also more in line with conventional operations than that of the variance-minimizing EV&SD_H solution.

Ray et al. [2013] and *Watkins and McKinney* [1997] draw similar conclusions from a water supply optimization problem where including standard deviation in costs as part of the objective function led to more reliable and sustainable results with respect to shortages, but increased vulnerability. Noting that minimizing variance penalizes outcomes both above and below the mean, *Ray et al*. [2013] and *Watkins and McKinney* [1997] reformulated their objective to incorporate a penalty for squared positive cost deviations from a target, a modification inspired by *Takriti and Ahmed* [2004]. Similarly, the worst first percentile only penalizes outcomes below the mean for maximization objectives and above the mean for minimization objectives. Both of these alternative objective formulations are able to achieve more stable policies without excessively compromising mean performance, as *Ray et al*. [2013] and *Watkins and McKinney* [1997] found that the cost deviations penalty resulted in less variable direct costs without increasing vulnerability with respect to shortages.

The conclusions from Figure 5 highlight the importance of evaluating rival framings of how stakeholder objectives should be translated into quantitative performance measures when designing water resources systems optimization problems. While engaging stakeholders in the problem formulation process is important for ascertaining their qualitative objectives and such participatory modeling has improved decision-making for water resources applications [*Palmer et al*., 1990], this process alone does not guarantee the design of effective policies, as it is not obvious a priori what the best mathematical

characterization of stakeholder risk preferences will be. For example, if we had only translated stakeholder objectives into the metrics utilized by the WC formulation, policies optimized to that formulation would seem acceptable. However, since multiple quantitative translations of stakeholder objectives and preferences have been tested here, it is clear that these policies may not provide sufficient flood protection, a consequence that would have otherwise gone undiscovered. Thanks to this rival framings analysis, stakeholders in the Red River choosing from alternative policies can now better see the hydropower production, deficit and flood levels they would expect both on average and once every 100 years under policies from different formulations. Further analysis illustrating simulated behavior with different policies could better illustrate when and how severe periods of flooding, drought, and high or low hydropower production might be.

4.3. Impacts of Problem Framing and Preference on Control Policies and Flood Dynamics

For the second step of the verification process, we have selected solutions from different preference regions of each formulation's Pareto approximate set to analyze more deeply. In the supporting information, we examine the reservoir operations that result from these policies to determine how each solution is able to achieve its objective values. In general, we find that the best flood solutions maintain the lowest storages across the reservoirs to retain capacity to capture large flood events, while the best hydro solutions maintain the highest storages to have a high head differential for greater power production. Solutions favoring other objectives maintain intermediate storage levels. Comparing solutions from similar preference regions across formulations, the WC solutions tend to maintain the highest storages, explaining why they perform well on every formulation's hydropower objective, and poorly on the WP1 Flood objective.

In addition to examining the operations at each of the reservoirs, it is informative to visualize how these operations result in different responses downstream at Hanoi. For this analysis, we highlight solutions from the WC and WP1 formulations to distinguish the effects of these two variants of risk-averse optimization problems, and in particular two variants of flood control objectives. To illustrate a discrete number of policies spanning a wide range of preferences, we select the best hydro solution, best flood solution, and a compromise solution from each formulation. It has been long noted that multiobjective participatory planning in water resources systems is critical for discovering candidate compromise policies [*Maass et al*., 1962; *Cohon and Marks*, 1975; *Matalas and Fiering*, 1977; *Haimes and Hall*, 1977]. Recent advances in visual analytics have enhanced the interactive and collaborative exploration of candidate compromise solutions using techniques such as brushing objectives on stakeholders' performance criteria [*Basdekas*, 2014; *Herman et al*., 2014; *Huskova et al*., 2016; *Groves et al*., 2016]. Analyzing the behavior of compromise solutions from different formulations and how they compare with extreme solutions, one can see how stakeholders will make better, more-informed decisions if choosing compromise solutions from a number of different problem formulations.

Figure 6 shows where the three selected solutions from each formulation lie with respect to each of the objectives on a parallel axis plot (plots a and b), as well as how their storage trajectories differ at Hoa Binh (plots c and d). Generally speaking, the storage trajectories of the compromise solutions lie between those of the best hydro and best flood solutions. For each of these solutions, we use the simulations from the validation set of 100,000 years of synthetic inflows to estimate the probability density function (PDF) of the water level at Hanoi over time. Figure 7 shows these estimates in log space for each of the solutions, with high probabilities shaded red, moderate probabilities yellow, and low probabilities blue. A dotted line is drawn at the first alarm level of 6 m, a dashed line at the second alarm level of 11.25 m, and a solid line at the dike height of 13.4 m.

Across all solutions, the general shape of the time-varying PDFs is similar. The water level of the high-probability density region in red increases from May to July due to the monsoonal rains. As the rains subside at the end of the season, the water levels begin to fall and low levels are maintained throughout the dry season until the beginning of the next calendar year. Then, in response to the high releases from the reservoirs to meet the agricultural water demand during the planting season (see Figure 1), water levels at Hanoi briefly spike and then fall again before the next monsoon season begins. The shape of the second spike varies by solution, illustrating the different ways that the water demand can be met and the reservoirs emptied in advance of the monsoon. The shape of the spike for the best WP1 flood solution is particularly interesting as there seems to be a bifurcation with two different release magnitudes depending on how much water needs to be released to meet the agricultural water demand.

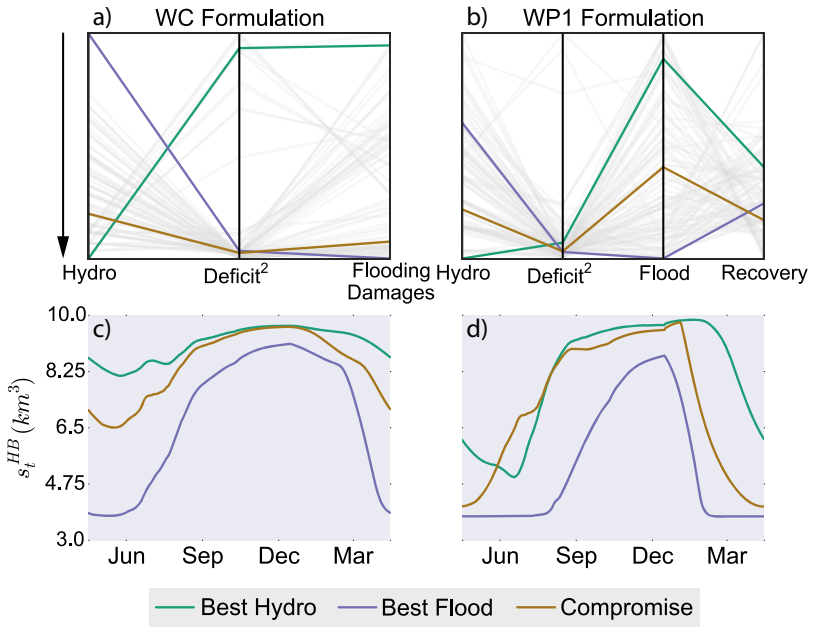

Figure 6. Best hydro solution (green), best flood solution (purple), and compromise solution (brown) from (left column: a and c) WC formulation and (right column: b and d) WP1 formulation. The top row (Figures 6a and 6b) shows the location of these solutions in the objective space of their problem formulations, while the bottom row (Figures 6c and 6d) shows their storage trajectories at Hoa Binh.

While the general shape of the time-varying PDFs is similar across solutions, there are some important differences. Comparing the three solutions from the WC formulation in the top row (plots a–c), these differences appear to be minor, but across the WP1 solutions in the bottom row (plots d–f), noticeable differences emerge. The best flood solution from the WP1 formulation (plot d) actively attempts to reduce the peak water level at Hanoi by maintaining moderately high water levels throughout the monsoon

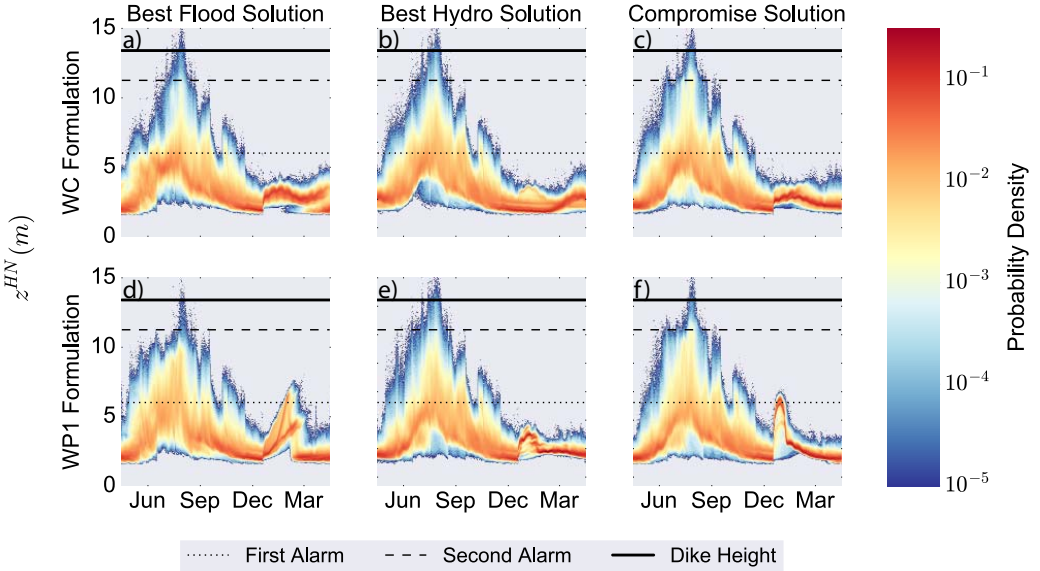

Figure 7. Probabilistic trajectories of the water level at Hanoi. (top row: a–c) Trajectories for three selected WC solutions and (bottom row: d–f) three selected WP1 solutions. The dotted line represents the first alarm level of 6 m, the dashed line the second alarm level of 11.25 m, and the solid line the dike height of 13.4 m. The best WP1 flood solution (Figure 7d) has a higher probability of crossing 6 m than the other solutions, exhibiting less resilience, but a lower probability of crossing 11.25 m, exhibiting less vulnerability. The probabilistic behavior of the compromise WP1 solution (Figure 7f) lies between that formulation's best flood (Figure 7d) and best hydro (Figure 7e) solutions, while differences are harder to see for the best WC solutions (Figures 7a–7c).

season. In contrast to the best flood solution from the WC formulation (plot a), it has a much greater probability of crossing 6 m, but a lower probability of crossing 11.25 m. This again highlights the trade-off between flood resilience and vulnerability. It should be noted, though, that reducing the probability of crossing 11.25 m also appears to reduce the probability of overtopping the dikes at 13.4 m, as the best flood solution from the WC formulation (plot a) also overtops the dikes more often than the best flood solution from the WP1 formulation (plot d), although this is a rare event for both solutions.

Comparing the best WP1 flood solution (plot d) to the best WP1 hydro solution (plot e), one can see the probabilistic effects of different preferences on the water level at Hanoi over time. Under operations with the best WP1 hydro solution (plot e), the time-varying density of the water level at Hanoi more closely resembles that of the best WC hydro solution (plot b), with water levels exceeding 6 m less often earlier in the monsoon season than the best WP1 flood solution (plot d). However, this results in a greater probability of exceeding both the second alarm level of 11.25 m and the dike height of 13.4 m. The WP1 compromise solution (plot f), strikes a balance between the two, crossing 11.25 m more often than the best WP1 flood solution (plot d), but less often than the best WP1 hydro solution (plot e). The fact that these differences are less obvious between the WC solutions in plots a–c, whose probabilistic water level dynamics most closely resemble those of the WP1 hydro solution, highlights that the WC formulation is actually far less conservative with respect to flooding than intended, and instead maximizes hydropower production. Consequently, stakeholders simply choosing a compromise solution among a set of nondominated policies from a single formulation would make a poor decision if only the WC formulation were used to design operating policies.

While the storage and release trajectories in concert with the time-varying PDFs in Figure 7 provide some understanding of how the operations lead to this coincident behavior, it is helpful to visualize these state trajectories jointly through a state space diagram [*Nayfeh and Balachandran*, 2008]. State space diagrams are useful for examining how a system evolves in time. For example, one can observe whether and when a system converges to a steady state, bifurcates into separate trajectories, or exhibits periodic orbital behavior [*Nayfeh and Balachandran*, 2008]. Applied here, the state space diagram may enhance our understanding of the stability of different operating policies, and how they achieve their objective values.

Figure 8 shows the probabilistic state space diagram for the compromise solutions from the WC (plot a) and WP1 (plot b) formulations, to further highlight the benefits of testing multiple problem formulations to guide stakeholders in discovering effective compromise solutions. Tracing the high-probability density regions in dark red, one can see that for the WC compromise solution in plot a, total reservoir storage initially increases without significantly increasing the water level at Hanoi. This is because the reservoirs are not releasing much of what comes in, trying to maintain high storages for hydropower production. However, once the reservoirs reach maximum capacity, they are forced to increase releases and the water level at Hanoi quickly rises. Notice this occurs before the total storage capacity has been reached. This indicates that this policy is not making full use of all of the reservoirs. Supporting information Figure S2 suggests that this is because the WC solutions maintain high storages at Hoa Binh to maximize hydropower production and attempt to use the smaller Thac Ba reservoir for flood protection. The state space diagram in Figure 8a suggests that this is insufficient, as the larger reservoirs fill to capacity and are forced to spill water downstream, while the smaller reservoirs have unused capacity that is not being fully exploited for flood protection. Consequently, under the lower probability events in yellow and light blue, water levels exceed 11.25 m over a range of total storages from about 15–25 km^3 and overtop the dikes at storages between about 22 and 25 km^3.

The compromise solution from the WP1 formulation exhibits very different joint state behavior. Tracing again the highest probability streak in red, one can see that storage levels increase in concert with the water level at Hanoi. This is because the reservoirs do not store up everything that comes in, but release some water to maintain more storage space for potential future flood events. Most of the time, the water level at Hanoi reaches the first alarm level of 6 m and levels off there. As the monsoon subsides, the water level at Hanoi drops while storage remains high. This is because little flow comes during the dry season, so the reservoirs store up as much water as possible to meet the agricultural demand. In the occasional wet year when the WP1 compromise solution reaches maximum storage, water levels also rise sharply as for the WC compromise solution, but this happens less often and only at maximum total system storage, indicating that this solution is better able to make use of the full system capacity for flood protection. As a result, the probability of crossing 11.25 m is lower, as seen by the greater blue shade above this line compared to the

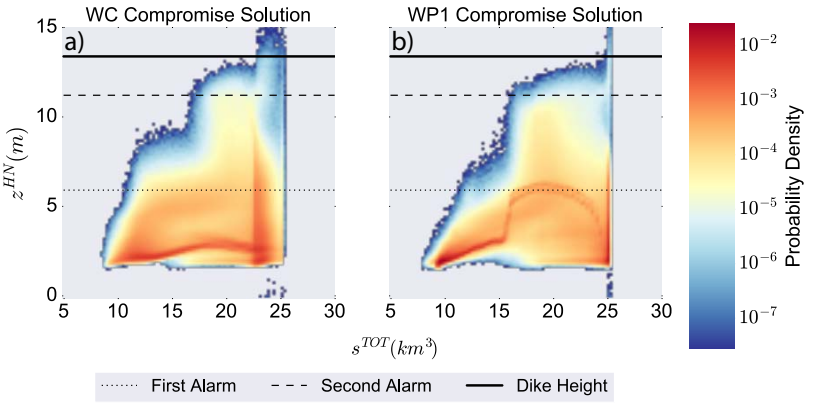

Figure 8. Joint probability density of total storage (*x* axis) and water level at Hanoi (*y* axis) under operations with the (a) WC compromise solution and (b) WP1 compromise solution. The WC compromise solution initially fills the reservoirs without releasing much, maintaining low water levels at Hanoi until the largest reservoirs reach maximum storage and must spill, causing the water level at Hanoi to jump. The WP1 compromise solution fills the reservoirs more slowly, releasing at the same time, causing water level and storage to rise simultaneously. This behavior makes better use of the full system capacity, reducing the probability of reaching maximum storage, and causing the water level at Hanoi to spike.

yellow shade above it for the WC compromise solution. Additionally, the WP1 compromise solution only ever results in overtopping when storage is at its maximum, while this occasionally occurs with the WC compromise solution before total storage capacity has been reached.

5. Conclusions

This study uses the multireservoir Red River system in Vietnam to illustrate that even modest changes in how objectives are quantified in a control problem can yield a surprising cascade of impacts on system performance. Consequently, it is important to test rival problem framings to determine the consequences of alternative quantitative abstractions of stakeholder objectives. In this system, where operating policies must balance the competing needs of flood management, hydropower production, and water supply for agriculture, we find that several commonly used problem framings can result in damaging unintended consequences. First, minimizing variance-based objectives often yields harsh consequences for expected performance while maximizing expected value objectives tends to expose systems to negative, high variance outcomes. In the context of reservoir controls, our results show that maximizing expected value objectives also has a tendency to yield over-fit control policies that do not generalize well out of sample. Finally, the Red River test case formulation with the greatest inherent negative consequences observed here is the worst case formulation commonly used in robust optimization [*Wald*, 1992; *Beyer and Sendhoff*, 2007] as it results in unstable policies that provide a poor representation of the system trade-offs and unintended modes of failure.

While it is known that the worst case over a large ensemble will degrade compared to the worst case over a smaller ensemble, this is not always of consequence if policies that minimize performance near the tails of an objective's distribution simultaneously minimize even more extreme values of the distribution. In this study, that is the case for the worst case hydropower and squared deficit objectives, but not for the worst case flood damages objective due to the noisy, unbounded penalty function used to approximate damages in expectation. An important outcome of our results is that minimizing this nonunique, nonlinear functional abstraction of flood damages in expectation inadvertently maximizes hydropower production, yielding levee overtopping in Hanoi for the 100 year flood level. This is true even if minimizing the worst case expected damages across an ensemble of multiyear streamflows. Fortunately, we find that the consequences of minimizing nonunique, nonlinear, worst case objectives can be overcome by formulating distinct, linear, worst first percentile objectives. In particular, minimizing the worst first percentile of the annual maximum water level solely targets large, infrequent events with a linear penalty, resulting in a stable objective that is able to simultaneously reduce expected damages, and the probability of observing even larger flood events. Additionally, it is able to reduce interannual variability without compromising expected performance as significantly as when including variance minimization as an objective.

These conclusions have significant implications for how reservoir operations are optimized for flood protection. Minimizing expected flood damages is a common objective in reservoir operations, but it may not adequately reduce system hazard. However, minimizing the more effective worst first percentile maximum annual flood objective while also maximizing expected hydropower production is only possible with simulation optimization frameworks such as EMODPS. Furthermore, EMODPS facilitates coordinated control across multiple reservoirs without suffering from the curse of dimensionality. The high-dimensional, multiobjective Red River control problem explored in this study is representative of the contextual and mathematical challenges that will be faced in a broad range of global multireservoir systems. Our ability to discover and appropriately manage the water, energy, and food trade-offs within these systems can be greatly advanced with the parameterization-simulation-optimization approach demonstrated here. Future work should focus on improving how information feedbacks and scalable control frameworks can be used to rigorously evaluate rival problem framings for managing complex river basins balancing evolving multisectoral demands, ecological impacts, and changing hydrologic extremes.

Appendix A: Objective Formulations

This appendix provides a detailed, mathematical description of the objectives in equation (10) under each of the four problem formulations. Recall from equation (8) that the d-th objective, J_d, is calculated by aggregating a daily metric, $g_d(t, i)$, over a T-year simulation (indexed by t) using some operator, Φ, and then filtering the result over an ensemble of N of these simulations (indexed by i) using some statistic, Ψ. For the WC formulation, objectives are calculated across $N = 50$ ensemble members in which simulations are of length $T = 20$ years, while for the other formulations $N = 1000$ ensemble members and $T = 1$ year.

A1. Hydropower Production

Across all formulations, total daily hydropower production $\eta_{t,i}$ from the four reservoirs in the i-th ensemble member, $g_{Hydro}(t, i)$, is averaged over the simulation length of each ensemble member:

$$\Phi_{Hydro}(i) = \mathbb{E}_{365T}[g_{Hydro}(t,i)] = \frac{1}{365T}\sum_{t=1}^{365T}\left[\sum_{j=1}^{4}\eta_{t,i}^{j}\right]. \tag{A1}$$

J_{Hydro}^{WC} is then calculated as the minimum value of $\Phi_{Hydro}(i)$ across the N ensemble members, J_{Hydro}^{WP1} as the first percentile, and J_{Hydro}^{EV} as the average:

$$J_{Hydro}^{WC} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Hydro}(i)] = \min_{i\in(1,\ldots,N)}[\Phi_{Hydro}(i)], \tag{A2}$$

$$J_{Hydro}^{WP1} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Hydro}(i)] = \underset{i\in(1,\ldots,N)}{\text{quantile}}\{\Phi_{Hydro}(i), 0.01\}, \text{ and} \tag{A3}$$

$$J_{Hydro}^{EV} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Hydro}(i)] = \mathbb{E}_N[\Phi_{Hydro}(i)] = \frac{1}{N}\sum_{i=1}^{N}\Phi_{Hydro}(i). \tag{A4}$$

A2. Squared Water Supply Deficit

Across all formulations, the daily squared water supply deficit in the i-th ensemble member, $g_{Deficit^2}(t, i)$, is first averaged over the simulation length of each ensemble member:

$$\Phi_{Deficit^2}(i) = \mathbb{E}_{365T}[g_{Deficit^2}(t,i)] = \frac{1}{365T}\sum_{t=1}^{365T}D_{t,i}^{2}. \tag{A5}$$

$J_{Deficit^2}^{WC}$ is then calculated as the maximum value of $\Phi_{Deficit^2}(i)$ across the N ensemble members, $J_{Deficit^2}^{WP1}$ as the 99th percentile, and $J_{Deficit^2}^{EV}$ as the average:

$$J_{Deficit^2}^{WC} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Deficit^2}(i)] = \max_{i\in(1,\ldots,N)}[\Phi_{Deficit^2}(i)] \tag{A6}$$

$$J_{Deficit^2}^{WP1} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Deficit^2}(i)] = \underset{i\in(1,\ldots,N)}{\text{quantile}}\{\Phi_{Deficit^2}(i), 0.99\}, \text{ and} \tag{A7}$$

$$J_{Deficit^2}^{EV} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Deficit^2}(i)] = \mathbb{E}_N[\Phi_{Deficit^2}(i)] = \frac{1}{N}\sum_{i=1}^{N}\Phi_{Deficit^2}(i). \tag{A8}$$

A3. Flood Damages and Vulnerability

In the WC formulation, the daily value of the flooding objective in the *i*-th ensemble member, $g^{WC}_{Flood}(t,i)$, is calculated using the penalty function displayed in Figure 3, which approximates damages as a function of the water level at Hanoi, $z^{HN}_{t,i}$. The damage function, $F(z^{HN}_{t,i})$, is a piecewise polynomial described by the following equation:

$$F(z^{HN}_{t,i}) = \begin{cases} 0, & z^{HN}_{t,i} \le 6.0\,\text{m} \\ \frac{75,000}{5.25}(z^{HN}_{t,i}-6), & 6.0\,\text{m} < z^{HN}_{t,i} \le 11.25\,\text{m} \\ 1.5\times10^{6}(z^{HN}_{t,i})^4 - 7.00\times10^{7}(z^{HN}_{t,i})^3 & \\ +1.22\times10^{9}(z^{HN}_{t,i})^2 - 9.45\times10^{9}z^{HN}_{t,i}, & z^{HN}_{t,i} > 11.25\,\text{m} \\ +2.74\times10^{10} & \end{cases} . \tag{A9}$$

Within each ensemble member, the daily value of the damage function is averaged over the simulation length:

$$\Phi^{WC}_{Flood}(i) = \mathbb{E}_{365T}[g^{WC}_{Flood}(t,i)] = \frac{1}{365T}\sum_{t=1}^{365T} F(z^{HN}_{t,i}). \tag{A10}$$

J^{WC}_{Flood} is then calculated as the maximum value of $\Phi^{WC}_{Flood}(i)$ across all *N* ensemble members:

$$J^{WC}_{Flood} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi^{WC}_{Flood}(i)] = \max_{i\in(1,\ldots,N)}[\Phi^{WC}_{Flood}(i)]. \tag{A11}$$

In the WP1, EV and EV&SD_H formulations, the flooding objective is framed as a flood vulnerability objective rather than a flood damage objective. Within each ensemble member *i*, $g^{WP1}_{Flood}(t,i)$ is defined as the daily water level at Hanoi in excess of 11.25 m, and $\Phi^{WP1}_{Flood}(i)$ is defined as the maximum value of $g^{WP1}_{Flood}(t,i)$ over the simulation length:

$$\Phi^{WP1}_{Flood}(i) = \max_{t\in(1,\ldots,365T)}[g^{WP1}_{Flood}(t,i)] = \max_{t\in(1,\ldots,365T)}[\max(z^{HN}_{t,i}-11.25\,\text{m}, 0)] \tag{A12}$$

J^{WP1}_{Flood} is then calculated as the 99^{th} percentile of $\Phi^{WP1}_{Flood}(i)$ across the *N* ensemble members and is constrained to be $\le$2.15 m:

$$J^{WP1}_{Flood} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi^{WP1}_{Flood}(i)] = \underset{i\in(1,\ldots,N)}{\text{quantile}}\{\Phi^{WP1}_{Flood}(i), 0.99\} \text{ and} \tag{A13}$$

$$J^{WP1}_{Flood} \le 2.15\,\text{m}. \tag{A14}$$

A4. Flood Resilience/Recovery Time

The WP1, EV, and EV&SD_H formulations all include an additional flooding objective to the flood vulnerability objective which represents the inverse of flood resilience. This objective, $J_{Recovery}$, indicates the time to "recover" once the water level at Hanoi exceeds 6 m. First, the within-ensemble average recovery time, $\Phi_{Recovery}(i)$, is calculated as:

$$\Phi_{Recovery}(i) = \frac{\sum_{t=1}^{365T} I_{t,i}}{\sum_{t=1}^{365T} v_{t,i}} \tag{A15}$$

where

$$I_{t,i} = \begin{cases} 0, & z^{HN}_{t,i} \le 6\,\text{m} \\ 1, & z^{HN}_{t,i} > 6\,\text{m} \end{cases} \text{ and} \tag{A16}$$

$$v_{t,i} = \begin{cases} 1, & t=1 \text{ and } z^{HN}_{t,i} > 6\,\text{m or} \\ & t>1,\ z^{HN}_{t,i} > 6\,\text{m and } z^{HN}_{t-1,i} \le 6\,\text{m}. \\ 0, & \text{otherwise} \end{cases} \tag{A17}$$

$I_{t,i}$ is an indicator variable signifying if the water level at Hanoi is above 6 m, while $v_{t,i}$ is an indicator variable signifying if a 6 m flood event has just begun. In the WP1 formulation, the 99th percentile value of

$\Phi_{Recovery}(i)$ across the N ensemble members is minimized, while in the EV and EV&SD_H formulations the average is minimized:

$$J^{WP1}_{Recovery} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Recovery}(i)] = \underset{i\in(1,\ldots,N)}{\text{quantile}}\{\Phi_{Recovery}(i), 0.99\} \text{ and} \tag{A18}$$

$$J^{EV}_{Recovery} = \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Recovery}(i)] = \mathbb{E}_N[\Phi_{Recovery}(i)] = \frac{1}{N}\sum_{i=1}^{N}\Phi_{Recovery}(i). \tag{A19}$$

A5. Standard Deviation of Annual Hydropower Production

In the EV&SD_H formulation, $J^{EV\,\&\,SD_H}_{Hydro\,Std}$ is defined as the standard deviation in average annual hydropower production, $\Phi_{Hydro}(i)$, across the N ensemble members:

$$\begin{aligned} J^{EV\,\&\,SD_H}_{Hydro\,Std} &= \underset{i\in(1,\ldots,N)}{\Psi}[\Phi_{Hydro}(i)] = \text{std}_N[\Phi_{Hydro}(i)] \\ &= \left[\tfrac{1}{N-1}\sum_{i=1}^{N}(\Phi_{Hydro}(i) - J^{EV}_{Hydro})^2\right]^{1/2}. \end{aligned} \tag{A20}$$

This is the only formulation which explicitly includes the interannual variability in hydropower production as an objective.

Acronyms

ANN	Artificial Neural Network
cap	capacity
CCFSC	Central Committee for Flood and Storm Control
EMODPS	Evolutionary Multi-Objective Direct Policy Search
EV	Expected Value
EV&SD_H	Expected Value & Standard Deviation of Hydropower
HB	Hoa Binh reservoir
HN	Hanoi
IWRP	Institute of Water Resources Planning
IMRR	Integrated and sustainable water Management of Red Thai Binh Rivers system in changing climate
lat	lateral flow
MARD	Ministry of Agriculture and Rural Development
MOEA	Multi-Objective Evolutionary Algorithm
MOIT	Ministry of Industry and Trade
RBF	Radial Basis Function
SL	Son La reservoir
TB	Thac Ba reservoir
TOT	total
TQ	Tuyen Quang reservoir
WC	Worst Case
WP1	Worst First Percentile

Notation

A	Number of RBFs in RBF policies
B	Number of inputs to RBF policies
b	Radii of RBF policies
C	Constraint on value of flooding objective function
c	Centers of RBF policies
D	Daily water supply deficit
e	Evaporation rate
F	Value of damage function on a given day

g	Value of objective function on daily time step
I	Indicator variable for days on which Hanoi water level exceeds 6 m
J	Objective function value across all ensemble members
M	Number of outputs of RBF policies
N	Number of ensemble members
q	Flow
r	Reservoir release
S	Surface area of reservoir
s	Reservoir storage
T	Number of years per ensemble member
t	Time
u	Policy-prescribed release
W	Water demand
w	Weights of RBF policies
x	Vector of inputs to RBF policies
z	Water level
η	Hydropower production per reservoir
θ	Vector of RBF parameters w, c, and b
ν	Indicator variable for days on which Hanoi water level first exceeds 6 m
τ	Tide level
Υ	Water volume in the canals
Φ	Operator for the aggregation of g over time
Ψ	Statistic used to filter noise of objective function across ensemble members
ϱ	Total inflow to the canals

Acknowledgments
The authors would like to thank Jonathan Lamontagne and Jery Stedinger for helpful discussions related to this work. This study was partially supported by the National Science Foundation (NSF) through the Network for Sustainable Climate Risk Management (SCRiM) under NSF cooperative agreement GEO-1240507 and the Penn State Center for Climate Risk Management. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding entities. All data used in this study are from the Ministry of Agriculture and Rural Development (MARD) of Vietnam and have been collected during the IMRR project (http://xake.elet.polimi.it/imrr/). Because the model code includes governmental information on hydropower plants, it cannot be made public. Please contact the authors to obtain specific data or model results.

References

Arrow, K. J. (1950), A difficulty in the concept of social welfare, *J. Polit. Econ.*, *58*(4), 328–346.

Asian Development Bank (2016), Vietnam: Energy Sector Assessment, Strategy, and Road Map, Mandaluyong City, Philippines. [Available at https://www.adb.org/documents/viet-nam-energy-sector-assessment-strategy-and-road-map.]

Basdekas, L. (2014), Is multiobjective optimization ready for water resources practitioners? Utility's drought policy investigation, *J. Water Resour. Plann. Manage.*, *140*(3), 275–276, doi:10.1061/(ASCE)WR.1943-5452.0000415.

Bernardi, D., M. Than Bui, M. Micotti, Q. Dinh, X. Nguyen, L. Schippa, R. Schmitt, V. Truong, P. Nam Vu, and E. Weber (2014), Report d5.1: Identifying the model, technical report, Integrated and sustainabale water Management of Red-Thai Binh River System in a changing climate. [Available at http://xake.elet.polimi.it/IMRR_Reports/D5.1Report.pdf.]

Beyer, H.-G., and B. Sendhoff (2007), Robust optimization: A comprehensive survey, *Comput. Methods Appl. Mech. Eng.*, *196*(33–34), 3190–3218, doi:10.1016/j.cma.2007.03.003.

Bosomworth, K., P. Leith, A. Harwood, and P. J. Wallis (2017), What's the problem in adaptation pathways planning?: The potential of a diagnostic problem-structuring approach, *Environ. Sci. Policy*, *76*, 23–28.

Castelletti, A., F. Pianosi, and R. Soncini-Sessa (2012a), Stochastic and robust control of water resource systems: Concepts, methods and applications, in *System Identification, Environmental Modelling, and Control System Design*, edited by L. Wang and H. Garnier, pp. 383–401, Springer. [Available at https://link.springer.com/book/10.1007/978-0-85729-974-1.]

Castelletti, A., F. Pianosi, X. Quach, and R. Soncini-Sessa (2012b), Assessing water reservoirs management and development in northern vietnam, *Hydrol. Earth Syst. Sci.*, *16*, 189–199.

Cohon, J., and D. Marks (1975), A review and evaluation of multiobjective programming techniques, *Water Resour. Res.*, *11*(2), 208–220.

Coumou, D., and S. Rahmstorf (2012), A decade of weather extremes, *Nat. Clim. Change*, *2*(7), 491–496.

De Kort, I. A., and M. J. Booij (2007), Decision making under uncertainty in a decision support system for the red river, *Environ. Modell. Software*, *22*(2), 128–136.

Dinh, N. Q. (2015), Multi-objective evolutionary algorithm, dynamic and non-dynamic emulators in the design of optimal policies for water resources management, PhD thesis, Politecnico di Milano, Milan, Italy. [Available at https://www.politesi.polimi.it/handle/10589/101048.]

Dittrich, R., A. Wreford, and D. Moran (2016), A survey of decision-making approaches for climate change adaptation: Are robust methods the way forward?, *Ecol. Econ.*, *122*, 79–89.

Giuliani, M., and A. Castelletti (2016), Is robustness really robust? how different definitions of robustness impact decision-making under climate change, *Clim. Change*, *135*(3–4), 409–424.

Giuliani, M., D. Anghileri, A. Castelletti, P. N. Vu, and R. Soncini-Sessa (2016a), Large storage operations under climate change: Expanding uncertainties and evolving tradeoffs, *Environ. Res. Lett.*, *11*(3), 035009.

Giuliani, M., A. Castelletti, F. Pianosi, E. Mason, and P. M. Reed (2016b), Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations, *J. Water Resour. Plann. Manage.*, *142*(2), 04015050.

Giuliani, M., J. D. Quinn, J. D. Herman, A. Castelletti, and P. M. Reed (2017), Scalable multi-objective control for large scale water resources systems under uncertainty, *IEEE Trans. Control Syst. Technol.*, (99), 1–8, doi:10.1109/TCST.2017.2705162. [Available at http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Julianne%20D.%20Quinn.QT.&newsearch=true.]

Groves, D. G., K. Kuhn, J. Fischbach, D. R. Johnson, and J. Syme (2016), *Analysis to Support Louisiana's Flood Risk and Resilience Program and Application to the National Disaster Resilience Competition*, RAND, Santa Monica, Calif. [Available at http://www.rand.org/content/dam/rand/pubs/research_reports/RR1400/RR1449/RAND_RR1449.pdf.]

Hadka, D., and P. Reed (2013), Borg: An auto-adaptive many-objective evolutionary computing framework, *Evol. Comput.*, *21*(2), 231–259.
Hadka, D., and P. Reed (2015), Large-scale parallelization of the Borg multiobjective evolutionary algorithm to enhance the management of complex environmental systems, *Environ. Modell. Software*, *69*, 353–369.
Haimes, Y. Y., and W. A. Hall (1977), Sensitivity, responsivity, stability, and irreversability as multiple objectives in civil systems, *Adv. Water Resour.*, *1*(2), 71–81.
Hall, J., D. Grey, D. Garrick, F. Fung, C. Brown, S. Dadson, and C. Sadoff (2014), Coping with the curse of freshwater variability, *Science*, *346*(6208), 429–430.
Hallegatte, S., M. Bangalore, L. Bonzanigo, M. Fay, T. Kane, N. Ulf, J. Rozenberg, D. Treguer, and A. Vogt-Schilb (2015), *Shock Waves: Managing the Impacts of Climate Change on Poverty*, World Bank, Washington, D. C. [Available at https://openknowledge.worldbank.org/bitstream/handle/10986/22787/9781464806735.pdf?sequence=13&isAllowed=y.]
Hansson, K., and L. Ekenberg (2002), Flood mitigation strategies for the red river delta, in *Proceeding of the 2002 Joint CSCE/EWRI of ASCE International Conference on Environmental Engineering, An International Perspective on Environmental Engineering, Niagara Falls, Ont., Canada, July*, pp. 21–24, Canadian Society for Civil Engineering, Montreal, Quebec. [Available at https://inis.iaea.org/search/searchsinglerecord.aspx?recordsFor=SingleRecord&RN=38094773.]
Hashimoto, T., J. R. Stedinger, and D. P. Loucks (1982), Reliability, resiliency and vulnerability criteria for water resource system performance evaluation, *Water Resour. Res.*, *18*(1), 14–20.
Herman, J., H. Zeff, P. Reed, and G. Characklis (2014), Beyond optimality: Multistakeholder robustness tradeoffs for regional water portfolio planning under deep uncertainty, *Water Resour. Res.*, *50*, 7692–7713, doi:10.1002/2014WR015338.
Hoppe, R. (2011), *The Governance of Problems: Puzzling, Powering and Participation*, Policy Press, Bristol, Great Britain. [Available at http://press.uchicago.edu/ucp/books/book/distributed/G/bo13442662.html.]
Huntington, T. G. (2006), Evidence for intensification of the global water cycle: Review and synthesis, *J. Hydrol.*, *319*(1), 83–95.
Huskova, I., E. S. Matrosov, J. J. Harou, J. R. Kasprzyk, and C. Lambert (2016), Screening robust water infrastructure investments and their trade-offs under global change: A London example, *Global Environ. Change*, *41*, 216–227.
Kasprzyk, J. R., P. M. Reed, B. R. Kirsch, and G. W. Characklis (2009), Managing population and drought risks using many-objective water portfolio planning under uncertainty, *Water Resour. Res.*, *45*, W12401, doi:10.1029/2009WR008121.
Kasprzyk, J. R., P. M. Reed, G. W. Characklis, and B. R. Kirsch (2012), Many-objective de novo water supply portfolio planning under deep uncertainty, *Environ. Modell. Software*, *34*, 87–104, doi:10.1016/j.envsoft.2011.04.003.
Kasprzyk, J. R., P. M. Reed, and D. M. Hadka (2015), Battling Arrow's paradox to discover robust water management alternatives, *J. Water Resour. Plann. Manage.*, *142*(2), 04015,053.
Kawachi, T., and S. Maeda (2004), Optimal management of waste loading into a river system with nonpoint source pollutants, *Proc. Jpn. Acad. Ser. B*, *80*(8), 392–398.
Kirsch, B. R., G. W. Characklis, and H. B. Zeff (2013), Evaluating the impact of alternative hydro-climate scenarios on transfer agreements: A practical improvement for generating synthetic stream flows, *J. Water Resour. Plann. Manage.*, *139*(4), 396–406, doi:10.1061/(ASCE)WR.1943-5452.0000287.
Koutsoyiannis, D., and A. Economou (2003), Evaluation of the parameterization-simulation-optimization approach for the control of reservoir systems, *Water Resour. Res.*, *39*(6), 1170, doi:10.1029/2003WR002148.
Labadie, J. W. (2004), Optimal operation of multireservoir systems: State-of-the-art review, *J. Water Resour. Plann. Manage.*, *130*(2), 93–111.
Le Ngo, L., H. Madsen, and D. Rosbjerg (2007), Simulation and optimisation modelling approach for operation of the Hoa Binh reservoir, Vietnam, *J. Hydrol.*, *336*(3), 269–281.
Lund, J. R. (2002), Floodplain planning with risk-based optimization, *J. Water Resour. Plann. Manage.*, *128*(3), 202–207.
Maass, A., M. M. Hufschmidt, R. Dorfman, H. A. Thomas, Jr., S. A. Marglin, and G. M. Fair (1962), *Design of Water-Resource Systems: New Techniques for Relating Economic Objectives, Engineering Analysis, and Governmental Planning*, Harvard Univ. Press, Cambridge.
Majone, G., and E. S. Quade (1980), *Pitfalls of Analysis*, John Wiley, Chichester, New York, Brisbane and Toronto.[Available at http://pure.iiasa.ac.at/1228/1/XB-80-108.pdf.]
Malekmohammadi, B., R. Kerachian, and B. Zahraie (2009), Developing monthly operating rules for a cascade system of reservoirs: Application of Bayesian networks, *Environ. Modell. Software*, *24*(12), 1420–1432.
Matalas, N. C., and M. B. Fiering (1977), Water-resource systems planning, in *Climate, Climatic Change, and Water Supply. Studies in Geophysics*, pp. 99–110, Natl. Acad. of Sci., Washington, D. C. [Available at https://www.nap.edu/catalog/185/climate-climatic-change-and-water-supply.]
Nayfeh, A. H., and B. Balachandran (2008), *Applied Nonlinear Dynamics: Analytical, Computational and Experimental Methods*, John Wiley, Weinheim, Germany. [Available at http://onlinelibrary.wiley.com/doi/10.1002/9783527617548.fmatter/pdf.]
Needham, J. T., D. W. Watkins Jr., J. R. Lund, and S. Nanda (2000), Linear programming for flood control in the Iowa and Des Moines rivers, *J. Water Resour. Plann. Manage.*, *126*(3), 118–127.
Nguyen, T. C., N. H. Do, T. Nguyen, and K. Egashira (2002), Agricultural development in the red river delta, Vietnam-water management, land use, and rice production, *J. Fac. Agr. Kyushu U.*, *46*(2), 445–464.
Nowak, K., J. Prairie, B. Rajagopalan, and U. Lall (2010), A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow, *Water Resour. Res.*, *46*, W08529, doi:10.1029/2009WR008530.
Olsson, L., et al. (2014), Livelihoods and poverty, in *Climate Change 2014: Impacts, Adaptation and Vulnerability. Part A: Global and Sectoral Aspects, Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by C. B. Field, eds., pp. 793–832, Cambridge University Press, Cambridge, United Kingdom and New York, N. Y. [Available at https://www.ipcc.ch/pdf/assessment-report/ar5/wg2/WGIIAR5-Chap13_FINAL.pdf.]
Orlovski, S., S. Rinaldi, and R. Soncini-Sessa (1984), A min-max approach to reservoir management, *Water Resour. Res.*, *20*(11), 1506–1514.
Palmer, R. N., W. J. Werick, A. MacEwan, and A. W. Woods (1990), Modeling water resources opportunities, challenges, and tradeoffs: The use of shared vision modeling for negotiation and conflict resolution, in *Proceedings of ASCE's 26th Annual Conference on Water Resources Planning and Management*, Am. Soc. Civ. Eng.
Pareto, V. (1896), *Cours D'Economie Politique*, F. Rouge, Lausanne.
Ray, P. A., D. W. Watkins Jr, R. M. Vogel, and P. H. Kirshen (2013), Performance-based evaluation of an improved robust optimization formulation, *J. Water Resour. Plann. Manage.*, *140*(6), 04014,006.
Reed, P. M., and D. Hadka (2014), Evolving many-objective water management to exploit exascale computing, *Water Resour. Res.*, *50*, 8367–8373, doi:10.1002/2014WR015976.
Reed, P. M., and J. R. Kasprzyk (2009), Water resources management: The myth, the wicked, and the future, *J. Water Resour. Plann. Manage.*, *135*(6), 411–413.

Reed, P. M., D. Hadka, J. D. Herman, J. R. Kasprzyk, and J. B. Kollat (2013), Evolutionary multiobjective optimization in water resources: The past, present and future, *Adv. Water Resour.*, *51*, 438–456.
Roy, B. (1990), Decision-aid and decision-making, *Eur. J. Oper. Res.*, *45*, 324–331.
Roy, B. (2010), Robustness in operational research and decision aiding: A multi-faceted issue, *Eur. J. Oper. Res.*, *200*(3), 629–638.
Smith, R., J. Kasprzyk, and L. Dilling (2017), Participatory framework for assessment and improvement of tools (ParFAIT): Increasing the impact and relevance of water management decision support research, *Environ. Modell. Software*, doi:https://doi.org/10.1016/j.envsoft.2017.05.004, in press.
Soncini-Sessa, R., J. Zuleta, and C. Piccardi (1990), Remarks on the application of a risk-averse approach to the management of "el Carrizal" reservoir, *Adv. Water Resour.*, *13*(2), 76–84.
Soncini-Sessa, R., E. Weber, and A. Castelletti (2007), *Integrated and Participatory Water Resources Management-Theory*, vol. 1, Elsevier, Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney and Tokyo. [Available at https://www.elsevier.com/books/integrated-and-participatory-water-resources-management-theory/soncini-sessa/978-0-444-53013-4.]
Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993), Frequency analysis of extreme events, in *Handbook of Hydrology*, edited by D. R. Maidment, chap. 18, McGraw-Hill, New York.
Stirling, A. (2008), "Opening up" and "closing down" power, participation, and pluralism in the social appraisal of technology, *Sci. Technol. Hum. Val.*, *33*(2), 262–294.
Taguchi, G. (1986), *Introduction to Quality Engineering: Designing Quality into Products and Processes*, Asian Productivity Organization, Tokyo.
Takriti, S., and S. Ahmed (2004), On robust optimization of two-stage systems, *Math. Program.*, *99*(1), 109–126.
Tanaka, S. K., T. Zhu, J. R. Lund, R. E. Howitt, M. W. Jenkins, M. A. Pulido, M. Tauber, R. S. Ritzema, and I. C. Ferreira (2006), Climate warming and water management adaptation for California, *Clim. Change*, *76*(3–4), 361–387.
Thomas, Jr., H. A., and M. B. Fiering (1962), Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, in *Design of Water-Resource Systems: New Techniques for Relating Economic Objectives, Engineering Analysis, and Governmental Planning*, edited by A. Maass, et al., Harvard Univ. Press, Cambridge.
Trenberth, K. E. (2011), Changes in precipitation with climate change, *Clim. Res.*, *47*(1–2), 123–138.
Tsoukias, A. (2008), From decision theory to decision aiding methodology, *Eur. J. Oper. Res.*, *187*(1), 138–161.
Vinh Hung, H., R. Shaw, and M. Kobayashi (2007), Flood risk management for the Rua of Hanoi: Importance of community perception of catastrophic flood risk in disaster risk planning, *Disaster Prev. Manage.*, *16*(2), 245–258.
Wald, A. (1992), Statistical decision functions, in *Breakthroughs in Statistics*, edited by S. Kotz, and N. L. Johnson, pp. 342–357, Springer, Berlin, Heidelberg, New York. [Availble at https://link.springer.com/content/pdf/10.1007%2F978-1-4612-0919-5.pdf.]
Walker, W. E., P. Harremoës, J. Rotmans, J. P. van der Sluijs, M. B. van Asselt, P. Janssen, and M. P. Krayer von Krauss (2003), Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support, *Integr. Assess.*, *4*(1), 5–17.
Watkins, D. W., and D. C. McKinney (1997), Finding robust solutions to water resources problems, *J. Water Resour. Plann. Manage.*, *123*(1), 49–58.
Windsor, J. S. (1973), Optimization model for the operation of flood control systems, *Water Resour. Res.*, *9*(5), 1219–1226.
World Bank (2016), *High and Dry: Climate Change, Water and the Economy*, Washington, D. C.
Yakowitz, S. (1982), Dynamic programming applications in water resources, *Water Resour. Res.*, *18*(4), 673–696, doi:10.1029/wr018i004p00673.
Yeh, W. W.-G. (1985), Reservoir management and operations models: A state-of-the-art review, *Water Resour. Res.*, *21*(12), 1797–1818.
Zatarain Salazar, J., P. M. Reed, J. D. Herman, M. Giuliani, and A. Castelletti (2016), A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control, *Adv. Water Resour.*, *92*, 172–185.
Zeff, H. B., J. R. Kasprzyk, J. D. Herman, P. M. Reed, and G. W. Characklis (2014), Navigating financial and supply reliability tradeoffs in regional drought management portfolios, *Water Resour. Res.*, *50*, 4906–4923, doi:10.1002/2013WR015126.
Zeleny, M. (1981), On the squandering of resources and profits via linear programming, *Interfaces*, *11*(5), 101–107.
Zeleny, M. (1989), Cognitive equilibrium: A new paradigm of decision making?, *Hum. Syst. Manage.*, *8*, 185–188.