POLITECNICO
MILANO 1863

Politecnico di Milano
Department of Electronics, Information, and Bioengineering
Doctoral Program In Information Technology

Feature Representation Learning in complex water decision making problems

Doctoral Dissertation of:
Marta Zaniolo

Advisor:
Prof. Andrea Castelletti

Co-Advisor:
PhD. Matteo Giuliani

Tutor:
Prof. Fabio Dercole

The Chair of the Doctoral Program:
Prof. Barbara Pernici

XXXII Cycle

A mia nonna Imelda,
che ha lottato per se stessa
e per tutte le donne.

Abstract

The success of a control policy highly relies by its feature representation, i.e., the information set it is conditioned upon. In real world control problems, defining an appropriate feature representation is a complex task, given the coexistence of multiple interacting processes whose relevance for the control task is often unclear. In this thesis, we address the control problem of water resources systems, where a dam release policy is designed accounting for multiple water demands. This decisional problem is challenged by the presence of non-linearities, strong disturbances, possible alternative problem framings, and multiple conflicting objectives. Currently, the control rules of most water reservoirs are conditioned upon basic information systems comprising reservoir storage and time index, however, the value of a more informative feature representation is generally undisputed. We capitalize on recent advances in monitoring and forecasting water availability to develop novel feature representation learning strategies to enhance water systems resilience towards their crucial vulnerabilities, including droughts, critical phases in reservoir development (i.e., construction and filling), and multisectoral conflicts. Additionally, in multi-purpose systems, different control targets might be heterogeneous in their dynamics and vulnerabilities, and likely benefit from a tailored feature representation that varies across different objectives tradeoffs.

We revise current literature on feature representation learning, and propose a taxonomy comprising *a priori*, *a posteriori*, and *online* approaches. For each approach, we propose novel contributions targeting the control problem of multipurpose water systems. Among the methodological contributions included in this thesis, (1) we propose FRIDA, a feature extraction-based framework to design basin-tailored drought indexes, and (2) we employ FRIDA index to inform water reservoir operations; (3) we extend the concepts of feature representation learning beyond pure control applications to a problem of dam planning and filling; (4) we use Artificial Intelligence to capture the state of multiple climate signals to improve seasonal forecast in a framework named CSI; (5) we propose an original multi-objective neuro-evolutionary algorithm, NEMODPS, that evolves tradeoff-tailored policy architectures, and (6) we combine

it with a feature selection routine to learn a policy representation online and tradeoff-dynamically.

A common thread of the outcomes generated in this collection of works is that learning an appropriate policy information set is an asset to improve water system performance, especially by targeting its most critical failures. Specifically, by mitigating the damages associated with hydrological extremes (e.g., drought emergencies), critical stages reservoir development (i.e., construction and filling), and social tensions deriving from conflicts between different users and their demands.

Part of this research has appeared (or will appear) in the following journal publications:

(1) Zaniolo, M., Giuliani, M., Castelletti, A.F., Pulido-Velazquez, M., 2018b. Automatic design of basin- specific drought indexes for highly regulated water systems. Hydrology and Earth System Sciences 22, 2409-2424. *(Chapter 2)*;

(2) Zaniolo, M., Giuliani, M., Castelletti, A., 2019. Data-driven modeling and control of droughts. IFAC- Papers On Line 52, 54-60. *(Chapter 3)*;.

(3) Zaniolo, M., Giuliani, M., Burlando, P., Castelletti, A., 2020a When timing matters - misdesigned dam filling impacts hydropower sustainability. Nature Sustainability (under review). *(Chapter 4)*;

(4) Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., Block, P., 2019. Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. Water Resources Research 55, 9133-9147. *(Chapter 5)*;

(5) Zaniolo, M., Giuliani, M., Castelletti, A., 2020b. Neuro-evolutionary direct policy search for multi-objective optimal control. IEEE transactions on neural networks and learning systems (under review). *(Chapter 6)*;

(6) Zaniolo, M., Giuliani, M., Castelletti, A., 2020c. Dynamic retrieval of informative inputs for multi-sector reservoir policy design with diverse spatio-temporal objective scales. Environmental Modeling and Software (in preparation). *(Chapter 7)*.

Sommario

Il successo di una politica di controllo dipende fortemente dalla sua rappresentazione, ovvero dall'insieme di variabili con cui è informata. In problemi di controllo nel mondo reale, la definizione di una rappresentazione appropriata è un compito complesso, data la coesistenza di più processi interagenti la cui rilevanza per il problema di controllo è spesso poco chiara. In questa tesi, affrontiamo il problema di controllo dei sistemi di risorse idriche, in cui una politica di rilascio della diga è progettata tenendo conto di molteplici domande idriche. Questo problema decisionale è complicato dalla presenza di non linearità , forti disturbi, possibili formulazioni alternative del problema, e molteplici obiettivi contrastanti. Attualmente, le regole di controllo della maggior parte dei bacini idrici sono condizionate su sistemi informativi basilari che considerano l'invaso del serbatoio e un indice del tempo, d'altra parte, il valore di una rappresentazione della politica più ricca e informativa è generalmente indiscusso. Sfruttiamo i recenti progressi nel monitoraggio e nella previsione della disponibilità di acqua per sviluppare nuove strategie di apprendimento della rappresentazione della politica per migliorare la resilienza dei sistemi idrici rispetto a vulnerabilità cruciali tra cui siccità , fasi critiche nello sviluppo di serbatoi (ad esempio costruzione e riempimento), e conflitti tra diversi settori. Inoltre, nei sistemi caratterizzati da molteplici usi della risorsa idrica, diversi obiettivi di controllo potrebbero essere eterogenei nelle loro dinamiche e vulnerabilità, e dunque trarre vantaggio da una rappresentazione delle caratteristiche su misura che varia a seconda dei diversi obiettivi.

Analizziamo la letteratura recente sull'apprendimento della rappresentazione della politica, e proponiamo una tassonomia che comprende approcci *a priori*, *a posteriori*, e *online*. Per ogni approccio, proponiamo contributi metodologici originali, mirati al problema del controllo dei sistemi idrici a molti-obiettivi. Tra i contributi metodologici inclusi in questa tesi, (1) proponiamo FRIDA, una procedura basata sull'estrazione di variabili per progettare indici di siccità di bacino su misura, e (2) impieghiamo l'indice FRIDA per informare le operazioni di gestione di una diga; (3) estendiamo i concetti di apprendimento della rappresentazione della politica oltre applicazioni di puro controllo a un problema

di pianificazione e riempimento di un serbatoio idrico artificiale; (4) utilizziamo tecniche di Intelligenza Artificiale per analizzare lo stato di diversi segnali climatici per migliorare le previsioni stagionali, secondo una procedura originale chiamata CSI; (5) proponiamo un nuovo algoritmo neuro-evolutivo a molti obiettivi, NEMODPS, che evolve un' architettura della politica su misura per gli obiettivi e relativi compromessi; (6) combiniamo NEMODPS con una strategia di selezione di variabili che apprende una rappresentazione della politica online, e dinamica rispetto agli obiettivi.

Un filo conduttore dei risultati generati in questa raccolta di lavori è che l'apprendimento di un adeguato set di informazioni per informare la politica si configura come una valida risorsa per migliorare le prestazioni del sistema idrico, in particolare rispetto alle sue vulnerabilità più critiche. Nello specifico, mitigando i danni associati agli estremi idrologici (ad esempio le siccità), fasi critiche di sviluppo del serbatoio (costruzione e riempimento), e tensioni sociali derivanti da conflitti tra i diversi utenti idrici.

Parte della ricerca presentata in questa tesi è apparsa, o apparirà, nelle seguenti pubblicazioni:

(1) Zaniolo, M., Giuliani, M., Castelletti, A.F., Pulido-Velazquez, M., 2018b. Automatic design of basin- specific drought indexes for highly regulated water systems. Hydrology and Earth System Sciences 22, 2409-2424. *(Capitolo 2)*;

(2) Zaniolo, M., Giuliani, M., Castelletti, A., 2019. Data-driven modeling and control of droughts. IFAC- Papers On Line 52, 54-60. *(Chapter 3)*;.

(3) Zaniolo, M., Giuliani, M., Burlando, P., Castelletti, A., 2020a When timing matters - misdesigned dam filling impacts hydropower sustainability. Nature Sustainability (under review). *(Capitolo 4)*;

(4) Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., Block, P., 2019. Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. Water Resources Research 55, 9133-9147. *(Capitolo 5)*;

(5) Zaniolo, M., Giuliani, M., Castelletti, A., 2020b. Neuro-evolutionary direct policy search for multi-objective optimal control. IEEE transactions on neural networks and learning systems (under review). *(Capitolo 6)*;

(6) Zaniolo, M., Giuliani, M., Castelletti, A., 2020c. Dynamic retrieval of informative inputs for multi-sector reservoir policy design with diverse spatio-temporal objective scales. Environmental Modeling and Software (in preparation). *(Capitolo 7)*.

Acknowledgements

Several people contributed to my PhD education, and the research presented in this dissertation. First, I would like express my sincere gratitude to my supervisor, Professor Andrea Castelletti, for giving me the opportunity to receive a PhD eduction, and for guiding me every step of the way with dedication and trust. I soon realized how lucky I was to be part of his mentoring vision, one that exposes PhD students to many opportunities to take part in international conferences, engage with other research groups, and disseminate their work. I am thankful to Professor Matteo Giuliani, my co-advisor, always present, and incredibly helpful. Thank you for the countless times you made yourself available to brainstorm ideas and discuss my doubts. My gratitute goes to Professor Paul Block for collaborating with me during my visiting in the University of Wisconsin, Madison, a time of intense professional and personal growth I hold dear to my heart.

I'm very thankful to the reviewers of this thesis, Professor Patrick Reed and Professor Julianne Quinn, for taking the time to read my work and offering their suggestions. Their inputs have been essential to add value to the final version of this dissertation, and their words of appreciation humbled me . My PhD was supported by DAFNE, an H2020 European project. Witnessing and contributing to the developent of such a large and ambitious project taught me a lot. My gratitude goes to all DAFNE members, and Professor Paolo Burlando as project leader.

Lastly, I would like to thank the members of the Environmental Intelligence Lab, Alessandro, Angelo, Federica, Federico, Jazmin, Marco for sharing coffee breaks, lunches, and the occasional aperitivo. I'm grateful for your friendship and the supportive environment we created. I watched you and myself become better researchers together, and you made every day in the office much better.

I will keep out of this page the immense gratitude I feel for the people I love outside the academic world, but I here vow to thank each one of them personally.

Marta

Contents

List of Figures

List of Tables

1

Introduction

On August 6th, 2012, NASA's Curiosity Rover landed on Mars as part of the Mars Science Laboratory mission. Curiosity's aim is to investigate Martian climate and geology, and search for possible past traces of water and planet habitability in preparation for human exploration. Due to the long travel time of signals between Earth and Mars, the rover is not directly controlled by NASA's scientists, but is autonomous in the planet exploration. Curiosity is equipped with a control unit that interacts with the Martian environment and learns how to navigate it, trying to cover large portions of land, collect environmental samples, while maintaining stability, energy efficiency, and avoiding the planet's spots of soft soil (Welch et al., 2013). Unprecedentedly, Curiosity is also autonomous in its target selection, and is able to identify, without any input from Earth, the relevant rock and soil targets to collect and send back to NASA's computers (Francis et al., 2017).

Mathematically, the control problem of Curiosity can be described as a highly complex Markov Decision Process (MDP) (Zilberstein et al., 2002; Huntsberger et al., 2005) that would be infeasible to solve with a traditional exact control problem formulation (e.g., Dynamic Programming). In these cases, Reinforcement Learning (RL, Sutton et al., 1998) approaches have been established as valid alternatives to handle peculiar challenges of the problem at hand, e.g., the presence of numerous continuous state and control variables, uncertainty in problem formulation, existence of significant disturbances, and multiple conflicting objectives.

Examples of control tasks characterized by one or more of these challenges are numerous, and encompass different domains, e.g., autonomous robot control (Francis et al., 2017), self-learning artificial videogame players (Szita, 2012),

self-driving cars (Shalev-Shwartz et al., 2016; Sallab et al., 2017), and the operation of data centers (Knight, 2017). A peculiar example of a control problem that can comprise all the above challenges is the problem of designing a release policy for a water reservoir accounting for multiple water demands (e.g., Castelletti et al., 2008a, 2010a, 2012, 2013), which will be the main focus of this thesis.

Albeit different in nature, the common characteristic of the cited control problems is the presence of an agent whose goal is to learn an optimal action policy via interactions with an environment. In particular, the agent performs actions on the environment and collects the reward signals associated with the environmental transition to a new state. By trials-and-errors, and continuous collection of environmental samples, the agent's aim is to learn an action policy, or control policy, that maximizes its cumulated reward.

In the described learning mode, namely online learning, the agent becomes proficient in the assigned task by trials-and-errors, directly performing actions on an environment. Recently, there has been growing interest in deploying RL algorithms to new, unconventional environments such as healthcare (Shortreed et al., 2011), education (Mandel et al., 2014), or water management (Castelletti et al., 2012). In these high-stakes situations, experimenting arbitrary actions directly on the environment can easily generate unacceptable social costs. Instead, a dataset of information on the system is pre-collected or generated via simulation, and employed to train the agent on a virtual environment until a good policy is designed, namely offline learning (Mandel et al., 2016).

The most attractive characteristic of the RL framework is its highly autonomous nature in exploring the environment and discovering how to interact with it, yet, the computational time and number of interactions required to learn an optimal control policy is exponential in the number of features that condition an agent's control policy. For instance, considering the control of a water infrastructure operating in a hydrological basin, the entire problem environment comprises the water available in the whole basin under multiple forms: artificial and natural lakes, snow, groundwater, and topsoil water content. Additionally, part of the problem environment is also the atmospheric conditions insisting on the watershed, including temperature, winds, air humidity, cloud cover, and precipitation (Turner and Galelli, 2016; Denaro et al., 2017a). An agent could potentially consider in its policy learning all signals detectable in the environment, including derived spatial and temporal aggregations, indices, modeled variables, and forecasts. Yet, likely, only a limited subset of features are relevant for the task. On the one hand, overloading the agent with irrelevant or redundant information hampers the policy search process and results in increased computational time and decreased policy quality. On the other hand, the agent should have access to enough information to predict the environmental response to its actions, and learn a sufficiently complex policy (Hachiya and

Sugiyama, 2010). The problem of defining an appropriate subset of features that constitutes a low dimensional and highly informative policy input set is called feature representation learning (e.g., Maillard et al., 2011). Specifically, a good feature representation as defined by Böhmer et al. (2015) must be i) Markovian, i.e., contain enough information to predict the expected state transition associated to an action, given the current state, ii) able to describe appropriately the current state of the environment, iii) generalizable to unseen states, and, iv) sufficiently low dimensional.

The recent breakthroughs in Deep RL techniques (i.e., employing deep neural networks in RL) has increased the capacity of controllers to interact with high-dimensional feature spaces (e.g., Mnih et al., 2015; Lillicrap et al., 2015) by autonomously learn directly from large datasets of raw observations, apparently mitigating the need for feature representation (Si et al., 2017). However, in order to produce reliable policies, such methods require a large amount of training data to discern true causal connections from the spurious ones that may verify in large datasets (Rusu et al., 2016). This data can be very costly to obtain, or simply not available in real world control problems. Besides, without an informative and compact feature representation, it is unclear how the learning agent is synthesizing the cloud of raw information into a policy, thereby producing blackbox results that are strictly case-specific (Lesort et al., 2018). Besides, the benefits of learning a reduced feature representation go beyond the simple control policy design. For instance, it can generate insights on the control problem by identifying task-relevant features and drivers. Those insights can be generalized to similar applications beyond the specific problem, and across domains (de Bruin et al., 2018), for instance by guiding feature representation learning on related tasks, or even by informing the planning of a cost-effective monitoring network in a new environment with similar characteristics. As a consequence, for a more efficient and generalizable learning process, even Deep RL routines can take advantage from operating on a compact and informative feature representation, instead of a high-dimensional policy input space (Munk et al., 2016; Lesort et al., 2018).

While being a defining topic since the earlier works on RL (Dominey, 1995; Sutton et al., 1998), feature representation learning is still an open issue especially in real-world control applications. First, as the environments where autonomous agents operate become more complex (i.e., contain more sensors of increased diversity), the extraction of relevant features is increasingly challenging (Nouri and Littman, 2010). Additionally, real world control problems are generally characterized by the coexistence of multiple conflicting objectives, a major challenge for feature representation learning, given the need to identify a policy representation that is optimal across objectives (Liu et al., 2014).

In the next section, we present a literature review of relevant past experience

in feature representation for RL. Subsequently, Section 1.2 presents the aim of the thesis, and Section 1.3 its structure.

1.1 Feature Representation Learning

The problem of feature representation in RL is generally tackled by pairing Feature Extraction (FE) algorithms with Policy Search (PS) methods (Liu et al., 2015; Lesort et al., 2018).

FE are a family of techniques that transform an original dataset into a more compact, while still highly informative dataset (Cunningham, 2008). FE algorithms are widely used in supervised learning applications that aim at reproducing a given target variable. Three classes of FE algorithms are generally recognized in supervised learning, namely filters, wrappers and embedded methods (see, e.g., Cunningham, 2008, and reference therein). Filter algorithms evaluate the relevance of a feature based on the intrinsic properties of the dataset, generally by computing an error metric between feature and target (e.g., Yang and Pedersen, 1997; Sharma, 2000a). Wrappers evaluate the fitness of a variable ensemble in reproducing the target given a learning machine (e.g., Taormina et al., 2016). Embedded methods include dimension reduction as part of the learning (e.g., Gashler et al., 2008). However, such taxonomy does not generalize well to RL applications in the absence of a target variable.

In the following review, we consider the case of a complex control problem, for which deriving an appropriate feature representation is not an obvious task, and a good control policy must be learned offline, via a precollected dataset of environmental signals that are considered as candidate policy inputs. This dataset comprises heterogeneous and multi-source observed and modeled variables, disturbances, and their manipulations, e.g., linear and non-linear combinations, forecasts, and aggregations, at a variety of time and spatial scales.

In the proposed taxonomy, we identify *a priori*, *a posteriori*, and *online* approaches to pairing FE and PS for feature representation and policy learning.

- **A priori.**

  In the *a priori* approach, the FE step is antecedent and independent from the PS step. The dimensionality of the dataset of candidate features is reduced on the basis of intrinsic properties of the controlled system, and does not depend on the policy search process (Fig. 1.1). Three main *a priori* strategies can be identified.

  The first strategy employs dimension reduction techniques to map the dataset of candidate features into a lower dimensional latent space that retains most of its information content, e.g., via autoencoders (e.g., Morimoto et al., 2008), or Principal Component Analysis (Nouri and Littman,

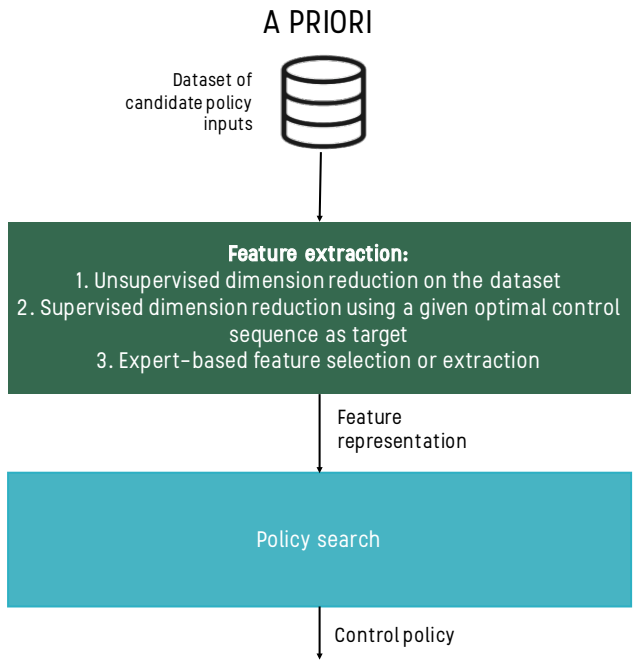

Figure 1.1: *Feature representation learning taxonomy, a priori approach: feature extraction precedes policy design.*

2010). Such strategies are suitable for environments presenting high redundancy, but, on the other hand, do not detect the presence of irrelevant variables that would add dimensionality and complexity to the feature representation without contributing to its information content.

In the second strategy, the optimal control sequence is known or can be computed in advance for a given disturbance realization (e.g., via Deterministic Dynamic Programming). This sequence is used to derive the most informative policy drivers by formulating the feature representation problem as a supervised FE problem, where the optimal control sequence is employed as target variable (Kroon and Whiteson, 2009; Giuliani et al., 2015; Denaro et al., 2017a).

The third strategy uses expert-based feature selection (e.g., Akrour et al., 2012) or extraction (e.g., Sturtevant and White, 2006; Giuliani and Castelletti, 2019) to design a problem-specific representation. In some cases, the expert-based representation can be rigorous and motivated by physical characteristics of the controlled system. In the example of a water reservoir operated for flood control, the key information to the operations is the time necessary to create a buffer volume for a possible incoming flood

that can be derived from physical properties, e.g., reservoir or spillway capacity (Saavedra Valeriano et al., 2010; Raso et al., 2014).

In general, a priori feature representation is advisable whenever there is sufficient knowledge on the system and the task to confidently devise a dataset dimension reduction without needing any feedback from policy design. This method, in fact, does not offer any guarantees on the optimality of the chosen representation.

- **A posteriori.** This approach evaluates the suitability of a feature repre-

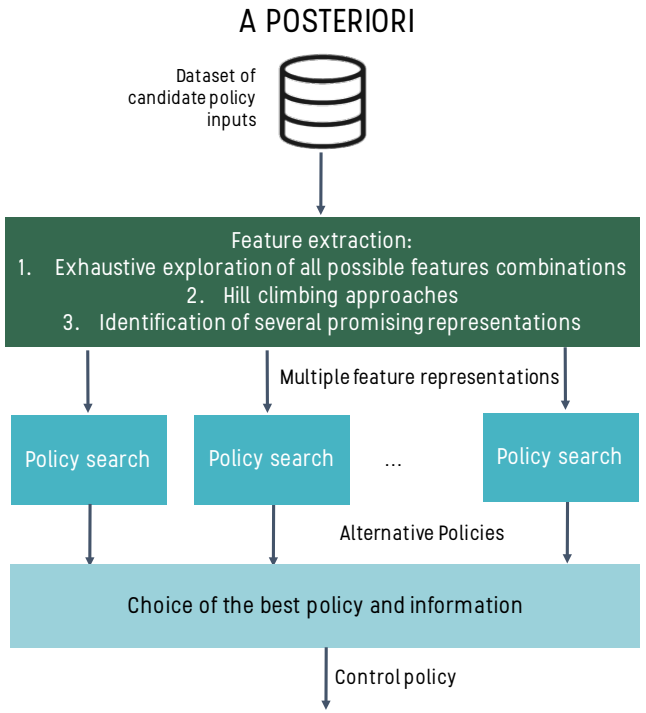

Figure 1.2: *Feature representation learning taxonomy, a posteriori approach: the optimal policy representation is selected based on the performance of the policy conditioned upon it.*

sentation by assessing the performance of the policy conditioned upon it. In general, multiple policies are designed with alternative feature representations, and the desired representation is identified as the one producing the best performing policy (Fig. 1.2).

In principle, the entire combinatorial space of features subsets could be exhaustively explored, yielding to an optimal solution. However, given a subset of f features, the possible representations are 2^f, which rapidly be-

comes computationally intractable for non-trivial datasets (see, e.g., Gaudel and Sebag, 2010).

Alternatively, several works have developed strategies to navigate the combinatorial space of feature combinations guiding the exploration in the direction of efficient subsets. For instance, hill-climbing approaches incrementally add features to the representation by testing a given minimal input set with each additional candidate feature, and retaining the most successful one (Wright et al., 2012). In some cases, this forward selection is mixed with backward selection (Zhang, 2009; Tan et al., 2013).

Finally, an initial *a priori* reduction can be applied to select a limited number of candidate representations that are then exhaustively compared *a posteriori*. Reduction techniques can once again be purely statistical, e.g., via autoencoders (Keller et al., 2006; Parr et al., 2007) or expert-based (Giuliani et al., 2016a; Castelletti et al., 2016).

In general, *a posteriori* feature representation is significantly more computationally burdensome than the a priori counterpart, with a computational effort that grows exponentially with the dimension of the candidate representation set. Yet, an exhaustive *a posteriori* search can be performed with virtually no pre-existing knowledge of the task, and guarantees the optimality of the derived feature representation.

- **Online.** In the on line approach, the FE step is simultaneous and codependent from PS, and an efficient policy representation is learned in conjunction to the policy (Fig. 1.3). In some cases, FE is completely embedded and indistinguishable from PS, which operates representation and policy learning simultaneously and automatically. For instance, evolutionary routines are employed to evolve populations of individual policies with diverse representations by selecting and mating the most promising ones until convergence (James and Tucker, 2004; Whiteson et al., 2005; Tan et al., 2012). Other works approximate value function and information set simultaneously (Kolter and Ng, 2009; Vigorito and Barto, 2009; Tangkaratt et al., 2016). While this fully embedded approach is demonstrated effective in some cases, the problem complexity it can handle is generally limited (Loscalzo et al., 2015).

  Alternatively, several authors suggest to separate representation and policy learning routines while maintaining interactions via a feedback loop that allows information exchange between these routines, as a way to lighten the search (Wahlström et al., 2015; Munk et al., 2016; Shelhamer et al., 2016; de Bruin et al., 2018; Lesort et al., 2018). This iterative *online* approach interleaves FE phases throughout the PS process, using progressively refined feature representations to support policy learning. In these

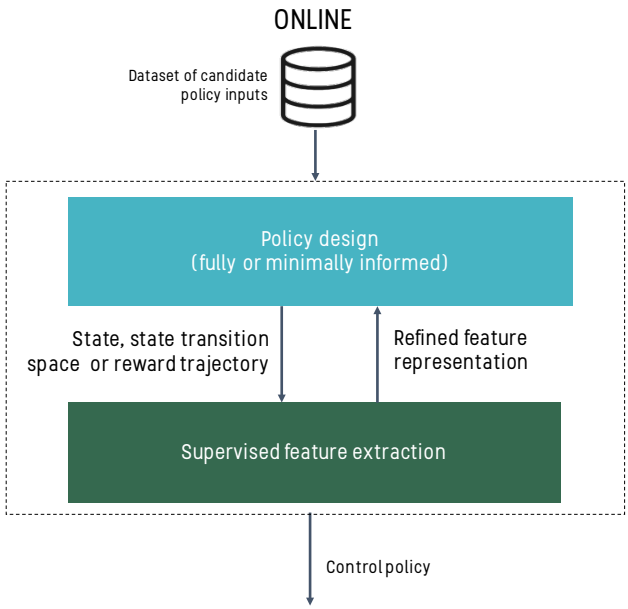

Figure 1.3: *Feature representation learning taxonomy, online approach: policy search is interleaved with feature extraction routines.*

routines, the policy representation is either progressively complexified, whenever it is possible to identify a meaningful minimal policy representation (e.g., Castelletti et al., 2011a), or progressively simplified from an initial representation comprising the entire input set (e.g., Van Hoof et al., 2016). Representations are updated during the search via supervised learning, by extracting features that approximate the state space (Curran et al., 2016; Alvernaz and Togelius, 2017), state-transition space (Assael et al., 2015; Van Hoof et al., 2016), or the reward trajectory (Munk et al., 2016; Oh et al., 2017) of the policy learned thus far (for a comprehensive review, see Lesort et al., 2018). The adjusted representation is then employed to refine policy search in a feedback loop between the two routines. While state-based feature extraction can identify variables that are relevant for environmental transitions, these can sometimes be irrelevant or distracting for the task, e.g., variables that change in response to external drivers, with no relation with the control task (Loscalzo et al., 2012). Reward-based selection is, instead, helpful to disentangle task-oriented features, ultimately deriving a more task-tailored representation (Lesort et al., 2018).

Computationally, iterative *online* methods require the PS process to be repeated or adjusted multiple times with refined representations, making

them more expensive than *a priori* methods. However, with respect to the *a posteriori* approach, they can handle much larger information sets by employing supervised feature extraction that is significantly faster than testing multiple policy input combinations.

1.2 Objective of the thesis

This thesis contributes to the literature of *a priori*, *a posteriori*, and *online* feature representation learning for the control problem of multi-purpose reservoir systems. The problem of defining optimal operating rules for water reservoirs has long been the object of studies. Since the seminal work of Rippl (1883), the field has been active for decades with several fundamental contributions (e.g., Yakowitz, 1982; Yeh, 1985; Simonovic, 1992; Wurbs, 1993; Labadie, 2004; Reed et al., 2013). However, this control problem still remains intellectually intriguing given the inherent difficulties emerging from its formulation and solution, ones that are rarely addressed in the broader literature of Reinforcement Learning. Among them, (see, e.g., Castelletti et al., 2008a) 1) high non-linearities emerge in the controlled system and the objective functions; 2) strong uncertainties and disturbances affect the system and cannot be neglected; 3) the problem formulation and framing is often non-univocal (Quinn et al., 2017); 4) multiple and conflicting demands coexist; 5) water systems are facing new unprecedented pressures related to climate change and demand growth. Most existing water systems are currently operated with static rules conditioned on basic feature representations including the day of the year and the reservoir storage, although there has long been consensus around the fact that enriching the information set could enhance the system performance (e.g., Kelman et al., 1990; Kim and Palmer, 1997a; Faber and Stedinger, 2001).

The coexistence of multiple control objectives is a major challenge to feature representation learning in multi-purpose water reservoir systems as common operating targets, e.g., flood protection and irrigation supply, can be vastly heterogeneous in their dynamics and vulnerabilities. Flood events are generally caused by the onset of fast and intense wet meteorological extreme events, while irrigation supply failures are the result of a prolonged period of water shortage caused by slow-developing dry hydrological extremes, i.e., droughts. In these systems, learning an appropriate feature representation becomes more intricate. On the one hand, a flood-conservative policy benefits from short-term information on the peak inflow magnitude and timing, on the other, a water supply-prone policy should be informed by the onset of a prolonged water shortage to activate hedging strategies. Moreover, the tradeoff space between these two opposite solutions is populated by an ensemble of policies that balances opposite control targets. One policy input set is thus inadequate to repre-

sent the entire space of tradeoffs in multipurpose systems, and feature representation should be searched tradeoff-dynamically, to appropriately characterize the entire set of alternative control behaviors.

When the problem complexity is high, *optimal* feature representation might become an elusive concept. In our contributions, instead of chasing optimality per se, we identify a specific system criticality, and we search a feature representation that has the potential to enhance the system resilience towards the targeted failure. In the following collection of works, we improve water system performance targeting drought emergencies, multisectoral conflicts, and critical phases in dam development (i.e., filling).

This thesis is organized in three parts, following the taxonomy of *a priori*, *a posteriori*, and *online* approaches. Each Part proposes novel methodological contributions, discussed in the relative introduction, including:

- FRIDA, FRmework for Index-based Drought Analysis, that automatically designs an index representing a surrogate of the drought conditions of a basin, computed by selecting and combining relevant information about the water circulating in the system via feature extraction. FRIDA is specifically targeted to highly regulated water systems, where natural water availability is conditioned by the operation of water infrastructures (e.g., dams, diversions, pumping wells) and traditional drought indicators fail in detecting critical events.
- A novel methodology that supports sustainable dam planning addressing the critical initial filling phase of large dams. The core novelty of this work consists in informing filling timing and operations anticipating hydrological variability by analysing the climate oscillations that affect the region.
- CSI, Climate State Intelligence, a framework designed to capture the state of multiple climate signals from global datasets of oceanic temperatures to improve seasonal forecasts. These forecasts are designed and employed to inform water system operations at the basin scale.
- NEMODPS, NeuroEvolutionary Multi-Objective Direct Policy Search, a novel DPS routine that conjunctively searches the policy functional class and its parameterization in a hyperspace containing policy architectures and coefficients. NEMODPS specifically addresses MO problems producing a tradeoff-dynamic architectural selection in one single run.
- AFS-NEMODPS, where NEMODPS is combined with an Automatic Feature Selection routine to conjunctively and progressively refine feature representation and policy design. Additionally, AFS-NEMODPS targets MO problems, by tailoring feature representation and policy architecture to different objective tradeoffs.

1.3 Thesis structure

This thesis is organized as follows:

- **Part 1.** In this Part we explore via *a priori* feature representation the potential of an unconventional feature representation that includes a measure of a basin's drought state. In the first contribution, we develop a framework for the automatic construction of basin-tailored drought indices via feature extraction, namely, FRIDA (Chapter 2). In the second, FRIDA is employed to design a drought index to inform the control of the Lake Como dam, in an expert-based, *a priori*, fashion (Chapter 3). Lastly, long term drought forecasts designed via big data analysis are applied to a problem of water reservoir planning and initial dam filling, expanding the concepts of informed problem representation beyond pure control applications (Chapter 4).

- **Part 2.** In Part 2 we compare, in *a posteriori* fashion, alternative feature representations that comprise raw, processed, and highly processed data, whereby highly processed data represent a more immediately interpretable information for policy design, but, on the other hand, are affected by higher modeling errors. The proposed analysis provides insights on how the level of input data processing can affect the resulting policy performance (Chapter 5).

- **Part 3.** This Part proposes a methodological contribution to iterative *online* feature representation tailored to MO problems in two aspects. First, we present a novel MO policy search routine that supports the online and tradeoff dependent optimization of the policy architecture, namely NEMODPS (Chapter 6). NEMODPS is also applied in the second contribution in combination with an Automatic Feature Selection routine to accomodate changes in the policy input set and learn a tradeoff-dynamic feature representation (Chapter 7).

- **Conclusions** In Chapter 8, we summarize the achievements of this PhD thesis, providing general conclusions, ideas and opportunities for further research.

Part I

A Priori Feature Representation for Multi-Objective Reinforcement Learning

This Part comprises three Chapters contributing to Multi-Objective (MO) *a priori* feature representation learning for a problem of control, and planning and control of multi-purpose water resources systems.

In the former, we explore the potential of enriching the information set conditioning the control policy of a water reservoir by accounting for the basin drought state. Droughts are responsible for water systems failures worldwide; however, there is no universal consensus on how to define the indicesnomena, or measure their duration and intensity. Additionally, in highly regulated water systems, drought detection is challenged by the interplay of processes with inconsistent dynamics and origins (natural or anthropic), making this indicesnomenon basin-specific (Mishra and Singh, 2011).

The first work proposed here (Chapter 2) contributes a novel FRamework for Index-based Drought Analysis (FRIDA), for the design of basin-specific drought tailored to highly-regulated water systems. The framework is validated with a benchmark analysis against the successful institutional index-based drought management system of the Spanish Jucar basin.

In the second work (Chapter 3), FRIDA is applied to the case study of the Lake Como system, in northern Italy, and the resulting drought index demonstrates good detection skills. Lake Como, operated for both flood protection and irrigation supply, suffered unprecedented and recurrent drought events in the last several years that disproportionately damaged the irrigation sector. This situation is likely to worsen in the future as a result of an ongoing climate change trend that amplifies dry extremes in the region. In an *a priori* expert-based feature representation learning experiment, we re-designed the control of the lake's dam by including the drought index in the control policy representation. While we have no guarantees of its optimality, a drought index is a promising asset for policy design in the region, as it distills a multiplicity of hydro-meteorological processes into a single value representing their contribution towards droughts, and directly targets the cause of the system's recent failures. The performance of the drought-informed policy is contrasted with that of a control policy relying on a basic representation, showing a Pareto front-wide improvement that has the potential to inform drought management strategies and reduce the conflicts between irrigation and flood sectors, yielding cross-sectoral benefits.

The third work expands the concepts of feature representation beyond pure control applications, to a problem of dam planning and initial filling. When a new dam is built, the reservoir behind it is filled withholding a substantial volume of water from downstream users, often generating critical water shortages and upstream-downstream tensions (Chapter 4). Previous studies on dam filling considered standard non-adaptive strategies designed to impound (or release) a fixed reservoir inflow percentage or absolute volume (King and Block,

2014; Wheeler et al., 2016). However, in climates that are highly affected by climatic oscillations, hydrological variability plays a key role: if the filling occurs during a drought, enhanced impacts are experienced by all sectors. The novelty of our work consists in expanding the basic dam planning and filling problem representations by considering information on long and medium term hydroclimatic variability. In particular, we consider the case study of the Omo-Turkana Basin, where the impacts of teleconnections have long been recognized (e.g., Lanckriet et al., 2015). In a retrospective analysis of the recent filling of Gibe III dam, we found how enriching the planning and filling problem representation shows great potential in reducing basin-wide conflicts.

This Part contains the following works:

- Zaniolo, M., Giuliani, M., Castelletti, A.F., Pulido-Velazquez, M., 2018b. Automatic design of basin- specific drought for highly regulated water systems. Hydrology and Earth System Sciences 22, 2409-2424;
- Zaniolo, M., Giuliani, M., Castelletti, A., 2019. Data-driven modeling and control of droughts. IFAC- Papers On Line 52, 54-60;
- Zaniolo, M., Giuliani, M., Burlando, P., Castelletti, A., 2020a When timing matters - misdesigned dam filling impacts hydropower sustainability. Nature Sustainability (under review).

2

Automatic design of basin-specific drought indexes for highly regulated water systems

Abstract[1]

Socio-economic costs of drought are progressively increasing worldwide due to undergoing alterations of hydro-meteorological regimes induced by climate change. Although drought management is largely studied in the literature, traditional drought indexes often fail in detecting critical events in highly regulated systems, where natural water availability is conditioned by the operation of water infrastructures such as dams, diversions, and pumping wells. Here, ad-hoc index formulations are usually adopted based on empirical combinations of several, supposed-to-be significant, hydro-meteorological variables. These customized formulations, however, while effective in the design basin, can hardly be generalized and transferred to different contexts. In this study, we contribute FRIDA (FRamework for Index-based Drought Analysis), a novel framework for the automatic design of basin-customized drought indexes. In contrast to ad-hoc, empirical approaches, FRIDA is fully-automated, generalizable, and portable across different basins. FRIDA builds an index representing a sur-

[1]This work has been published as: Zaniolo, M., Giuliani, M., Castelletti, A.F., Pulido-Velazquez, M., 2018b. Automatic design of basin- specific drought indexes for highly regulated water systems. Hydrology and Earth System Sciences 22, 2409-2424

rogate of the drought conditions of the basin, computed by combining all the relevant available information about the water circulating in the system identified by means of a feature extraction algorithm. We used the Wrapper for Quasi Equally Informative Subset Selection (W-QEISS), which features a multi-objective evolutionary algorithm to find Pareto-efficient subsets of variables by maximizing the wrapper accuracy, minimizing the number of selected variables, and optimizing relevance and redundancy of the subset. The preferred variable subset is selected among the efficient solutions and used to formulate the final index according to alternative model structures. We apply FRIDA to the case study of the Jucar river basin (Spain), a drought-prone, highly regulated Mediterranean water resource system, where an advanced drought management plan relying on the formulation of an ad-hoc State Index is used for triggering drought management measures. The State Index was constructed empirically with a trial-and-error process begun in the '80s and finalized in 2007, guided by the experts from the *Confederación Hidrográfica del Júcar* (CHJ). Our results show that the automated variable selection outcomes align with CHJ's 25 years-long empirical refinement. In addition, the resultant FRIDA index outperforms the official State Index in terms of accuracy in reproducing the target variable and cardinality of the selected inputs' set.

2.1 Introduction

A drought is a slowly-developing natural phenomenon that occurs in all climatic zones and can be defined as a temporary significant decrease of water availability (Tallaksen and Van Lanen, 2004; Van Loon and Van Lanen, 2012). Drought impacts can propagate to virtually every water-related sector, such as farming and livestock production, industry, power generation, and public water supply (Spinoni et al., 2016). During the period 1976-2006, droughts in Europe affected more than 11% of the population, and their economic cost was estimated to exceed €100 billion, considering damages endured by consumers, tourism, industry, energy, and agricultural sectors. Moreover, climate change is expected to produce longer, more frequent and severe drought events, especially in southern Europe (Giorgi and Lionello, 2008; Spinoni et al., 2016; Marcos-Garcia et al., 2017). Recent drought cost trends show a significant increasing tendency, reaching an average of €6.2 billion/year in the years 1991-2006 (EU, 2007). These estimates, however, only account for the economic damages, (i.e., situations in which a water deficit induced by droughts affects production, sales and business in a variety of sectors), neglecting environmental and social costs (Spinoni et al., 2016). A comprehensive quantification of drought impacts is, in fact, complicated by the considerable lag occurring between the realization of dry climatic conditions and the impacts on economy

and society (Changnon, 1987; Stahl et al., 2016).

We can distinguish four types of droughts: meteorological, agricultural, hydrological, and operational (or anthropogenic) drought, depending on the time horizon and the variable of interest. (Heim Jr, 2002; Mishra and Singh, 2010; Pedro-Monzonìs et al., 2015; Spinoni et al., 2016). The development chain of droughts through time is exemplified in Figure 2.1.

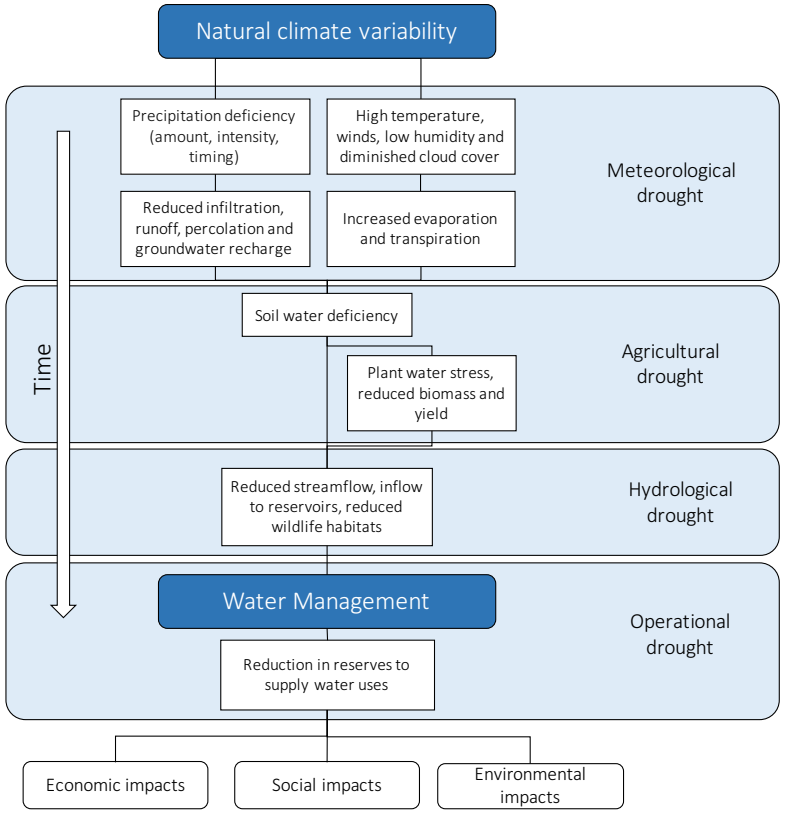

Figure 2.1: *Development chain of droughts through time. Meteorological drought, defined as a lack of precipitation over a region for a certain period of time, develops in the short term. Agricultural drought accounts for the plants and crops water stress; develops in the medium term. Hydrological drought, defined as a period of low streamflow in watercourses, lakes and groundwater level below normal, develops in the long term. Operational drought, defined as a period with anomalous supply failures in a developed water exploitation system, develops in the long term. Figure adapted from Spinoni et al. (2016) to include Operational drought.*

A meteorological drought is defined as a lack of precipitation over a region for a certain period of time (Mishra and Singh, 2010). It develops over the short term (1-3 months) and can extend on longer periods, and is usually associated

with the global behavior of the atmospheric circulation (Pedro-Monzonìs et al., 2015). Precipitation is always the core variable to characterize this drought type, with most meteorological drought indexes based on precipitation only (Byun and Wilhite, 1999; McKee et al., 1993). In some cases, especially in regions where droughts can be strongly influenced by evapotranspiration, additional variables such as temperature trends are also considered (Vicente-Serrano et al., 2010; Lorenzo-Lacruz et al., 2010).

Agricultural drought affects, and is defined through, the state of soils and crops in the medium term (3-6 months) (Pedro-Monzonìs et al., 2015). This drought type manifests itself with dryness in the root zone and, although rainfall deficiency is a primary cause, precipitation alone is often not enough to describe it. Approaches to characterize agricultural droughts focus on monitoring soil water balance and the subsequent deficit (Palmer, 1965; Narasimhan and Srinivasan, 2005; Hao and AghaKouchak, 2013). The factors involved in this case include vegetation type, soil water holding capacity, wind intensity, evapotranspiration rate, and air humidity (Heim Jr, 2002). In regulated systems, agricultural droughts can be usually restrained with irrigation (Keyantash and Dracup, 2002).

Hydrological drought is defined as a period of exceptionally low flows in watercourses, and lakes and groundwater levels below normal (Dracup et al., 1980; Van Loon and Van Lanen, 2012). Related indicators mainly focus on streamflow, as the by-product of every hydro-meteorological process taking place in water catchments (Heim Jr, 2002; Vicente-Serrano and López-Moreno, 2005). More comprehensive indexes can also include snowpack extent, reservoir storage, and groundwater level (Shafer and Dezman, 1982; Keyantash and Dracup, 2004; Staudinger et al., 2014). This drought takes place after a prolonged time of low precipitation and deficient soil moisture and its effects are witnessed in the long-term (6-12 months) (Zargar et al., 2011).

These three categories refer to droughts as a natural hazard, i.e., a threat of a naturally occurring event that negatively effects people or the environment (Gustard et al., 2009; Van Loon and Van Lanen, 2013; Laaha et al., 2016). On the other hand, particularly in highly regulated contexts, a dry spell may be caused by natural scarcity of precipitation as well as inconsiderate overuse and/or mismanagement of water resources. Another interesting way to approach drought analysis is, therefore, through the concept of operational (or anthropogenic) drought. Operational drought is defined as a period with anomalous supply failures in a developed water system (Pedro-Monzonìs et al., 2015). It is caused by a combination of two factors: lack of water resources and excess of demand (Mishra and Singh, 2010; AghaKouchak, 2015b). Moreover, it can be further worsened by an inadequate design and management of the water exploitation system and its operating rules (Mishra and Singh, 2010). Operational droughts

indicators aim at comparing water availability to human water needs and serve as a measure of water well-being, rather than a measure of natural fluctuation as in the case of meteorological, agricultural, and hydrological indicators (Sullivan et al., 2003; Rijsberman, 2006). In the computation of operational drought indicators, the available water is often represented by the streamflow, or a fraction of it, and the water need is usually quantified by a standard per capita or by a fixed nominal demand (Falkenmark et al., 1989; Raskin et al., 1997). Depending on the application scope, operational drought indicators are either river basin specific (Garrote et al., 2007; Haro-Monteagudo et al., 2017) or used in studies covering continental or global areas with an annual time resolution (Yang et al., 2003; Oki and Kanae, 2006; Alcamo et al., 2007; Kummu et al., 2010).

When considering a highly regulated water system, i.e., a system where natural water availability is altered by the presence and operation of water infrastructures, traditional drought indicators (e.g., SPI, Standardized Precipitation Index; SPEI, Standardized Precipitation and Evapotranspiration Index; SRI, Standardized Runoff Index) present different shortcomings. On the one hand, meteorological, agricultural, and hydrological indexes often fail in representing drought conditions when regulated lake releases and/or groundwater pumping filter water availability and play a role in magnifying or smoothing drought impacts. Anthropized systems have, in fact, a demonstrated ability to endure meteorological droughts for months, or even years, without suffering consequences, i.e., without incurring in a situation of water shortage perceived by the users. An effective planning and management of water resources enables such systems to wisely exploit the combined storage capacities of surface and groundwater reserves and restrain drought (Rijsberman, 2006; Haro et al., 2014a). On the other hand, operational drought indexes are often designed to operate analysis over coarse spatiotemporal resolutions, thus resulting unsuitable for a real time basin level drought detection, characterization, and management. Highly regulated systems need ad hoc index formulations tailored on basin characteristics (Wanders et al., 2010; AghaKouchak, 2015a), combining human-controlled variables (e.g., reservoirs and groundwater levels) with uncontrolled hydro-meteorological variables (e.g., precipitation, temperature, natural inflows) to reflect both regulation effects and natural fluctuations in the basin.

A paradigmatic example of a practical and systematic policy for the identification and mitigation of operational droughts is provided by Spain, where public River Basin Management authorities (Confederaciones Hidrográficas) are bind by Law (Ministerio del Medio Ambiente, 2000) to design basin-specific State Indexes associated with each main river basin (Ie, *Índice de Estado*). Most of the basins in Spain are highly regulated and these State Indexes are computed as a weighted average of relevant observed variables at selected control

points, e.g., precipitation, streamflow, reservoir level, and groundwater level. Each river basin authority has designed its customized formulation for the State Index which reflects the hydroclimatic conditions and the water uses of the region (Estrela and Vargas, 2012). The value of the State Indexes is monitored monthly and used to trigger water demand and supply measures when entering a drought period, according to the district Drought Management Plan (DMP) (Garrote et al., 2007; Gómez and Blanco, 2012; Haro et al., 2014a).

Each DMP and the relative State Index formulation is the result of a long collaborative process including public participation, basin experts, and stakeholders, and providing an effective multi-sector partnership approach for managing drought risk (Carmona et al., 2017). State Indexes are the result of a long trial-and-error process mostly begun in the eighties, through which the variable choice and combination have been progressively adjusted to best suit the basin drought management requirements. In the case of the Jucar basin, for instance, the final form of the associated index was established in 2007 with a report by the *Confederación Hidrográfica del Júcar* (CHJ, 2007b), after 25 years of refinements. This long empirical process produced an index formulation tailored for the Jucar system, which cannot be generalized to different contexts. Similarly, other main Spanish river basins (e.g., Duero, Ebro, and Guadalquivir river basins) underwent an analogous process and formulated their own State Indexes (CHD, 2007; CHE, 2007; CHG, 2007).

Since their establishment in 2007, State Indexes have represented the most consistent and extensively applied paradigm of index-based drought management. Thus, I_es constitute the state of the art for basin-customized operational drought indexes. A reasonable research question is whether the empirical process leading to their design can be formalized, automated, and easily exported to different water systems.

In this study, we contribute the FRamework for Index-based Drought Analysis (FRIDA), which allows the automatic construction of basin-customized drought indexes for highly regulated water systems. In contrast to traditional empirical approaches, FRIDA uses an advanced feature extraction method that completely automatizes and generalizes the variable selection process for the construction of the index. The selected variables are then combined into a new index that can effectively represent the state of water resources in the basin as well as support the characterization of drought conditions. The feature extraction step is key in FRIDA as it guides the construction of a skillful (highly accurate) and parsimonious (with low input dimensionality) drought index by performing the selection of the best input subset to build a model of a predefined target output representing the drought conditions in the basin.

Specifically, FRIDA is structured in three steps. First, we define a target variable, an appropriately chosen water deficit acting as a proxy for the drought

conditions of the considered basin (e.g., water supply deficit, soil moisture deficit), and a dataset of hydro-meteorological variables and traditional drought indicators. Second, we identify Pareto optimal subsets of variables balancing predictive accuracy and parsimony. In this study, we employed the Wrapper for Quasi-Equally Informative Subset Selection (W-QEISS) to perform this operation (Karakaya et al., 2015; Taormina et al., 2016). Traditional variable selection algorithms are conceived to select only one optimal subset of predictors, while W-QEISS identifies one subset with the highest predictive accuracy, and multiple subsets with similar information content, thus providing more informative results. Moreover, W-QEISS includes two metrics of relevance and redundancy in the search process in addition to the commonly used objectives of accuracy and cardinality, fostering the diversification among the provided solutions (Sharma and Mehrotra, 2014). Third, we choose the preferred predictor subset among the non-dominated solutions based on accuracy, cardinality (i.e., dimensionality), and, possibly, additional factors, including cost and availability of the variable observations. The subset is finally used to calibrate a chosen model class with respect to the target variable, and the drought index is thus completed.

The potential of the proposed framework is demonstrated on the highly regulated Mediterranean basin of the Jucar river, in eastern Spain, where the State Index-based drought management system provides an ideal benchmark for testing FRIDA index (Andreu et al., 2009; Haro et al., 2014b; Pedro-Monzonís et al., 2014; Macian-Sorribes and Pulido-Velazquez, 2017; Haro-Monteagudo et al., 2017; Carmona et al., 2017). The Jucar State Index provides guidelines for FRIDA application. First, it facilitates the target variable choice and candidate variable retrieval, and, second, it allows the validation of FRIDA predictors selection, and index design steps. FRIDA and State indexes are compared in terms of accuracy in reproducing the drought conditions of the basin, number of variables required for their computation, and general reliability and portability of the methods. The outcome of this analysis consists in demonstrating the validity of a completely automated procedure (i.e., no information on system topology or basin characteristics is required) in recognizing the main drought drivers, and predicting a deficit with accuracy and limited computational effort.

2.2 Methods and tools

2.2.1 Framework for Index-based Drought Analysis

The FRamework for Index-based Drought Analysis (FRIDA) designs drought indexes in three steps as reported in Figure 2.2.

The Identification of basin characteristics is a preliminary empirical process, which consists in the selection of a target variable and the collection of candi-

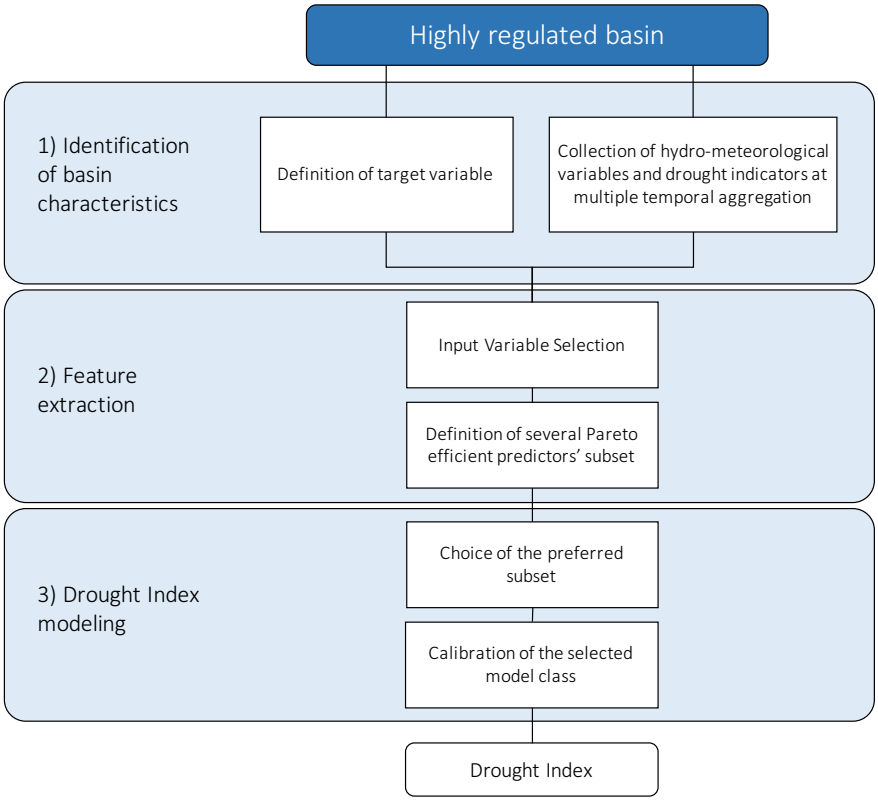

Figure 2.2: *FRamework for Index-based Drought Analysis (FRIDA): 1. Identification of basin characteristics, 2. Feature Extraction, 3. Drought Index modeling.*

date predictors. The target variable is an appropriately chosen water deficit, representative of the actual drought conditions in the basin (e.g., water supply deficit, soil moisture deficit). The dataset of predictors contains the candidate features to reproduce the target variable and consists of observed hydrometeorological variables and composite drought indicators over different spatio-temporal scales.

Target variable and candidate predictors constitute the input to the Feature Extraction step, the second building block of the framework. This block employs an Input Variable Selection (IVS) algorithm that explores the space of candidate predictors to select Pareto efficient subsets of predictors with respect to multiple assessment metrics. Most commonly, these metrics quantify the subset accuracy in reproducing the target and the parsimony (i.e. the cardinality of the subset), crucial characteristics for an operational index expected to balance precision and ease-of-use. In some cases, also relevance and redundancy can be considered in order to explore the input space more effectively. In particular, the metric of relevance favors highly informative subsets (i.e., constituted by predictors that are highly correlated with the target), while the redundancy metric ensures low intra-subset similarity. The objectives of relevance and redundancy are essential to stimulate the search process towards the identification of a diversified and comprehensive set of solutions, which would not be achieved optimizing cardinality and accuracy only.

In this work, we use an advanced IVS algorithm called Wrapper for Quasi-Equally Informative Subset Selection (W-QEISS). W-QEISS provides as output a number of efficient subsets that are collected in a Selection Matrix.

In the Drought Index modeling block, the preferred efficient solution is selected by the user, balancing the trade-off between competing objectives, and, possibly, considering additional operative needs neglected in the IVS search (e.g., cost and reliability of the variable monitoring). Lastly, an appropriate regressor is fit to the sample data set of Pareto efficient inputs and the target variable. The choice of model class is determined by the application of interest. In general, highly non-linear learning machines like Artificial Neural Networks (ANNs) provide a good balance between accuracy and flexibility. On the other hand, such black-box models lack of intuitive interpretability and might result unsuitable for applications that affect several stakeholders and require a wide acceptance of the tool to be employed (Estrela and Vargas, 2012). In these cases, a simpler model (e.g., a linear model) might be preferred, as it grants an immediate understanding of the physical meaning, though at the price of poorer approximation skills.

2.2.2 Feature Extraction via Wrapper for Quasi-Equally Informative Subset Selection

Feature extraction techniques, employed in the second block of FRIDA, are an ensemble of data pre-processing algorithms that transform the original input data set into a more compact, while still highly informative, subset (Cunningham, 2008). Among the feature extraction algorithms, Input Variable Selection (IVS) techniques specifically address the problem of the reduction of the input space by identifying the relevant predictors to be used to calibrate a model of the target variable (Bowden et al., 2005). There are two main classes of IVS techniques: Filters and Wrappers. Filters evaluate the relevance of each variable separately, computing an error metric on the features (Yang and Pedersen, 1997; Sharma, 2000a; Galelli and Castelletti, 2013). Wrappers, on the other hand, assess the relevance of a variables ensemble, evaluating the prediction performance of a given learning machine calibrated on the input set, and thus considering the interactions and dependencies between variables (Guyon, 2003). In terms of performance, Wrappers are often more accurate than Filters, although computationally more intensive (Galelli et al., 2014).

In this study, we used the Wrapper for Quasi-Equally Informative Subset Selection (Karakaya et al., 2015; Taormina et al., 2016). The W-QEISS algorithm receives as input the set $\mathbf{X}$ of candidate predictors, i.e., $\mathbf{X} = \{x_i, \ldots, x_{n_X}\}$ and the trajectory y of the target variable. The algorithm is composed of three main steps (Karakaya et al., 2015), as synthesized in Figure 2.3:

- Step 1: a set $\mathbf{A} \subseteq \mathbf{X}$ of Pareto-efficient solutions is built according to the four-objective functions of relevance $f_1(\cdot)$, redundancy $f_2(\cdot)$, cardinality $f_3(\cdot)$, and accuracy $f_4(\cdot)$. A global multi-objective optimization algorithm is employed to explore the space of the possible subsets. In this study, we use the self-adaptive Borg MOEA (Hadka and Reed, 2013), which has shown to outperform other benchmark evolutionary algorithms in terms of number of solutions returned, ability to handle many-objective problems, ease-of-use, and overall consistency across a suite of challenging multi-objective problems (Reed et al., 2013). A learning machine is used to compute the predictive accuracy f_4 of each set. In this study, we employ the Extreme Learning Machines (ELMs) (Huang et al., 2006a), belonging to the family of Artificial Neural Networks, which were shown to provide a good performance in terms of accuracy and flexibility in a variety of problems while resulting up to thousand times faster than benchmark feed-forward ANNs (Huang et al., 2012). ELMs, in fact, bypass the time consuming gradient-based search of optimal neurons parameters required by traditional ANN techniques, by defining randomly parameterized hidden nodes, and subsequently optimizing their output weights. Such optimization is solved through a one-step matrix product and essentially amounts

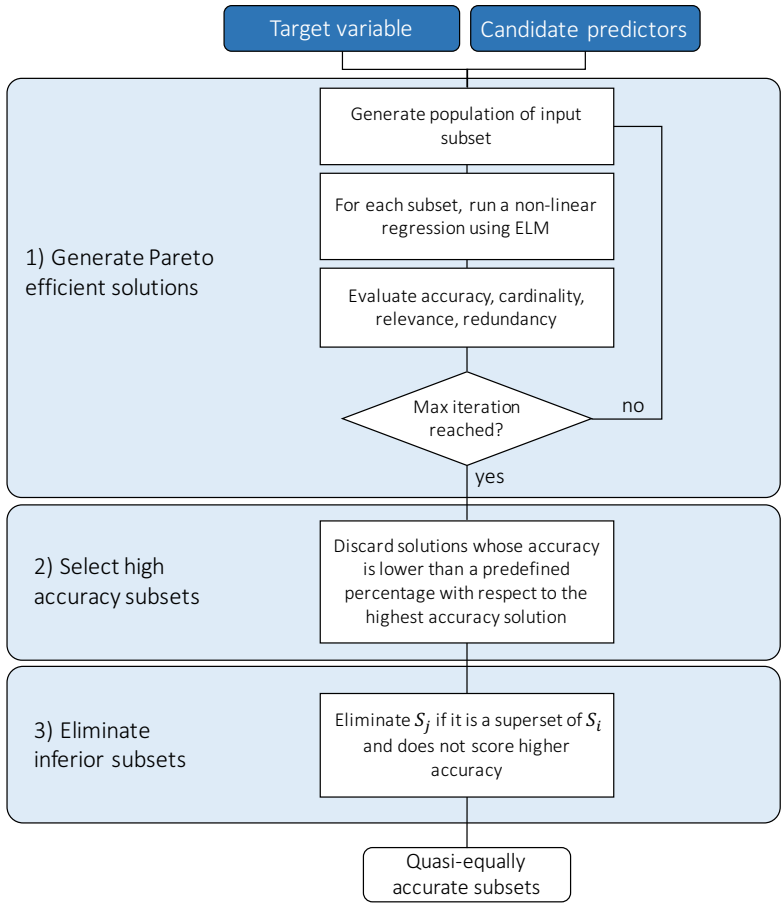

Figure 2.3: *W-QEISS flowchart. Step 1: generate Pareto efficient solutions with respect to the four objectives of relevance, redundancy, cardinality, and accuracy; Step 2: select high accuracy subsets; Step 3: eliminate inferior subsets.*

to learning a linear model.

However, we do not expect the choice of the learning machine or MOEA to be crucial for the attainment of the result. A different benchmark MOEA (e.g., NGSAII, MOEAD, eps-MOEA) is likely to achieve a comparable result, although requiring a possibly significant effort in the manual calibration of the evolution parameters, which is automated in Borg MOEA. Similarly, other ANN techniques could in principle be substituted to ELM, although incrementing the computational time to possibly unbearable levels, given the multiple calibration and validation processes reiterated in WQEISS.

- Step 2: Among the Pareto-efficient subsets, the maximum value of accuracy f_4^* is identified, associated with subset $\mathbf{S}_{f_4^*} \subseteq \mathbf{A}$. Then, solutions with significantly lower accuracy are discarded and from ensemble $\mathbf{A}$, obtaining $\mathbf{A}_\epsilon$. The ensemble $\mathbf{A}_\epsilon$ contains quasi-equally informative subsets with respect to $\mathbf{S}_{f_4^*} \subseteq \mathbf{A}_\epsilon \subseteq \mathbf{A}$, i.e., subsets that have (almost) the same predictive accuracy with respect to a given model class. When the dataset of candidate variables presents significant correlation among features, numerous subsets characterized by a wide range of cardinalities are generally available to achieve a relative small range of accuracies. This is often the case in environmental problems, where spatial and temporal correlation of hydro-meteorological variables and associated indicators is significant. Therefore, at this stage, the accuracy metric is used to retain accurate solutions only, provided that they feature different cardinalities and predictors combinations.
  Formally, on the basis of an predefined small value of δ, $\mathbf{S}_i$ is δ-quasi equally informative to subset $\mathbf{S}_{f_4^*}$ if

$$f_4(\mathbf{S}_i) \geqslant (1-\delta) f_4^* \qquad \text{for} \quad 0 \leqslant \delta \leqslant 1 \tag{2.1}$$

- Step 3: The final ensemble $\mathbf{A}_\delta^*$ is computed after the elimination of the inferior subsets. The subset $\mathbf{S}_j$ is considered inferior to $\mathbf{S}_i$, if it is a superset of $\mathbf{S}_i$, and does not score higher accuracy. Formally
  $\mathbf{S}_i \subset \mathbf{S}_j$ and $f_4(\mathbf{S}_i) \geqslant f_4(\mathbf{S}_j)$.
  In this step, all subsets contained in $\mathbf{A}_\delta$ are compared in order to find possible inferior subsets and eliminate them. By doing this, the final ensemble of δ-quasi equally informative subsets $\mathbf{A}_\delta^*$ is provided as output of the procedure and reported in a Selection Matrix.

The W-QEISS algorithm differs from a traditional IVS approach as it introduces the consideration that, for a given cardinality, multiple subsets of variables can have almost indistinguishable accuracy performance. The outcome of W-QEISS variable selection is thus not a single most accurate subset for each

cardinality, but a pool of δ-quasi equally accurate solutions among which the preference can be determined by other metrics not directly considered in the optimization (e.g., cost and reliability of the variable observation).

Another innovative feature of the W-QEISS approach relies on the formulation of a four objective optimization problem. Beside the two traditional objectives of accuracy ad complexity commonly employed in Wrappers, W-QEISS includes other two metrics of relevance and redundancy (Sharma and Mehrotra, 2014). The maximization of accuracy ensures a precise reproduction of the data, while the minimization of cardinality aims at simplifying the final models. These characteristics are key for an operational index, expected to balance precision and ease-of-use. Relevance and redundancy optimization is instead an asset for an effective subset search process, as it fosters the diversification of the solutions explored within the MOEA algorithm, guaranteeing low intra-subset similarity, and high information content of the solutions. A two-objective search based on cardinality and accuracy only would, in fact, identify optimal solutions, but at the same time disregard a number of quasi-equally informative subsets with an almost identical operational behavior. The identification of such alternative solutions, nevertheless, grants flexibility and a multiplicity of options for the expert-based choice of the preferred subset, where certain combinations of predictors can be favored according to case-specific operative purposes, e.g., more robust or less costly data gathering process, enhanced acceptability or immediacy of the index.

Three of the four objectives formulations make use of the Symmetric Uncertainty (SU), a measure of the dependence and similarity between two variables (Witten and Frank, 2005). SU assumes values between 0 (independent variables) and 1 (complete dependence) and is computed for two features A and B as:

$$SU(A,B) = \left[\frac{2 \cdot (H(A) + H(B) - H(A,B))}{H(A) + H(B)}\right] \tag{2.2}$$

where $H(\cdot)$ is the entropy of variable $(\cdot)$ (see for instance Scott (2012) for the definition).

WQEISS bases its objectives formulation on information theory, as discussed in Karakaya et al. (2015). Information theoretic criteria (e.g., SU, Mutual information, and Partial Mutual Information) do not assume any functional relationship between the variables and thus result well suited to quantify the dependence between two variables in any modeling context (MacKay, 2003). Other objectives formulations could in principle be explored, for instance substituting the use of Symmetric Uncertainty with more traditional correlation coefficients, although with the risk of losing generality by assuming linear dependence between variables.

The four assessment metrics are formulated as follows:

1. Relevance $f_1(\mathbf{S})$: to be maximized, is formulated as:

$$f_1(\mathbf{S}) = \sum_{x_i \in \mathbf{S} \subseteq \mathbf{X}} SU(x_i, y) \tag{2.3}$$

where the term $SU(x_i, y)$ represents the symmetric uncertainty between the feature x_i and the output y. The relevance is therefore a measure of the explanatory power of the features with respect to the output.

2. Redundancy $f_2(S)$: to be minimized, is formulated as:

$$f_2(\mathbf{S}) = \sum_{x_i \in \mathbf{S} \subseteq \mathbf{X}} SU(x_i, x_j) \tag{2.4}$$

where $SU(x_i, x_j)$ represents the SU between two features x_i and x_j. High redundancy thus means high similarity between the features. By minimizing the redundancy the algorithm ensures that the search will be oriented towards the selection of subsets with mutually dissimilar features.

3. Cardinality $f_3(\mathbf{S})$: to be minimized, is formulated as:

$$f_3(\mathbf{S}) = |\mathbf{S}| \tag{2.5}$$

where $|\mathbf{S}|$ is the number of predictors within the subset. Its minimization guarantees that the resulting model will not be unnecessarily complex.

4. Accuracy $f_4(\mathbf{S})$: to be maximized, is formulated as:

$$f_4(\mathbf{S}) = SU(y, \widehat{y}(\mathbf{S})) \tag{2.6}$$

where $SU(y, \widehat{y}(\mathbf{S}))$ is the correlation, measured in SU, between the observed output y and the prediction $\widehat{y}(\mathbf{S})$ obtained from the model.

2.3 Case Study: the Jucar river basin

The Jucar river basin occupies an area of 42,989 km^2 located in the eastern part of Spain (see Figure 2.4). The territory is mainly mountainous in the interior part, while the center-eastern section shows a vast plain system ending into the Mediterranean sea. The territory is characterized by various climatic conditions of which sub-humid and semi-arid are largely dominating. The main rivers of the area are Jucar, Mijares, and Turia, covering all together more than 80% of the total mean areal flow. The subterranean runoff is very relevant, providing 74% of the contribution to the river network (CHJ, 2007b).

Since the mean value of the total annual runoff (1,747 Mm^3 from 1940 to 2009) almost equals the annual water demand (1,640 Mm^3), water scarcity and

droughts have long been perceived as primary issues for agricultural, social, economic, and environmental reasons. On the other hand, meteorological droughts in the Jucar basin can be endured for several years without suffering any consequences, due to the highly regulated water system set in the area. There are three main large surface reservoirs in the region: Alarcón, Contreras, and Tous (maximum capacity: 1,118 Mm^3, 444 Mm^3, and 378.6 Mm^3, respectively). In addition, most aquifers in the basin are intensively exploited to support agricultural supply and are currently experiencing a significant depletion due to over-drafting, which, in turn, affects the rivers flow.

In such a highly regulated basin with long overyear storage, water scarcity is not a necessary condition derived from a meteorological drought (CHJ, 2007b; Carmona et al., 2017). Thus, traditional drought indexes fail in detecting the timing and severity of the incidence of a drought, and an ad-hoc monitoring system was conceived to properly capture the hydrological status of the catchment. The monitoring system is based on the formulation of a basin specific index, namely the State Index ($\mathbf{I}_e$, *Índice de Estado*). The State Index was constructed empirically by the Jucar river basin authority (CHJ), with the intent of highly correlate to water scarcity conditions in the basin, in order to support drought management and the implementation of the actions considered in the Drought Management Plan (CHJ, 2007b). For that purposes, the index is developed after identifying the water sources for every main demand in the basin and the selection of representative variables to characterize the status of those sources.

The total State Index $\mathbf{I_e}$ is computed as a weighted mean of 12 partial I_e. Partial I_es are obtained by normalizing hydro-meteorological indicators (V_i) belonging to the following categories (see Figure 2.4):

1. The mean monthly storage of one, or more reservoirs combined [Mm^3] (2 storage *indicators*);
2. The mean streamflow contribution of the last 3 months [Mm^3] (4 flow *indicators*);
3. The mean monthly piezometric level [m] (3 piezometer *indicators*);
4. The areal precipitation of the last 12 months [mm], computed averaging the values observed by multiple pluviometers (3 precipitation *indicators*).

Each *indicator* (*Vi*) is consequently normalized to obtain 12 partial I_e values:

$$I_e = \begin{cases} \frac{1}{2}\left[1 + \frac{Vi - Vm}{Vmax - Vm}\right] & \text{if } Vi \geqslant Vm \quad (2.7a) \\ \frac{Vi - Vmin}{2\,(Vm - Vmin)} & \text{if } Vi < Vm \quad (2.7b) \end{cases}$$

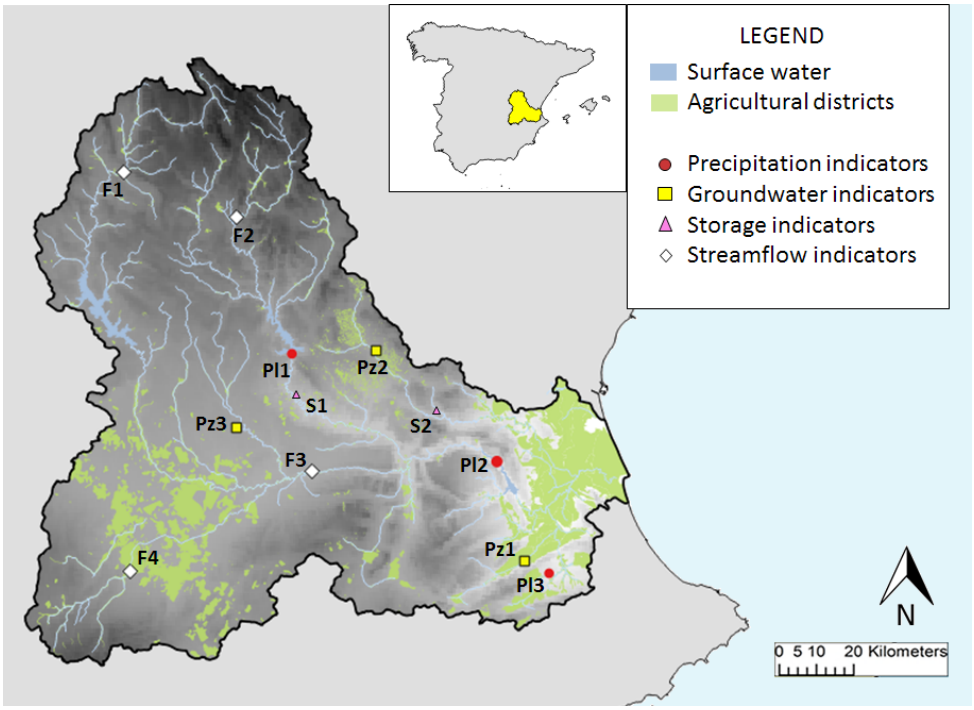

Figure 2.4: *Map of the Jucar Basin river network. The colored markers represent the variables considered for the State Index calculation. S: reservoir storage, F: streamflow, Pz: piezometer, Pl: pluviometer. Streamflow and piezometers markers are located in correspondence to the relative measurement station, while storage and pluviometers markers are put in the center of the polygon formed by connecting the multiple measurement points used for their computation.*

where *Vm*, *Vmax* and *Vmin* are the mean, maximum, and minimum values of each indicator time series. The storage and precipitation monthly time series are normalized with respect to maximum and minimum values of the considered month, while piezometers and river flows are normalized with respect to the complete historical time series. The partial I_es result as normalized indexes between 0 and 1, where $I_e > 0.5$ indicate higher than average value of *Vi*. Once the partial I_e have been computed, they are aggregated as a weighted sum to obtain the total $\mathbf{I_e}$. The weights are established according to the demand class associated to the indicator, ranging from class A (demand > 100 hm^3/year) to D (demand < 10 hm^3/year).

The Jucar river basin represents a Mediterranean drought prone highly regulated basin, featuring one of the most innovative and effective drought management systems, relying on the formulation of an empirically constructed basin specific drought index (Andreu et al., 2009; Haro et al., 2014b; Haro-Monteagudo et al., 2017; Carmona et al., 2017). As a consequence, it represents the state of the art for basin-customized operational drought indexes employed for drought restraining purposes, and a remarkable benchmark to test and validate the proposed FRIDA methodology.

2.4 Numerical results

For the presentation of the numerical results we follow the workflow proposed in Figure 2.2 . The length of the dataset available for the experiments is $N = 174$ data points, corresponding to monthly values in the period 1986-2000, and $n_x = 28$ number of candidate predictors were used (Zaniolo et al., 2018a). The parameterization of W-QEISS was adjusted using available guidelines given by Huang et al. (2006a), Karakaya et al. (2015), and a trial-and-error process. For Borg MOEA, we set the number of function evaluation (NFE) equal to 2 millions, while the number of hidden neurons in the ELM, presenting a sigmoidal activation function, was set to 30. A k-fold cross-validation process (with k = 10) was repeated 5 times and the average resulting value was used to estimate the predictive accuracy of each model. The W-QEISS experiment with such setting was run 20 times to filter out the random component of the process, and the results presented below are obtained by merging the Pareto fronts obtained by each repetitions into a final Pareto front of non-dominated solutions.

2.4.1 Identification of basin's characteristics

In the first report concerning the $\mathbf{I_e}$ development (CHJ, 2007a), the index was validated for the time span from January 1986 to June 2000 against the supply deficit recorded in the basin with respect to agricultural and urban water demand, and the procedure for the State Index computation was approved. To

ensure comparability between the $\mathbf{I_e}$ and the FRIDA constructed index, we decided to employ the same supply deficit as target variable for the application of FRIDA approach to the Jucar case study. The Jucar supply deficit employed in this work was simulated via AQUATOOL model (Andreu et al., 1996). The model can run in simulation mode with a monthly time step, and it is conceived in the form of a flow network with oriented connections reproducing water losses, hydraulic relations between nodes, reservoirs and aquifers, and flow limitations based on elevation. Within AQUATOOL, complex processes such as evaporation and infiltration are effectively reproduced. The modeled supply deficit, employed as target variable, represents the monthly nominal shortage of water conveyed to the irrigation districts, and is only quantifiable a posteriori, when the water shortage has already jeopardized the fields. On the other hand, a drought index can be constantly monitored, and thus represents a valuable management tool for restraining drought impacts and identifying effective drought management strategies.

The database of candidate input variables was assembled retrieving the available observed variables in the basin and computing traditional drought indicators at multiple time aggregations. The resulting candidate predictors, listed in Table 2.1, are the following:

- 2 temporal features: date of the measurement, and month of the year;
- 12 monthly observed variables, current inputs to the $\mathbf{I_e}$, reported in Figure 2.4: average monthly storage and groundwater levels, average three months river runoff, and cumulated areal precipitation over 12 months;
- 8 additional observed variables in the basin: outflows from, and inflows to, the main reservoirs, and mean monthly areal temperatures;
- 6 traditional drought indicators: Standardized Precipitation Index (SPI), and Standardized Precipitation and Evapotranspiration Index (SPEI). SPI and SPEI indicators are computed on mean monthly data over the entire basin for 3, 6, and 12 months time aggregations. SPI requires as input the precipitation, and SPEI requires precipitation and temperature, as it uses the difference between precipitation and potential ET as reference variable.

  Their values express the water availability conditions of a basin in terms of units of standard deviation from the mean: negative (positive) values indicate drier (wetter) conditions than average (see McKee et al. (1993); Vicente-Serrano et al. (2010) for details on definition and calculation of these indicators).

Table 2.1: *Set of candidate input features for the feature extraction step via W-QEISS.*

Feature type	Feature code	Description
Time information	Date	Date of the measurement
	Moy	Month of the year
State Index Inputs	S1	Cumulated storage of Alarcón, Contreras and Tous
	S2	Storage at Forata
	F1	Flow measurement in the upper basin
	F2	Flow measurement in the upper basin
	F3	Flow measurement in the middle basin
	F4	Flow at Jardín tributary
	Pl1	Pluviometer measurement in Contreras reservoir
	Pl2	Pluviometer measurement in Tous reservoir
	Pl3	Pluviometer measurement in Bellús reservoir
	Pz1	Piezometric level in the south-east
	Pz2	Piezometric level in the center
	Pz3	Piezometric level in the west
Observed variables	In A	Inflow to Alarcón reservoir
	In C	Inflow to Contreras reservoir
	In T	Inflow to Tous reservoir
	Out A	Outflow from Alarcón reservoir
	Out C	Outflow from Contreras reservoir
	T1	Temperature in the west
	T2	Temperature in the center
	T3	Temperature in the east
Indicators	SPI_3	SPI at 3 months time aggregation
	$SPEI_3$	SPEI at 3 months time aggregation
	SPI_6	SPI at 6 months time aggregation
	$SPEI_6$	SPEI at 6 months time aggregation
	SPI_{12}	SPI at 12 months time aggregation
	$SPEI_{12}$	SPEI at 12 months time aggregation

2.4.2 Feature extraction via W-QEISS

The result of the W-QEISS algorithm is not a single most-accurate set of variables for a given cardinality, but several quasi-equally informative subsets, whose accuracy is lower than the best one by a small percentage $\delta \cdot 100\%$. Figure 2.5 represents a Selection Matrix, which reports the composition of each alternative subset of predictors within 15% of accuracy with respect to the highest one. The value $\delta = 0.15$ was chosen since it provides a reasonable trade-off between the number of solutions and their accuracy. The accuracy is measured in symmetric uncertainty between the target variable and the ELM calibrated using the reported subset.

The alternative subsets are sorted in ascending order of cardinality (from top to bottom), and accuracy (within each cardinality level). A rectangular marker is placed at the intersection between the row identifying a given subset and the columns corresponding to the selected predictors. The marker color varies with the cardinality of the subset, with lighter shades of gray indicating smaller subsets. In this case the cardinality spans from 3 to 9 features. The highest accuracy is reported in red and recorded for subset number 14. The 5 corresponding selected predictors, marked on the horizontal axis with a blue background, are the following:

- Moy: month of the year;
- S1: total storage aggregated for the reservoirs Alarcón, Contreras, and Tous;
- F3: river flow measured in the Jucar middle basin, after the confluence with smaller rivers Jardín and Lezuza coming from south-west;
- Pz2: groundwater level measured at the Piezometer situated in central area of the basin, in correspondence of a rainfed agricultural area;
- SPEI_6: SPEI at 6 month time aggregation computed with precipitation and temperature data averaged for the whole basin.

From the analysis of the Selection Matrix, several insights can be gained from a modeling and from a decision-making viewpoints. To begin with, insights on predictors' relevance can be obtained from the detection of the vertical bars traced by joining markers across multiple rows. Uninterrupted bars indicate strongly relevant predictors that cannot be substituted by other input combinations without incurring into a substantial drop of predictive accuracy. This is the case of the cumulated storage of the three main reservoirs Alarcón, Contreras, and Tous (S1). This information is essential to the final model, as

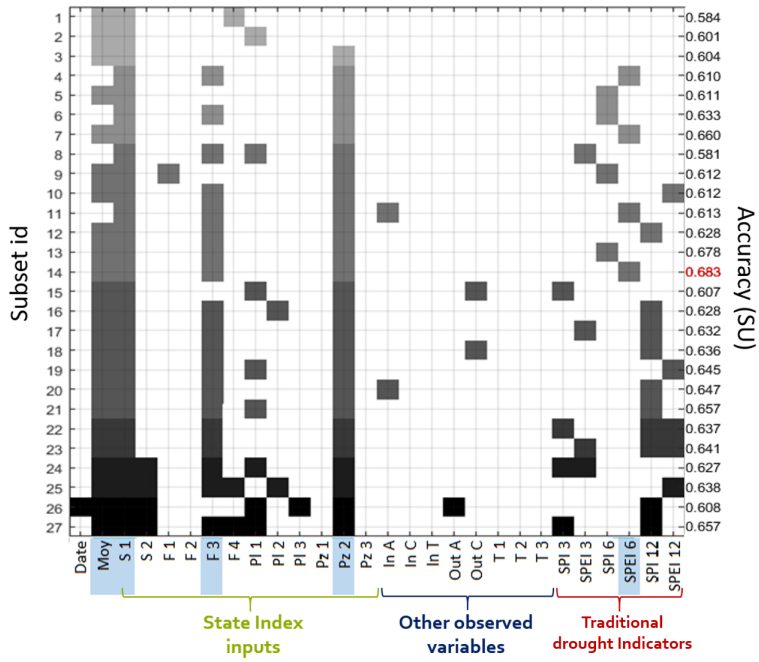

Figure 2.5: *Selection Matrix: the left vertical axis represents the subset number and the right vertical axis the corresponding accuracy measured in SU. A colored marker is put in correspondence of the variables, listed on the horizontal axis, selected by each subset. The shade of gray is an indication of the cardinality of the subset, lighter shades for lower cardinality. The highest accuracy is reported in red and the corresponding variables, constituting the most accurate subset, have a blue background.*

the exclusion of such predictors highly affects the model performance. Increasing gaps in the vertical bars are found when considering predictors with progressively weaker relevance, while irrelevant inputs are recognizable by isolated markers or their total absence. The variables Moy, F3, and Pz2 are considered relevant variables, as they are selected quite frequently, although high accuracy solutions exist that do not make use of all of them. Finally, the variable $SPEI_6$, while included in the most accurate subset, is overall present in 4 subsets only, whereas in other solutions with comparable accuracy it is replaced by different predictors, mainly carrying a similar precipitation-based information, such as pluviometer measures, or SPI, SPEI indicators at different time aggregations.

The presence of alternative subsets helps exploring the trade-off between multiple measures of predictive accuracy with respect to other metrics not directly considered in the optimization routine, an the choice of the preferred subset is determined by the index application. Given the cardinality, one can decide to sacrifice a small amount of predictive accuracy for an easier-to-yield (or more reliable) combination of predictors. For example, with a loss smaller than 1% in accuracy, subset 13 selects SPI_6 instead of $SPEI_6$. This possible replacement is interesting from an point of view as SPI is easier to compute than SPEI. In fact, SPI requires only the precipitation for its computation with respect to precipitation and temperature or evapotranspiration needed for the computation of SPEI. In addition, even after the preferred subset is chosen and the system is operating, knowing that one specific predictor can be replaced by one (or multiple) predictor(s) can aid the management in case of monitoring networks maintenance or instrument failure. When the main predictor is not observable, one can temporarily resort to alternative predictors incurring in a minimum loss of accuracy.

An additional consideration is related to the possibility to effectively address the uncertainty deriving from the choice of model inputs (Taormina et al., 2016). When multiple alternative subsets are provided, it is possible to explore the uncertainty related to the selection of predictors yielding similar accuracy. For instance, in this case study, we can observe that almost all subsets carry a groundwater and a rain information, but while the piezometric level is consistently provided by Pz2, the source of the precipitation information highly varies among the precipitation-based features (pluviometers or other SPI, SPEI indicators).

Finally, through the selection matrix analysis we can contrast the features selected by W-QEISS and the variables that constitute the State Index input set. Apart from sporadic single selections, all the observed variables not included in the State Index are consistently discarded by the W-QEISS as well, suggesting that the algorithm comes to the same conclusion as the Spanish experts considering inflows, outflows, and temperatures as non-relevant for the

description of the state of water resources in the Jucar river basin. Note that this result is a consequence of the use of the nominal agricultural demand to compute the target deficit. A temperature information is likely to become relevant if a real, weather-influenced, agricultural demand is employed instead. The feature month of the year is not explicitly an input to the State Index, nevertheless, an analogous information is implicitly included in the $\mathbf{I_e}$ through the normalization of the indicators described in equation 2.7. On the other hand, several features are considered in the $\mathbf{I_e}$, but generally neglected by W-QEISS selection. Among them, two out of three piezometers, the river flows upstream from the reservoirs, one pluviometer and the storage of Forata. These inputs probably result redundant due to their spatial correlation. Spatial variability is considered in the computation of $\mathbf{I_e}$ by including several spatially distributed observations of the main information categories: 2 measures of reservoir storages, 4 of river flows, 3 groundwater levels, and 3 precipitation measures. Conversely, the selection matrix supports the gain of a deeper understanding of the spatial interdependence of variables by identifying the best location for measuring the variables, spearing the need for several distributed measures. The highest accuracy-subset, in fact, selects only one variable out of each category: 1 storage, 1 river flows measure, 1 piezometer, and a spatially distributed precipitation information, i.e., SPEI_6 which replaces three areal pluviometers.

2.4.3 Drought Index Modeling

Among the pool of solutions, the choice of the preferred subsets is driven by the index application. For instance, an on-line use of the index that requires its frequent computation may benefit from an agile, easy-to-observe subset. With respect to the highest accuracy solution (subset 14), for instance, subset number 7 neglects predictor F3 thus presenting lower cardinality with an accuracy loss of only 3%. Similarly, the already mentioned subset 13 contains an easier-to-compute indicator (SPI instead of SPEI) with a negligible performance degradation. Nevertheless, for our methodological purpose we will employ the most accurate subset 14, as we are interested in discussing the potential of the method.

Concerning the model class choice, a highly flexible non-linear model is likely to yield the highest accuracy in reproducing the target. However, strong non-linearity and black-box behavior typically result in poor interpretability, a feature that is detrimental to the use of the index for management purposes as in the Jucar system, where restrictive measures in water use are activated when certain threshold values of the State Index are reached. As a consequence, the index outcome exerts a direct influence on many water-related activities requiring an easily interpretable and widely acceptable tool.

The calibration of a linear model on the chosen 5 dimensional subset seems

to be a good compromise between accuracy and transparency. As mentioned above, the feature Moy represents the succession of the months in the year, and is an expression of the seasonality of hydro-meteorological processes. Moy is constructed as the repetition of an array of numbers from 1 to 12 for the length of the considered time horizon, and thus presents a discontinuous shape: a slow and steady increase followed by a steep decrease in correspondence to the onset of a new year. While the non-linear models employed in the feature selection can effortlessly handle such an intermittent vector, linear models struggle with similar shapes. We therefore decided to account for the seasonality in the linear model indirectly, i.e., excluding Moy from the predictors set, but, consistently, considering seasonality by depurating the predictors of their annual cyclostationary mean.

The calibrated linear model representing the supply deficit is reported in Figure 2.6 and provides a very satisfying result, with an accuracy measured with the coefficient of determination in crossvalidation of $R^2_{FRIDA-linear} = 0.904$, significantly higher than the $R^2_{\mathbf{Ie}} = 0.739$ scored by the State Index, and a set of weights of immediate physical interpretability reported in Table 2.2. By inspecting the weights, one can notice that those assigned to the predictors Flow and $SPEI_6$ are very low, although not null, and the index trajectory is mainly determined by Storage and Piezometer values. S1 and Pz2, in fact, describe the trajectories of the main water reservoirs of the region, lakes and groundwater, whose fluctuations are the result of natural variability as well as human regulation, mainly for irrigation purposes.

Table 2.2: *Weights of the linear model calibrated on the optimal subset of predictors. The predictor Moy (month of the year), providing a seasonal information, is not directly included in the weights optimization but it is accounted for by depurating the variables of their annual cyclo-stationary mean.*

Predictor	Weight
Moy	/
Storage (S1)	0.721
Flow (F3)	10^{-9}
Piezometer (P2)	0.278
$SPEI_6$	10^{-9}

As a further analysis, we reiterated the model calibration and crossvalidation steps with a more complex, highly flexible model class, the ELM architecture, which scored an accuracy of $R^2_{FRIDA-ELM} = 0.907$. On the one hand, the arguably insignificant 0.005% improvement in accuracy of ELM with respect to the linear class, probably does not justify the loss of immediacy and trans-

parency induced by the transition to a black-box model. On the other hand, this experiment proves the robustness of the linear model in constituting the model class of choice for this drought index. In table 2.3 we report a more detailed comparison between State index, FRIDA-linear and FRIDA-ELM indexes with several accuracy metrics. The analysis of other metrics seem to reinforce the conclusions drawn by considering R^2 only: both FRIDA indexes (linear and ELM) outperform the State Index quite significantly, while the difference among them is negligible, although the non-linear index is always the top performing.

Table 2.3: *Accuracy of the State Index, FRIDA linear, and FRIDA ELM in reproducing the supply deficit, quantified in terms of coefficient of determination R^2, the Pearson correlation coefficient, the Root Mean Square Error (RMSE), and the fourth grade Root Mean Square Error (R4MS4E).*

Metric	State Index	Frida Linear	Frida ELM
R^2	0.7396	0.9036	0.9074
Pearson	0.8601	0.9506	0.9533
RMSE	0.2066	0.1135	0.1014
R4MS4E	0.2549	0.1475	0.1299

The reported metrics do not distinguish between errors above and below the target deficit. Indeed, we consider these two error types of comparable importance. On the one hand, the underestimation of a deficit value may find the water users unprepared to face a serious drought. On the other hand, the overestimation of drought conditions may ignite repeated false alarms that will compromise the index trustworthiness and its efficacy in triggering an alert state. Therefore, rather than penalizing an error above or below the target trajectory, we find more compelling to focus on errors in the most crucial drought situations i.e., at the maximum level of deficit recorded. One way of doing so is considering R4MS4E, as in Table 2.3, which penalizes errors in the deficit peaks. Another specific assessment tool for analyzing the indexes performance during critical droughts is the confusion matrix, reporting the classification performance of critical droughts, here arbitrarily defined as months reporting deficit values above the 85th percentile (Tables 2.4, 2.5, 2.6). The rows of the confusion matrix represent the instances in a predicted class while the columns represent the instances in an actual class. Consequently, the main diagonal reports the number of correctly classified points. Cells outside the main diagonal specify the errors: the value in the bottom-left cell (first column, second row) indicates a situation in which the index does not recognize an ongoing drought, while the value in the top-right cell (first row and second column) indicates the number

of false alarms. FRIDA-ELM confusion matrix seems to significantly exceed the competitors' performances by erroring only 0,57% of the times, as opposed to the 10,91% of I_e, and the 6,3% of FRIDA-linear.

Table 2.4: *State Index confusion matrix.*

SI-deficit	critical drought	normality
critical drought	131	18
normality	1	24

Table 2.5: *FRIDA-Linear confusion matrix.*

Frida Linear-deficit	critical drought	normality
critical drought	138	11
normality	0	25

Table 2.6: *FRIDA-ELM confusion matrix.*

Frida Linear-deficit	critical drought	normality
critical drought	147	2
normality	1	24

2.5 Conclusions

The purpose of this study is to contribute to the identification of drought management strategies able to improve the efficiency and resilience of drought prone regulated water systems. This problem is considered urgent as the analysis of climate trends shows that drought frequency and severity are intensifying all over in Europe, particularly in the Mediterranean area.

This work explores the potential of drought indexes as a management tool for the purpose of containing drought impacts. Since traditional indicators are often inadequate to characterize water availability conditions in highly regulated contexts, a novel framework for the construction of customized basin-specific drought indexes is proposed. This framework relies on the employment of a feature extraction technique, the Wrapper for Quasi Equally Informative Subset Selection (W-QEISS). Given a set of information collected in the basin, W-QEISS features a deep learning machine that automatically selects the most suitable input set for the construction of a model reproducing the target

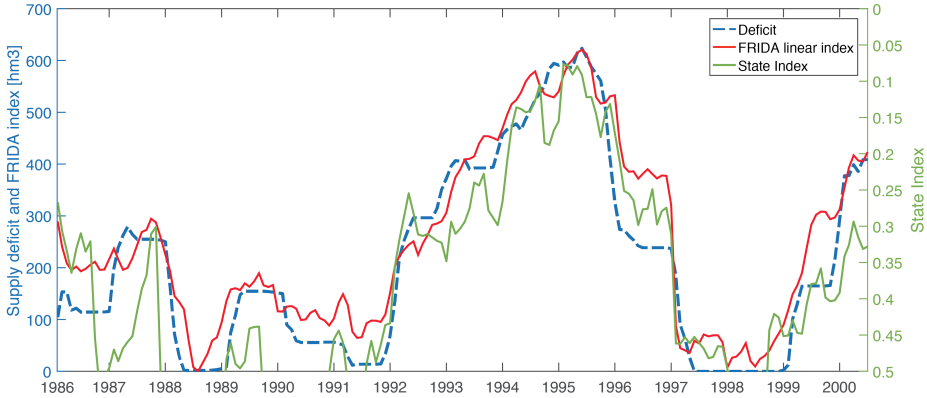

Figure 2.6: *Comparison between the FRIDA linear index (blue) and the state index (green) in reproducing the monthly aggregated supply deficit (red). FRIDA index presents an higher similarity with the deficit and only requires 5 inputs instead of the 12 required by the state index.*

variable, i.e., a ground truth representative for the state of water resources in the basin. Specifically, W-QEISS performs the search process in a four-dimensional metric space of predictive accuracy, cardinality, relevance, and redundancy. On top of that, W-QEISS algorithm is designed to identify one subset with the highest predictive accuracy and multiple subsets with similar information content (i.e., quasi equally informative subsets). This provides insights on the relative relevance of the variables and a deeper understanding of the underlying physical processes taking place in the basin. The choice of the preferred input set and model class balance accuracy and practicality of the index. The efficacy of FRIDA methodology is strongly dependent on data availability, in terms of predictors diversity and numerosity, and length of the time series. FRIDA is best applicable in contexts where an extensive monitoring system has been in place for long enough to allow a consistent and informative dataset for the index calibration. However, while some hydro-meteorological variables are easy to monitor and most often available (e.g., precipitation, temperature), the accessibility of soil moisture, groundwater table level, snowpack extent, air humidity etc., may represent a problem. When a key drought-driving variable for the context at hand is absent from the input set, the efficacy of FRIDA is undermined.

The application of the FRIDA in the Jucar river basin case study has successfully demonstrated the suitability of the framework to design a basin specific drought index. Firstly, the automatic variable selection yields an immediate and informative result, which presents strong similarities with the empirical expert-based variable set employed by the CHJ, while involving a significantly lower number of features (5 variables instead of the 12 required by the State In-

dex). Secondly, the newly computed FRIDA linear index outperforms the official Spanish State Index in terms of accuracy in reproducing the target variable, while maintaining immediate interpretability.

However, one of the reasons why the $\mathbf{I_e}$ enjoyed such wide acceptance among the Jucar stakeholders is related to the widely comprehensive approach employed for its construction. All water users, in fact, feel represented in the index through at least one variable being observed in the proximity of their water related activity, even if such variable is low-weighted or redundant when computing the basin-wide aggregated indicator. The FRIDA approach does not ensure such representation of all water users, although it appears as a more rigorous and efficient alternative to the inclusive CHJ approach. Moreover, FRIDA is a portable methodology, suitable for the many drought prone contexts in need of a drought management plan. In conclusion, the aim of arranging an effective framework for the construction of basin customized combined drought indexes can be considered achieved. The indexes constructed with FRIDA have proven to be an asset for (i) representing drought conditions in highly regulated basins, where traditional indexes tend to fail; (ii) gaining a deeper understanding of the hydro-meteorological processes taking place in the basin; and (iii) constituting a valid alternative to the Spanish approach for the State Index design, thus supporting appropriate drought management strategies, such as triggering drought restraining response measures.

The already valid results achieved by this study open new possibilities for the use of basin-specific drought indexes. Further research efforts could be addressed to exploring the potential of employing FRIDA indexes in directly informing water management operations. Additionally, the possibility of forecasting such indexes can be tested in order to timely prepare for upcoming dry seasons. We expect that the projection of a drought index fosters the adoption of a proactive (as opposed to the current reactive) approach in facing a drought. Proactivity promotes a shift from costly and often belated mitigation measures, to preventive actions that will grant flexibility to timely prepare to upcoming droughts, while reducing costs associated to drought impacts and restrictions.

Ultimately, FRIDA can represent an asset for improving the system resilience under a changing climate. Despite the fact that FRIDA is conditioned upon historical data, one can imagine that in the short term, drivers' interactions and relative role in causing a drought hold unchanged. In this case, the index formulation remains valid in the context of a changing climate. In the long term, nevertheless, this hypothesis may cease to hold, we thus suggest a frequent reiteration of FRIDA procedure to monitor the evolution of drivers and dynamics leading to a drought in the basin. For example, in a groundwater dominated system as the Jucar basin, the piezometer information is likely to remain essential in a future climate, but, at the same time, we can expect evapotranspiration

processes to increase their drought-propelling role, as climate change induces a general increase of temperatures. In other contexts, e.g., snow dominated catchments, the role of snow may lose priority due to a diminishing winter snowpack reserve. FRIDA will thus represent a valuable tool to support the analysis on the dynamic role of drivers in drought evolution under a changing climate.

Code availability: The complete dataset employed for the feature selection step can be downloaded open source from `http://doi.org/10.5281/zenodo.1185084` (Zaniolo et al., 2018a). A detailed description of FRIDA, including both data and codes, is available at `https://www.ei.deib.polimi.it/?page_id=779`.

3

Data-driven Modeling and Control of Droughts

Abstract[1]

In highly regulated water systems droughts are complex, basin-specific phenomena. The identification of drought drivers is challenged by the coexistence of possibly relevant processes with inconsistent dynamics and origins (natural or anthropic). FRIDA is a fully automated data-driven approach developed to extract relevant drought drivers from a pool of candidate hydrometeorological predictors at different time aggregations. Selected predictors are then combined into a basin-specific drought index to monitor the state of water resources in highly regulated contexts. The operational value of this index in improving water systems operations is quantified by designing a control policy informed by the index, and contrasting its performance with that of a baseline policy conditioned on basic information only. The approach is demonstrated on Lake Como, Italy, a multipurpose regulated lake operated for flood control and irrigation supply. Results show that the designed index is accurate in representing basin drought conditions, and the overall system performance can improve by nearly 20% when operations are informed with the basin-tailored drought index. The proposed framework is portable across different contexts, where basin-specific drought indexes can support drought characterization and control in a fully data-driven fashion.

[1]This work has been published as: Zaniolo, M., Giuliani, M., Castelletti, A., 2019. Data-driven modeling and control of droughts. IFAC- Papers On Line 52, 54-60

3.1 Introduction

Droughts in highly regulated water basins are prolonged periods of anomalous water supply failure (Pedro-Monzonìs et al., 2015). In these systems, the presence of man controlled water infrastructures (e.g., dams, pumping wells) can play a role in either magnifying drought impacts with inconsiderate water use, or in restraining them, when the available storage is effectively operated to hedge water shortages and distribute resources in space and time (Mishra and Singh, 2010). In order to design effective drought management strategies, it is key to timely recognize onset and severity of a dry spell. Nevertheless, drought identification is a challenging task, due to the interplay of several hydrometeorological phenomena characterized by disparate spatiotemporal dynamics, levels of observability, and relevance to drought conditions (AghaKouchak, 2015a). Traditional drought indexes (reviewed in e.g., Zargar et al., 2011) can contribute to this task for natural systems, however, in highly regulated basins the concurrence of natural fluctuations (e.g., below-average precipitation, above-average evaporation), and anthropic factors (excessive water demand, suboptimal management) results in basin specific dynamics where no single tool is suitable for all contexts. FRIDA is a FRamework for Index-based Drought Analysis for the automatic and fully data-driven design of basin-tailored drought indexes specifically targeted for highly regulated basins (Zaniolo et al., 2018b). The framework is based on a wrapper for feature extraction which selects relevant drought drivers from a pool of available candidate hydro-meteorological predictors and traditional drought indicators as to predict in the most effective way a target variable representative for the basin drought condition (e.g., water deficit). The feature extraction technique employs a multi-objective evolutionary algorithm to find Pareto-efficient subsets of variables by maximizing the wrapper accuracy, minimizing the number of selected variables (cardinality), and optimizing relevance and redundancy of the subset.

In this work, we contribute a holistic framework for drought modeling and control that combines FRIDA, for the construction of a customized drought index, with a data-driven control scheme (Formentin et al., 2013; Costanzo et al., 2016; Smarra et al., 2018) that employs the designed index to inform water reservoir control during droughts. The operational value of the proposed index is assess by re-designing the control policy by including the index its inputs, and contrasting the corresponding improvement with the performance of a policy relying on basic information only. We demonstrate the proposed methodology on Lake Como basin, Italy, where droughts have potentially severe implications on downstream crop yield.

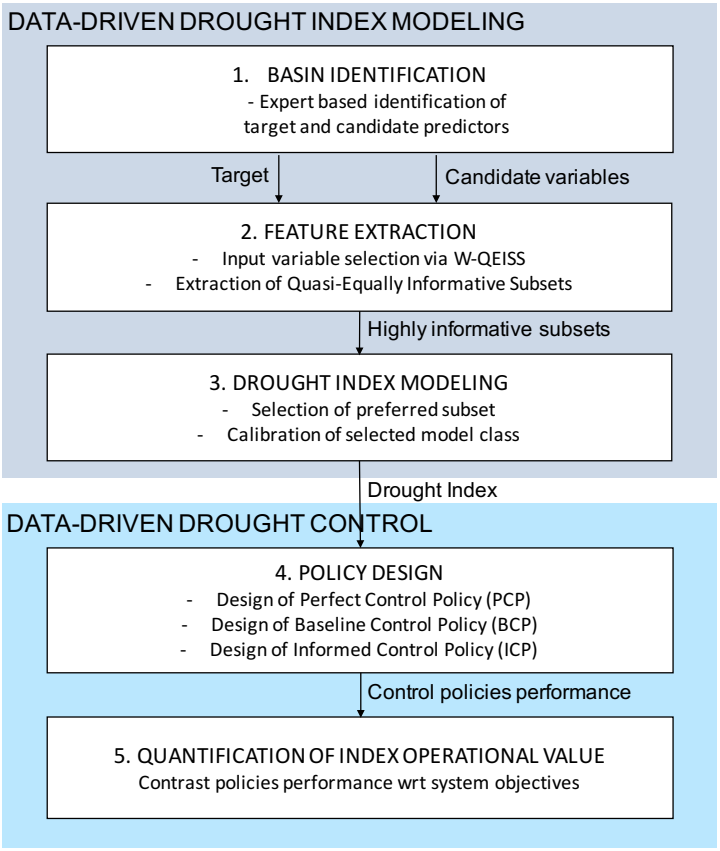

Figure 3.1: *Flowchart of the proposed framework for drought index design and evaluation.*

3.2 Methods and Tools

The proposed data-driven method is composed of two building blocks (Fig. 3.1): the first block, namely, data-driven drought index modeling, employs the Framework for Index-based Drought Analysis (FRIDA) for the automatic design of basin customized drought indexes in highly regulated basin (Zaniolo et al., 2018b). FRIDA drought index is the input to the second building block, which explores the operational advantage of using the index within a data-driven control scheme to condition water management operations.

3.2.1 Data-driven drought index modeling

In the data-driven drought index modeling block we employ FRIDA, composed of 3 steps. The first step is an empirical basin identification aimed at the def-

inition of a target variable and a set of candidate predictors for the basin of application. The target variable is a form of water deficit representative for the drought conditions in a basin (e.g., supply deficit with respect to the demand, soil moisture deficit). The set of candidate predictors is a collection of hydro-meteorological variables and traditional drought indicators available for the study site, at different time aggregations. The final drought index will be an appropriate combination of a subset of predictors to best represent the target variable, i.e., the basin drought conditions.
The second step is the Feature Extraction step, aimed at reducing the dimension of the predictors' space into a subset of selected variables with a minimum loss of information content. This step will define a number of highly informative subsets of predictors representing efficient inputs set for the Drought Index. It is performed via Wrapper for Quasi-Equally Informative Subset Selection (W-QEISS) (Karakaya et al., 2015). Unlike the more common filters for feature extraction which evaluate the relevance of each candidate variable separately, wrappers assess the relevance of a variables ensemble, computing the prediction performance of a given learning machine calibrated on the input set (Guyon, 2003). Moreover, W-QEISS does not only return the most accurate subset, but a set of quasi-equally informative subsets (i.e., with comparable accuracy performance) in order to support a data-driven analysis on predictors interactions and relative relevance (Karakaya et al., 2015).

The majority of wrappers solves a two-objectives optimization problem trading off regression accuracy and model complexity (Guyon, 2003). In the following formulation, $\mathbf{X}$ is the pool of candidate predictors, y the output variable, and $\mathbf{S}$ a subset of $\mathbf{X}$. SU (Symmetric Uncertainty) is a measure of how much information is shared between two features relatively to the entire information content of both features, and ranges between 0 (independent variables) to 1 (complete dependence) (Witten and Frank, 2005).

- **Predictive accuracy:** $f_1(S) = SU(y, \widehat{y}(S))$; to be maximized, defined as the Symmetric Uncertainty between the observed output y and the prediction $\widehat{y}(S)$ obtained from the model.
- **Model complexity:** $f_2(S) = |S|$; to be minimized, counts the number of predictors chosen, i.e., the subset cardinality. Its minimization guarantees that the resulting model will not be unnecessarily complex.

Beside these two traditional objectives, W-QEISS adds the optimization of two other metrics of relevance and redundancy concerning the degree of dependence between variables. The adoption of these metrics ensures high information content and low intra-subset similarity of the produced subsets (Sharma and Mehrotra, 2014).

- **Relevance:** $f_3(S) = \sum_{x_i \in S} SU(x_i, y)$; to be maximized, computes the

symmetric uncertainty between the feature x_i and the output y. It measures of the explanatory power of the feature with respect to the output.

- **Redundancy:** $f_4(S) = \sum_{x_i, x_j \in S, i<j} SU(x_i, x_j)$; to be minimized, computes the symmetric uncertainty among two predictors, x_i and x_j. Its minimization guides the search towards the selection of mutually dissimilar features.

The W-QEISS feature extraction is composed of three main steps. (1) A set A of Pareto-efficient solutions is built according to the above defined four objective functions. A global multi-objective optimization algorithm is employed to explore the space of the possible subsets, while a regression model is calibrated and used to compute the predictive accuracy f_1 of each set. (2) Among the Pareto-efficient subsets, the maximum value of accuracy f_1^* is identified. Then, on the basis of a pre-selected arbitrarily small value δ, all the subsets whose accuracy is lower than $(1-\delta)f_1^*$ are eliminated from the ensemble A of the non-dominated solutions, obtaining A_δ. (3) The final subset A_δ^* is computed after the elimination of the inferior subsets. The subset S_j is considered inferior to S_i if it is a superset of S_i and does not score higher accuracy. Formally: $S_j \supset S_i$ and $f_4(S_j) \leqslant f_4(S_i)$.

The implementation of W-QEISS requires the selection of the global multi-objective optimization algorithm, and of the regressor to be employed. In this study, we used Borg MOEA (Reed et al., 2013) as optimization algorithm, and the Extreme Learning Machines (ELM) (Huang et al., 2006a) as regressor. Borg MOEA is a sophisticated evolutionary algorithm which features a set of strategies to contrast the main shortcomings of evolutionary algorithms. Among them, the automatic adaptation of several evolutionary operators to promote the generation of efficient solutions, the ϵ-box dominance archive to contrast overfitting, and time continuation to avoid being trapped into search local minima (Hadka and Reed, 2013). ELM is designed as a single-hidden layer feedforward neural network. Though unlike traditional ANN, ELM does not perform a calibration of the nodes parameters, but randomly chooses hidden nodes and analytically determines the output weights of the hidden layer, thus bypassing the time-consuming calibration process of ANN (Huang et al., 2006a).

In the third step, one subset is selected among the quasi-equally informative, balancing the trade-off between competing objectives and, possibly, considering additional operative preferences neglected in the search (e.g., cost and reliability of the variable monitoring). Lastly, an appropriate regressor is fit to the chosen input set and the target variable, producing the Drought Index. The regressor choice is determined by the application of interest. In general, highly non-linear learning machines provide high accuracy and flexibility at the price of poor intuitive interpretability. On the other hand, linear models partly sacrifice approximation skills but grant an immediate understanding of the index

physical meaning.

3.2.2 Data-driven drought control

In the second building block of the framework, we investigate the potential of the Drought Index in conditioning water system operations by directly informing the regulation policy about the basin drought conditions.

The control problem of a water reservoir is traditionally formulated as a Markov Decision Process (MDP) where the state of the system $\mathbf{x}_t$ is the reservoir storage, while the decisions $\mathbf{u}_t$ concerns the volume of water to be released from the reservoir at each time step of the control horizon $[0, h]$. The state in $t+1$ is determined by state and decision at time t, and by stochastic disturbances affecting the system ε_{t+1}, (e.g., reservoir inflows) and generally described with a probability density function $\varepsilon_{t+1} \sim \phi_t$ in $t \in [0, h)$. Formally:

$$\begin{aligned} \mathbf{x}_{t+1} &= f_t(\mathbf{x}_t, \mathbf{u}_t, \varepsilon_{t+1}) \quad t \in [0, h) \quad \text{where:} \\ \mathbf{x}_t \in \mathbb{R}^{n_x}; &\quad \mathbf{u}_t \in \mathcal{U}_t(\mathbf{x}_t) \subseteq \mathbb{R}^{n_u}; \quad \varepsilon_t \in \mathbb{R}^{n_\varepsilon} \end{aligned} \tag{3.1}$$

In the adopted notation, the time subscript of a variable indicates the instant when its value is deterministically known. The MDP described in Equation (3.1) is discrete-time, periodic, nonlinear, and stochastic. A water reservoir control problem is generally multi-objective, and each objective (here assumed to be a cost) is formulated as the sequence of immediate costs $g^m_{t+1}(\cdot)$ associated with each state transition from time t to $t+1$ in $[0, h]$, and a penalty function $g^m_h(\mathbf{x}_h)$ associated with the final state. The step cost trajectory is aggregated over time with the operator Φ (e.g., the average $\Phi = \sum /(h+1)$) and the disturbances noise is filtered with the operator Ψ (e.g., expected value $\Psi = E$) as follows:

$$J^m = \Psi_{\varepsilon_1,\ldots,\varepsilon_h}\left[\Phi_{0,\ldots,h}\left(g^m_1(\mathbf{x}_0, \mathbf{u}_0, \varepsilon_1), \ldots, g^m_h(\mathbf{x}_h)\right)\right] \tag{3.2}$$

Solving the optimal control problem for the system defined in (3.1) thus requires to find the optimal control policy p^* (i.e., the periodic sequence of control laws $\mathbf{u}_t = \mu^*_t(\mathbf{x}_t)$ in period T) which minimizes the M-dimensional objective vector $\mathbf{J}$, i.e.:

$$\begin{aligned} p^* &= \arg\min_p \mathbf{J}(p, \mathbf{x}_0, \varepsilon^h_1) \\ &= \left[J^1(p, \mathbf{x}_0, \varepsilon^h_1), \ldots, J^M(p, \mathbf{x}_0, \varepsilon^h_1)\right] \end{aligned} \tag{3.3}$$

In step 4 of the framework, Problem 3.3 is solved three times for a Perfect, Baseline, and Informed policy, differing for the information set given to the policy.

1. Perfect Control Policy (PCP): designed assuming to have full and perfect information on the future at the moment at each decision step: $u_t = p(t, \mathbf{x}_t, \varepsilon_{[0,H]})$.

2. Baseline Control Policy (BCP): designed relying on a basic information set, e.g., time and state of the system: $u_t = p(t, \mathbf{x}_t)$.

3. Informed Control Policy (ICP): designed including the Drought Index as additional input to the control policy, i.e. $u_t = p(t, \mathbf{x}_t, DI_t)$.

The design of the Perfect Control Policy is a nonlinear deterministic optimization problem which is traditionally solved via Deterministic Dynamic Programming (DDP) (Bellman, 1957). Baseline and Informed policies instead require a stochastic optimization which does not assume any deterministic knowledge of future disturbances. The stochastic extension of DDP, namely Stochastic Dynamic Programming (SDP) is in this case severely limited by the so-called *curse of modeling* in designing control policies conditioned on exogenous information (e.g., the Drought Index) (Tsitsiklis and Van Roy, 1996). We therefore solve Problem (3.3) by means of Evolutionary Multi-Objective Direct Policy Search (Giuliani et al., 2016b), an approximate dynamic programming approach which allows to condition the control policy with non-modeled (exogenous) information in a fully data-driven fashion (Denaro et al., 2017a).

EMODPS performs a simulation-based optimization that directly operates in the policy space by exploring the parameter space Θ seeking the best parameterization for the control policy p_θ, defined within a given family of functions. The parameters are chosen in order to optimize the expected long-term cost defined by the objectives of the problem. Problem (3.3) thus becomes:

$$p_\theta^* = \arg\min_{p_\theta} \mathbf{J}_{p_\theta} \quad \text{s.t. } \theta \in \Theta \tag{3.4}$$

where the objective function $\mathbf{J}_{p_\theta}$ is defined as in Equation 3.2. Finding p_θ^* therefore corresponds to finding the best parameters θ^* for the class of policies p_θ, measured by the objectives $\mathbf{J}_{p_\theta}$. As search algorithm, we employ again Borg MOEA, which have been demonstrated to outperform other state-of-the-art MOEAs on challenging reservoir control problems (Zatarain et al., 2016), while for the choice of the control policy class we use Gaussian Radial Basis Functions networks (RBFs) given their demonstrated capacity to represent policies for a large class of MDPs (Busoniu et al., 2011). Finally, in step 5 of the procedure, the operational value of exogenous information can be assessed as the difference in system performance between a control policy based upon the exogenous information (ICP) and a policy relying on traditional information only (BCP) using the optimal policy (PCP) as reference. As suggested in Giuliani et al. (2015), we use the hypervolume indicator (HV), which captures both the convergence

of the Pareto front under examination $\mathcal{F}$ to the optimal one $\mathcal{F}^*$, as well as the representation of the full extent of tradeoffs in the objective space. The hypervolume metric allows set-to-set evaluations, measuring the volume of objective space Y dominated ($\preceq$) by the considered approximate set. HV assumes values between 0 to 1, where Pareto fronts with higher HV are considered better. The indicator is formally defined as:

$$\begin{aligned} HV(\mathcal{F}, \mathcal{F}^*) &= \frac{\int \alpha_{\mathcal{F}}(\mathbf{y})\,d\mathbf{y}}{\int \alpha_{\mathcal{F}^*}(\mathbf{y})\,d\mathbf{y}} \quad \text{where} \\ \alpha_{\mathcal{F}}(\mathbf{y}) &= \begin{cases} 1 & \text{if } \exists \mathbf{y}' \in \mathcal{F} \text{ such that } \mathbf{y}' \preceq \mathbf{y} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{3.5}$$

3.3 Study Site

The proposed Framework is demonstrated on Lake Como, a multipurpose regulated lake in the southern Alpine belt (Italy). The main tributary, and only emissary of the lake, is the Adda river, whose sublacual part originates in the southeastern branch of Lake Como and feeds 4 agricultural districts, eventually serving as a tributary for the Po river downstream. The southwestern branch of the lake, where the city of Como is located, constitutes a dead end exposed to flooding events. The hydrological regime is snow-rainfall dominated, characterized by scarce winter and summer inflows, a large snowmelt peak in late spring and and a secondary rainfall peak in autumn. The lake regulation has the dual aim of guaranteeing flood protection to the lake shores and supplying water to downstream users. Despite the regulation efforts, the late summer water demand is not always met, leading to conflicts among water users and negative impacts due to water shortage, which might be particularly critical in drought years. The system is modeled as a discrete-time, periodic, non-linear, stochastic MDP defined by a scalar state variable x_t (i.e., Lake Como storage), a control variable u_t representing the release decision, a vector of stochastic disturbances ε_{t+1} (net reservoir inflow), and a non-linear state-transition function $x_{t+1} = x_t - r_{t+1} + \varepsilon_{t+1}$, where the effective release r_{t+1} coincides with the release decision unless a correction is needed to respect physical or legal constraints. The release decision is daily extracted from the control policy $p(\cdot)$. On the basis of previous works (Anghileri et al., 2011), two sets of objectives are defined:

Flooding: the average number of annual flood days, defined as days in which the lake level is above the flood threshold $\bar{h} = 1.24$ m, i.e.,

$$J^{\text{flood}} = \frac{1}{N_y} \sum_{t=0}^{H-1} g_{t+1}^{\text{flood}}; \quad g_{t+1}^{\text{flood}} = \begin{cases} 1 & \text{if } h_{t+1} \geqslant \bar{h} \\ 0 & \text{if } h_{t+1} < \bar{h} \end{cases} \tag{3.6}$$

Irrigation: the daily average squared water deficit w.r.t. the daily downstream demand w_t, subject to the minimum flow constraint $q^{MEF} = 5\,m^3/s$ to guarantee environmental stakes, i.e.,

$$J^{irr} = \frac{1}{H}\sum_{t=0}^{H-1}(\beta_t \cdot max(w_t - (r_{t+1} - q^{MEF}), 0))^2 \tag{3.7}$$

where β_t is a time-varying coefficient accounting for the different impacts of the water deficit on the crop production depending on the growth stages, and the quadratic formulation of (3.7) aims to penalize severe deficits in a single time step, while allowing for more frequent, small shortages.

Among the irrigation districts served by the Adda river, the Muzza-Bassa Lodigiana is the largest (700 km^2) and associated with the highest water demand. Within the district, the main crop is maize. A large distributed-parameter model is available for the district to perform the simulation of the irrigation water distribution and the computation of the hydrologic balance in the crop root zone, in a regular grid of cells with 250 m of side length. The effective soil moisture of the crop is computed along with the required moisture, allowing the detection of periods of water stress. A detailed description of the model can be found in Giuliani et al. (2016c).

3.4 Results

In this chapter, we present the results of the application of the proposed methodology to the Lake Como case study, following the structure proposed in Fig. 3.1.

3.4.1 Data-driven drought index modeling results

The first step is the identification of a target variable, representative for the basin drought conditions, and a dataset of candidate drought predictors. In Lake Como system, droughts primarily affect the yearly crop yield of downstream irrigation districts. The target variable must be observable during the irrigation season and a reliable proxy of the annual crop yield recorded at the end of the season. The irrigation district model described in Section 3.3 can be used to simulate the water requirement $\bar{wr}_t^c$ of crops in cell c and the relative effective soil water content sw_t^c. Their difference is the water deficit of each cell $d_t^c = max(\bar{wr}_i^c - sw_t^c, 0)$, later shown to be an effective proxy of crop yield (Fig. 3.3). The estimated crop water requirements are only meaningful during the 22 weeks of crop (maize) growth season, ranging from week of the year number $WY_{beg} = 16$ to $WY_{end} = 37$, from mid-April to mid-September. The target variable D employed for the feature extraction experiments is the weekly cell deficit cumulated over the district spatial domain d_t^{SD}, smoothed

with a moving window of semiamplitude $s = 2$ weeks to filter the effects of conveyance and irrigation turns within the district units.

$$D = \frac{1}{WY_{end} - WY_{beg} - 2s} \sum_{w=WY_{beg}+s}^{WY_{end}-s} \left[\frac{1}{2s} \sum_{i=w-s}^{w+s} d_i^{SD} \right] \tag{3.8}$$

Observational data and traditional drought indicators at different time aggregations constitute the dataset of candidate variables, likely consisting in a large collection of correlated variables. After a preliminary correlation analysis between the candidate features and the ground truth, time aggregations that scored very low correlations were discarded, resulting in a 30 features dataset, reported in Table 3.1. The considered indicators belong to the family of the Standardized Precipitation Index (SPI) and are reported in the table in brackets. For a review of drought indicators see Zargar et al. (2011).

Feature type	**Feature name**	**Aggregation [weeks]**
Time	Week of the year	-
Variables	Precipitation	4
	Temperature	1
	Lake level	1
	Snow melt	4
	Snow Water Equivalent (SWE)	4
Indicators	SPI (precipitation)	1, 2, 4
	SPEI (precipitation, evaporation)	1, 2, 4
	SMRI (Snowmelt and rain)	4, 12, 26
	SWI (lake water level)	1, 2, 4
	SRI_in (lake inflow)	1, 2, 4
	SRI_out (lake outflow)	1, 2, 4
	SSI_sup (superficial soil moisture)	12, 26, 52
	SSI_tot (total soil moisture)	12, 26, 52

Table 3.1: *Set of candidate input features for WQEISS, comprehensive of time information, observed variables, and indicators. Following indicators' name, in brackets, the variable(s) needed for their computation are reported.*

On this dataset, the W-QEISS algorithm is run with the aim of extracting the most accurate and parsimonious subsets, while optimizing their redundancy and relevance. In particular, the W-QEISS routine was repeated 20 times for 2 millions function evaluations to filter out the random component of the algorithm, and a comprehensive Pareto front of quasi-equally informative subsets

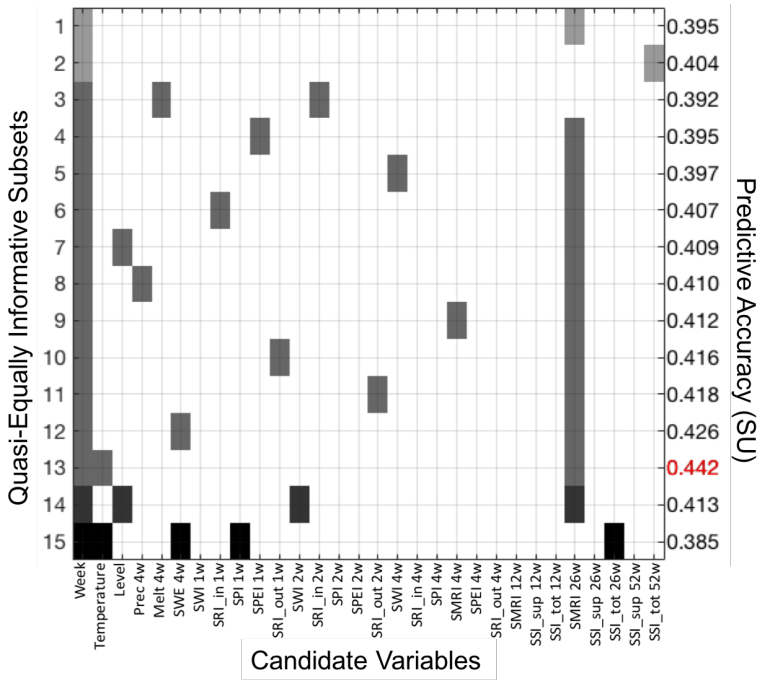

Figure 3.2: *Selection matrix: the left vertical axis represents the subset number and the right vertical axis the corresponding accuracy measured in SU. A colored marker is put in correspondence of the variable (horizontal axis) selected by each subset.*

was obtained by merging the fronts resulting from each seed. The outcome is represented in the Selection Matrix in Fig. 3.2 reporting the composition of each alternative subset of predictors within 15% of accuracy (measured in SU) with respect to the highest one.

The alternative subsets (vertical left axis) are sorted in ascending order of cardinality and accuracy. A rectangular marker is placed at the intersection between the row identifying a given subset and the columns corresponding to the selected predictors. The marker color varies with the cardinality of the subset, with lighter shades of gray indicating smaller subsets. We can observe that 11 out of 15 subsets have a cardinality of 3 variables, and the remaining subsets include 2, 4, and 5 predictors. The highest accuracy is recorded for the subset 13, reported in red, which selects as predictors Week of the year, Temperature (i.e., mean weekly temperature), and SMRI with an aggregation time of 26 weeks, equal to 6 months. Insights on predictors' relevance can be obtained from the observation of the vertical bars traced by joining the markers across multiple rows. Long and uninterrupted bars are expected for strongly relevant

predictors that cannot be substituted by other input combinations without incurring into a substantial drop of predictive accuracy. Increasing gaps are found when considering predictors with progressively weaker relevance.We observe that the information carried by the predictors Week of the year and SMRI 26w is strongly relevant as it is selected by 15 and 12 subset, respectively. The temperature information, despite belonging to the most accurate subset, is instead replaceable by most of the predictors with a negligible loss of accuracy. If we remove the snow information carried by SMRI 26w, SWE, or Melt, the wrapper accuracy degrades of at least 10% and selects another low frequency information, i.e., the soil water content over the entire year SSI_tot 52w. Despite the accuracy drop, such alternative may be interesting given that acquiring snow estimations is generally expensive and requires an extensive monitoring system.

The choice of the preferred subset for index computation is driven by the application, possibly considering additional objectives not explicitly included in the search (e.g., cost and reliability of variable retrieval, index immediacy). For our methodological purposes we select the most accurate subset (number 13) for the drought index modeling step. The resulting Drought Index (DI), obtained calibrating an ELM on the selected subset, is reported in Fig. 3.3. The figure background indicates the class of yearly yield recorded for the year, serving as an additional ground truth to validate the target variable. As expected, low levels of deficit coincide with abundant yield (green background), and vice-versa, indicating that the soil moisture deficit is an appropriate drought proxy for the case study from the point of view of the drought vulnerable stakeholders. DI successfully reproduces the target variable, scoring a correlation coefficient r=0.74, and symmetric uncertainty SU=0.44 over the growth season.

3.4.2 Data-driven drought control results

The optimal Lake Como operation is designed by solving Problem 3.3, where the vector of objectives $\mathbf{J}$ is defined in Equations (3.6) and (3.7). Such problem is solved three times for BCP, ICP, and PCP, employing different information sets. The best performing policy is, as expected, the PCP, represented with black diamonds. The associated Pareto front is sharp cornered, implying that with a perfect inflow foresight there is almost no conflict between the two operating objectives. In the absence of such perfect future knowledge, however, the use of additional information in the system operation has the potential to mitigate existing conflicts. ICP completely outperforms BCP solutions, demonstrating that the addition of the DI as an exogenous information for the policy design is a viable option for the improvement of the lake baseline operations. The improvement is primarily observed in the values of the irrigation object, as the ICP Pareto front appears to be a vertical shifted BCP front in the direction of

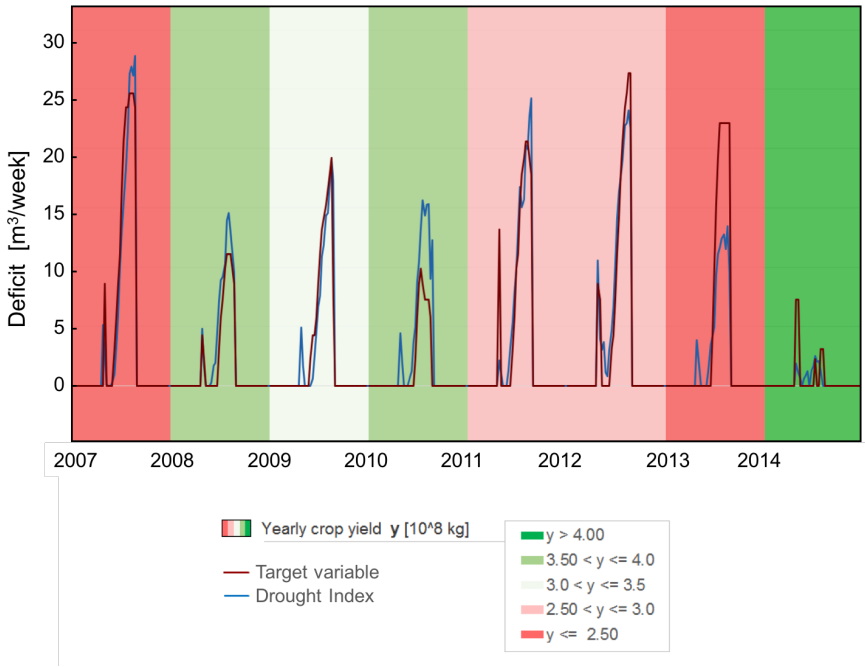

Figure 3.3: *The designed Drought Index compared against target variable and annual yield.*

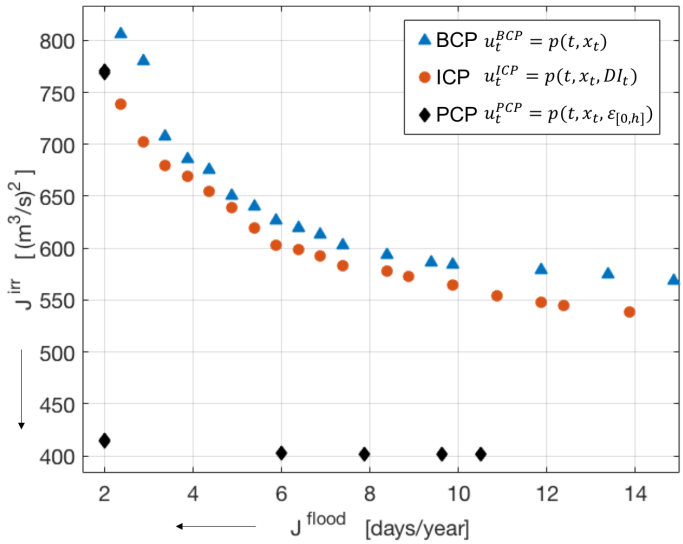

Figure 3.4: *Pareto Fronts representing the performance of BCP, ICP, and PCP in terms of water deficit (vertical axis) and flood days (horizontal axis).*

lower deficit values. The DI is indeed providing a crop state information which is particularly valuable to improve the water supply strategies.

The hypervolume indicator (HV) (eq. 5.4) supports a quantitative estimation of the operational improvement imputable to the Drought Index, by capturing the proximity and the distribution of the solutions of the baseline and informed Pareto fronts with respect to the reference (PCP). As PCP is considered the reference performance, its HV value is equal to 1. The HV of BCP is instead equal to 0.327, confirming the large space for improving current operations with additional information. The introduction of DI in the policy produces a value of HV for the ICP equal to 0.387, which means a 18.3% improvement with respect to the baseline solutions.

3.5 Conclusions

FRIDA is an automatic approach which can support drought analysis and characterization in complex highly regulated basins, in a data-driven fashion, i.e., without requiring knowledge of the physical processes taking place in the system.
Basin-tailored drought indexes can enhance water management skills in controlling droughts by warning the operations on the onset, severity, and persistency of a dry spell.
Exploiting available information in a data-driven control scheme can significantly improve the overall system performance at no cost.

Code availability: A detailed description of FRIDA, including codes, and a comprehensive dataset for a test case study, is available at `https://www.ei.deib.polimi.it/?page_id=779`. A code repository for the control policy design of Lake Como can be found at
`https://github.com/mxgiuliani00/LakeComo`

4

When timing matters - misdesigned dam filling impacts hydropower sustainability

Abstract[1]

Decades of sustainable dam planning efforts have focused on containing dam impacts in regime conditions, when the dam is fully filled and operational, overlooking potential disputes raised by the filling phase. Here, we argue that filling timing and operations can catalyze most of the conflicts associated to a dam's lifetime, which can be mitigated by adaptive solutions that respond to medium-to-long term hydroclimatic fluctuations. Our retrospective analysis of the contested recent filling of Gibe III in the Omo-Turkana basin provides quantitative evidence of the benefits generated by adaptive filling strategies, attaining levels of hydropower production comparable with the historical ones while halving the negative impacts to downstream users. Our results can inform a more sustainable filling of the new megadam currently under construction downstream Gibe III, and are generalizable to the almost 500 planned dams worldwide in regions influenced by global climate oscillations.

[1]This work is currently under review as: Zaniolo, M., Giuliani, M., Burlando, P., Castelletti, A., 2020 When timing matters - misdesigned dam filling impacts hydropower sustainability. Nature Sustainability (under review). This Chapter presents an unconventional structure with respect to the rest of the chapters in this thesis as required by Nature Sustainability guidelines. In particular, the Case Study and Results sections are merged, and the Methods section is located after the conclusions.

4.1 Introduction

Hydropower is the dominating renewable electricity source worldwide, accounting for the largest share of energy production and investments allocated to new projects (Zarfl et al., 2015). However, hydropower dam development does not occur without environmental and social costs (Zarfl et al., 2019). Efforts towards sustainable dam planning have addressed strategic dam sizing (Bertoni et al., 2017), individual dam location (Jozaghi et al., 2018; Schmitt et al., 2018), and basin wide siting of dam portfolios (Kondolf et al., 2018; Schmitt et al., 2019) to minimize long term impacts of such infrastructures. Yet, before starting electricity production, dam reservoirs must be filled, withholding a substantial fraction of the river streamflow from downstream users. The rapidity at which a reservoir is filled has direct implications for potential conflicts between upstream and downstream interests. In this phase, precaution towards downstream impacts requires transiting high percentages of inflow, resulting in multi-years, even decadal, filling transients (Yihdego et al., 2017). Conversely, upstream interests (e.g., hydropower production) favor fast impoundment of water, albeit generating critical periods of minimal streamflow downstream. Increasingly variable hydroclimatic regimes characterized by strong interannual oscillations further challenge the design of filling strategies as the same policy can yield very different results if it occurs during a wet or a dry spell.

Historically, the filling of large dams has generated serious international tension. In the Middle East, threats of an armed conflict were raised in 1992, when the filling of the Turkish Ataturk Dam on the Euphrates cut the water flow to downstream Syria and Iraq by 75% (Carkoglu and Eder, 2001). In 2019, the filling of the Ilsu dam on the nearby Tigris reinflamed tensions in the Middle East, in the midst of their unprecedented water, and humanitarian, crisis (Kucukgocmen, 2013). Similar transboundary tensions were generated by the filling of Gibe III, the "most controversial dam in Africa" (The Economist, 2016a), located in the Omo-Turkana Basin (OTB) shared by Ethiopia and Kenya. After Gibe III dam began impounding water (2015-2016), an upsurge of local and international groups contested the insufficiency of the summer flood pulse necessary to support downstream riparian activities, as well as a 2 meters level drop in the downstream Kenyan lake Turkana (Roth, 2019; Avery, 2017). Perhaps the most controversial case, given its global resonance, is the imminent filling of The Grand Ethiopian Renaissance Dam (GERD) on the Blue Nile. In 2020, at the beginning of the tenth year of negotiations, there is still no international filling agreement between Egypt, downstream, demanding guarantees on minimum GERD releases, and Ethiopia, upstream, resolved to maintain discretion in its operations (Gebre, 2019; BBC, 2019; The Economist, 2016b).

State-of-the-art efforts in cooperative upstream-downstream filling strategies have explored the effects of hydrological variability, climate change (King

and Block, 2014; Zhang et al., 2016), and coordination between co-riparian countries (Wheeler et al., 2016) on alternative filling strategies. However, these studies considered filling strategies designed to impound (or release) fixed fractions of inflow or absolute water volumes. In general, even if a preventive agreement is stipulated between the upstream and downstream parties, filling outcomes are still largely dependent on hydroclimatic variability: if the filling occurs during a drought, enhanced impacts are experienced by all downstream sectors. Here, we argue that this limitation can be addressed by advancing dam planning through the design of filling strategies that are adaptive to the alternation of wet spells and drought emergencies (Spinage, 2012) driven by global climate oscillations, aiming to minimize dam impacts during the critical initial filling phase. Accordingly, we propose a retrospective analysis of the recent filling of Gibe III to identify both a favorable filling timing, i.e., *when* to start the filling, and an effective filling policy, i.e., *how* to fill the reservoir. The reference provided by the contested historical filling of the reservoir allows us to investigate the potential of these adaptive solutions to address the tradeoff between upstream and downstream competing interests, and to quantify the role of hydroclimatic variability.

We find that Gibe III filling impacts were disproportionally amplified by an ongoing drought, and show how a more favorable dam filling timing could have been inferred in advance by observing long term climatic oscillations in the basin. Once the optimal timing is established, adaptive filling policies can be designed to better respond to natural hydroclimatic variability, thereby containing downstream alterations without damaging hydropower production. Since a new megadam, Koysha, is currently under construction downstream of Gibe III and is expected to begin filling in 2021 in correspondence with another expected multiyear dry spell, our results suggest the impacts of this new project may be amplified by these unfavorable hydroclimatic conditions, further jeopardizing regional activities and stability of the region.

4.2 Behind filling controversy - the role of timing

In recent years, Ethiopia's domestic electricity demand has witnessed a dramatic increase, propelled by an unprecedented growth in its GDP (Mondal et al., 2017) coupled with an increasing fraction of households connected to the national grid, from 23% in 2011 to 44% in 2017 (World Bank, 2019). Yet, Ethiopia's plans for the electricity sector in the near future are even more ambitious. By 2025, the country is striving for 100% electricity access (International Energy Agency, 2014), a 10-fold increase in power generation capacity since 2013 that would not only cover internal demand, but also allow a substantial electricity export (Asress et al., 2013) and a fully decarbonized economy (Paul and Weinthal,

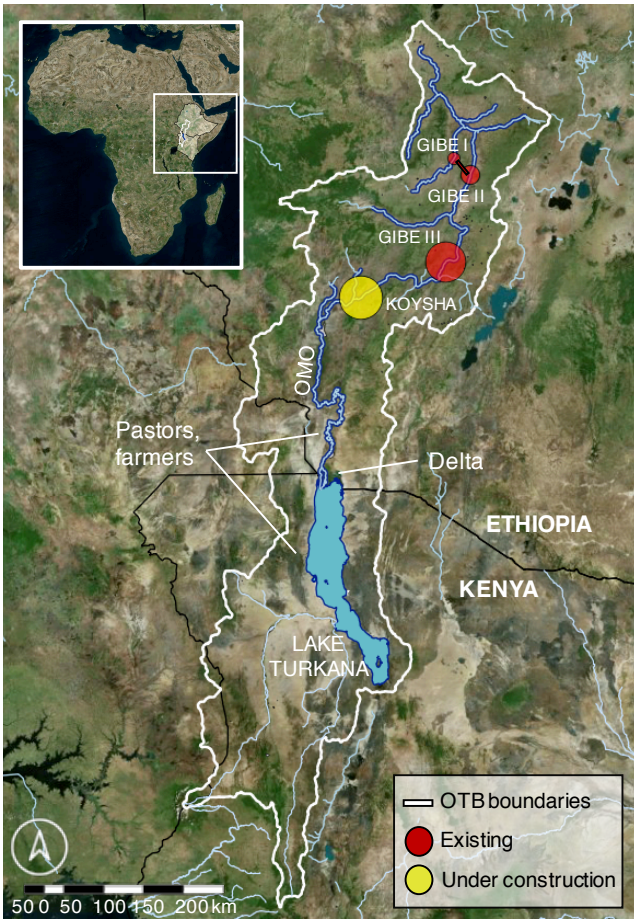

Figure 4.1: ***Geography of the Omo-Turkana Basin (OTB).*** *The Omo river collects the abundant rainfalls of the Ethiopian highlands and streams southwards through the Omo valley contributing about 90% of annual inflow to Lake Turkana, where its outlet forms a complex delta coincident with the Ethiopian-Kenyan border. About 500 thousands pastors and farmers inhabiting the area depend on the Omo or Turkana waters for their livelihoods. The Gibe-Koysha dam cascade regulates the river hydrology, comprising Gibe I and II, the recently completed Gibe III, and the under construction Koysha dam. Marker area is proportional to the dam's installed hydropower capacity.*

2019).

The key to becoming the green battery of Africa is accessing its exceptional renewable resource potential estimated around 60 GW of electric power producible from hydropower, wind, solar, and geothermal sources (Tessama et al., 2013), of which hydropower represents the largest share (45 GW) (Clapham, 2018). The Ethiopian Electric Power Corporation has thus embarked on an ambitious dam building program intended to exploit its abundant water reserves. Among the mega-infrastructures recently built or under construction we count the GERD, on the Blue Nile (Yihdego et al., 2017), along with Gibe III and Koysha, on the Omo river. Gibe III, commissioned in 2015, doubled Ethiopian hydroelectric installed hydropower capacity and has the potential to significantly alter Omo's streamflow regime with its massive reservoir volume of 14.7 billion m^3, corresponding to the average yearly river flow at the dam site. Different from to the other mega-infrastructures, Gibe III is already completed and currently operating at regime conditions, thus serving as a benchmark for alternative filling strategies against historical operations.

The Omo river is one of the largest and steepest Ethiopian rivers, and was a main target of dam expansion given its remarkable reserve of unharnessed hydropower potential. It originates in the Ethiopian Shewan highlands, and streams southwards through a mountainous area before slowing its pace as it meanders through the lower Omo valley (Figure 4.1). At the Ethiopian-Kenyan border, the river forms an extensive delta and contributes about 90% of the inflow to Lake Turkana, an endorheic lake of the Kenyan Rift Valley, and the world's largest desert lake (Avery and Tebbs, 2018).

A three-season meteorological year characterizes the regional climate: a rainy *Kiremt* season (June-September) which contributes the bulk of the annual precipitation through intense convective storm events, a dry *Bega* season (October-January) carried by Arabian desert winds, and a milder wet *Belg* season (February-May) induced by a wet air mass coming from the Indian Ocean (Spinage, 2012). In addition to seasonal variability, marked inter-annual climate variability affects the region as a result of the influence of large scale oscillation patterns in the atmospheric-ocean system (see e.g., Lanckriet et al., 2015; Degefu et al., 2017, and reference therein). Such teleconnections are responsible for frequent severe drought episodes recurring every 5 to 10 years that cause widespread famines in the country, including the catastrophic Ethiopian famine of the mid 1980s (Block and Webb, 2001).

The Omo river hydrology is characterized by a late summer flow peak that conveys the Kiremt rainfall, and reaching about 1000 m^3/s in the lower valley. Local ecosystems and activities largely depend on this flood pulse that enables recession agriculture practices and replenishes grazing lands for livestock, supporting the livelihoods of about 200,000 people in Southern Ethiopia. Reach-

ing Lake Turkana, the flood pulse sustains a biodiverse delta, and produces lake level oscillations that are vital to nutrient circulation, fish spawning, and the regeneration of lake shores grazing area for livestock, a crucial protein source for the 300,000 thousand people inhabiting the poorest region in Kenya (Avery, 2013).

A series of dams was built on the river, including Gibe I (187 MW), Gibe II (420 MW), and Gibe III (1870 MW). The dam cascade will be concluded with the addition of Koysha (2160 MW), currently under construction and expected to be completed in 2021. Gibe III received some opposition when the project was made public (Avery and Eng, 2012), but an unprecedented upsurge of national and international criticism erupted when the reservoir behind the dam started to impound water (The Economist, 2016a). Reports say that in 2015 and 2016, the flood pulse downstream of the dam did not occur or was severely dampened, and thus inadequate to serve its functions (Human Rights Watch, 2017; Avery, 2017), dramatically damaging the river-related ecosystems and downstream activities relying on it (UNESCO, 2018; IUCN, 2018). Simultaneously, during the filling of Gibe III, the level of Lake Turkana dropped 1.7 meters, of which over 1 meter fell in the first year (Avery and Tebbs, 2018). Were these dramatic impacts the inevitable price to pay for dam development, or was (more) sustainable filling possible?

To address this question, we analyzed the historical filling strategy and explored alternative options by changing both filling timing and operations. Since no official record of Gibe III operations during the filling is publicly available, we first reconstructed the historical strategy using satellite imagery and a simulation model of the OTB (see Methods and Supplementary Figure A.1). Overall, the reconstructed system dynamics (Figure 4.2) is coherent with news reports (e.g., Human Rights Watch, 2017; Avery, 2017), showing the largely impounded 2015 and 2016 Kiremt season streamflow, a fast level increase in Gibe III, and a steep drop in the level of Lake Turkana during the initial dam filling.

To understand the role of timing (i.e., when dam filling is initiated), we then performed a retrospective analysis by simulating the reconstructed historical filling policy, assuming this took place in different years featuring diverse hydroclimatic conditions. The annual cumulated precipitation in the basin from 1999 to 2018 shows a clear multiyear climatic oscillation that can be well approximated by the sum of three harmonics (Figure 4.3), associated with the ocean-atmospheric interactions persisting in the region (see Methods). Gibe III filling began in 2015, at the negative peak of a prolonged downwards phase in precipitation abundance; intuitively, this represented an unfortunate timing to immobilize a large water volume into a reservoir.

Looking at the climatic oscillations, we analyzed alternative timings corresponding to upwards (2007 and 2013) and downwards (2009 and 2015, which

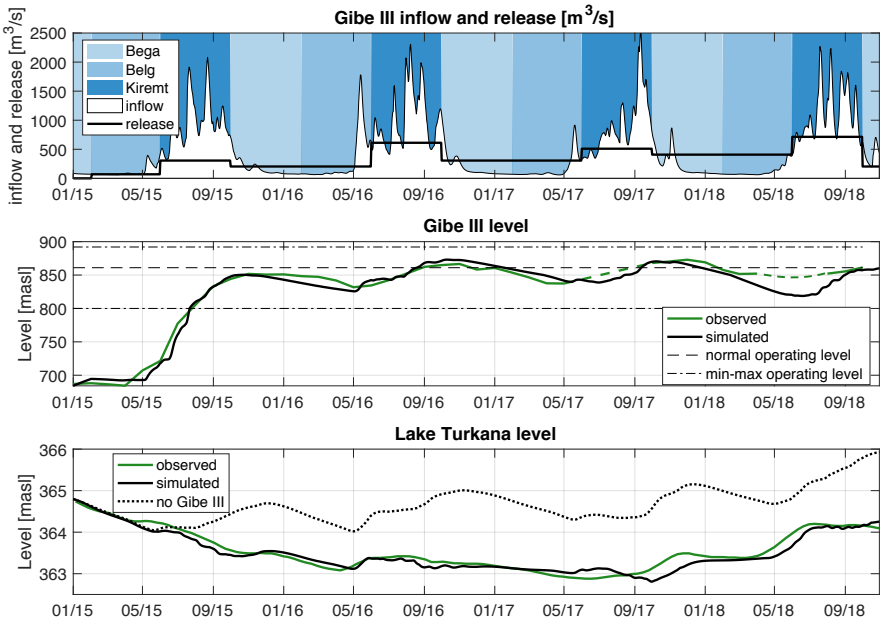

Figure 4.2: ***Reconstructed historical filling strategy.*** *Gibe III reservoir reached its normal operating level within its first two years of operations by impounding the near totality of the 2015 Kiremt season inflow, and a significant fraction of 2016's. In the two following years, the Gibe III level oscillated around its operational level as a consequence of a release pattern that increased low flows and reduced high flows with respect to natural Omo hydrology. Simultaneously, Lake Turkana suffered a two meter level drop with respect to a simulation of a scenario in which Gibe III was not built. While the trajectory of observed Lake Turkana levels is publicly available, the Gibe III level trajectory was reconstructed from Sentinel 2 image classification (see Methods).*

is the historical starting date) phases in precipitation abundance. Results show that the filling outcomes are strongly determined by the harmonic phase in which the filling starts (panels b in Figure 4.3). According to all considered indicators (see Methods and Supplementary Information for details about their formulation) reflective of both upstream interests (i.e., hydropower production, final Gibe III level) and downstream preservation (i.e., drop in Lake Turkana level, flood pulse magnitude), the worst timing to initiate dam filling would have been 2009, which corresponded with the onset of a multi-year dry spell. Conversely, starting the filling in 2013 would have benefited all sectors involved and contained the sharp intersectoral conflict observed in 2015. In particular, 2013 would have favored upstream water users yielding an additional 124 GWh/year in hydroelectricity, corresponding to a revenue of 17.36 Million USD in electricity export in its first 2 years of life, assuming the excess energy was sold to Kenya at the agreed upon price of 0.07 USD/kWh (Fitiwi Tekle, 2016). In addition, the drop in the level of lake Turkana could have been contained to 1 m instead of 1.63 m, and the flood pulse magnitude increased by 42% with respect to that actually observed.

This analysis shows that, in the future, the projection of the harmonic trends of precipitations could preventively inform a forward-looking planning of the timing of Koysha construction, in order to synchronize filling to a wet spell, rather than aggravating the expected natural water scarcity situation following a dry spell. Koysha filling is expected to start in 2021, again at the bottom of a steep decline in precipitation foreseen in the 2 previous years, thus likely magnifying the stress of a long running water shortage. Instead, beginning the filling one year later, at the inversion of the precipitation trend, would significantly reduce the impact downstream and produce more benefits upstream during the filling transient, however generating a one-year delay in hydropower production.

4.3 Forecast-informed adaptive filling

The historical filling operations would not have been able to achieve a sufficient flood pulse and still would have recorded a significant drop in the level of Lake Turkana even if the filling timing had been favorable (Figure 4.3b.3-4). These shortcomings motivate searching for alternative, adaptive filling strategies that better respond to the interannual hydroclimatic variability.

Taking advantage of advanced Machine Learning and data mining techniques, we synthetize global datasets of climate oscillations (Supplementary Table A.1) into a compact drought index forecast, namely the Standardized Precipitation and Evaporation Index (SPEI), which is representative of upcoming hydrometeorological anomalies at the Omo-Turkana basin scale (see Methods and Sup-

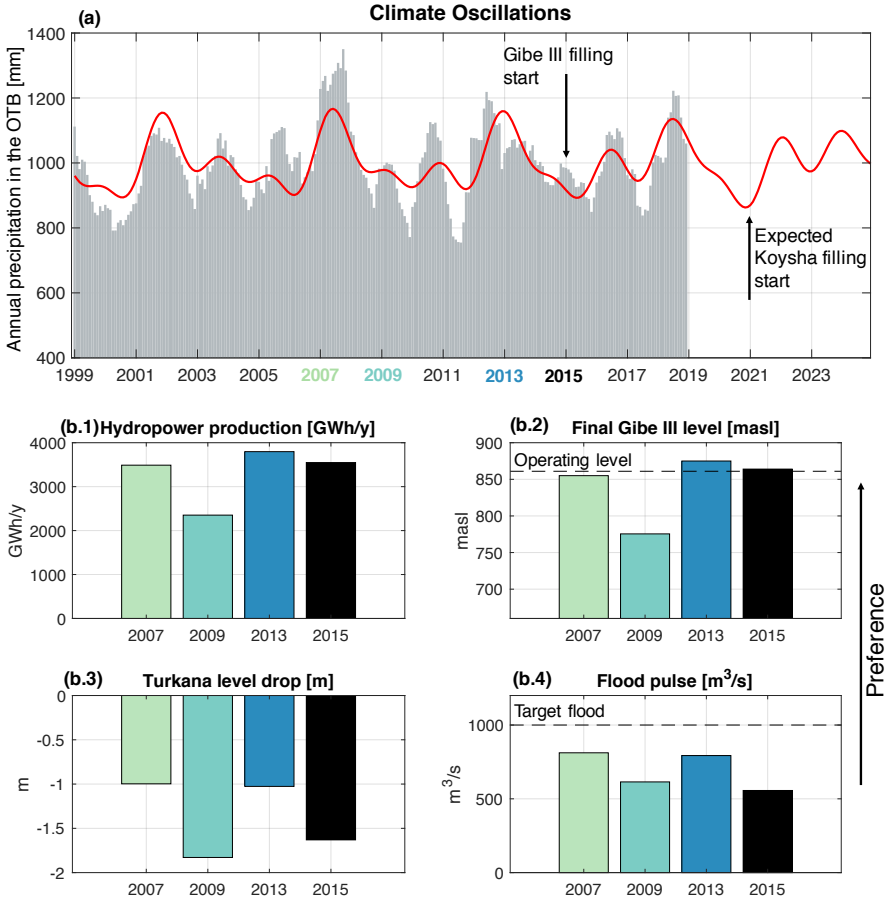

Figure 4.3: ***Climatic oscillations can inform a favorable timing for filling.*** *A pattern of harmonic climatic oscillations governs the magnitude of annual rainfall in the OTB (panel (a)). Filling the Gibe III reservoir during an upwards phase of water availability (e.g., 2013), instead of a downwards phase as historically, could have resulted in a more efficient, and less contentious filling (panels (b)). By projecting the harmonic trends into the future, we advise to delay the filling of Koysha by one year and to begin in 2022 instead of the planned 2021, as the additional stress caused by a poorly timed filling stress could have detrimental social and ecological consequences.*

plementary Figure A.2). Adaptive filling policies use the forecasted drought index to speed up the filling process during wet spells, and, conversely, increase releases during dry seasons to sustain downstream activities (see Methods).

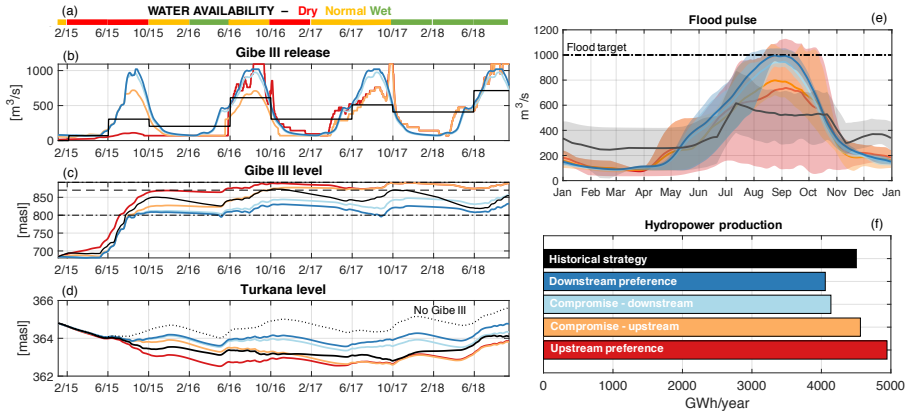

Figure 4.4: ***Adaptive filling strategies can reduce filling impacts.*** *The seasonal forecasts of the Standardized Precipitation and Evaporation Index expressed in terms of dry, normal, and wet conditions with respect to seasonal average (panel (a)) inform the designed adaptive filling strategies (panels (b,c)). Different colors correspond to adaptive strategies with different tradeoffs between upstream and downstream competing interests, blues for more environmentally inclined, and reds for hydropower-inclined strategies, while the historical strategy is represented in black. Adaptive strategies demonstrate the ability to significantly reduce downstream impacts on Lake Turkana (panel(d)) and average river hydrology (panel (e), where the shaded areas refer to the inter-annual variability) while remaining within a contained range of historically produced hydropower (panel (f)). The figure illustrates 4 different tradeoff solutions, while the complete set of results is reported in Supplementary Figure A.3.*

A total of over one hundred adaptive filling strategies were designed to provide a thorough exploration of the basin sectoral trade-offs (see Methods and Supplementary Figure A.3). In order to benchmark informed strategies with historical operations, we considered begining the filling in 2015 for all alternatives. The SPEI seasonal forecasts (Figure 4.4a) confirm that Gibe III filling started during a drought, but water availability conditions improved towards mid-2017. Tradeoffs are evident between upstream and downstream interests, whereby strategies attaining high hydropower production are also associated with large negative impacts downstream. Notably, the Downstream Preference policy ensures high Gibe III releases (panel (b)) especially in the first years (2015-2016), nearly halving lake Turkana level drawdown with respect to observed conditions (panel (d)) and preserving the natural flood pulse in the delta

(panel (e)). However, this policy is estimated to produce a 9% lower hydropower production with respect to the historical one, corresponding to lost revenue of 28.3 Million USD per year, in the first four years. Conversely, the Upstream Preference policy surpasses the historical hydropower production (+30.9 Million USD/year) by implementing a fast filling strategy that reaches the Gibe III operating level within the first year, at the cost of significant alterations to Lake Turkana levels. Finally, the Compromise-upstream policy achieves a historically equivalent hydropower production, while maintaining a significantly more natural hydrology downstream in terms of flood pulse, which is, on average, nearly 300 m^3/s higher than historically observed during the expected peak in late August.

The entire ensemble of adaptive policies produced (Supplementary Figure A.3) thoroughly explores the space of compromises and trade-offs between the conflicting interests coexisting in the OTB. Hydropower production of the adaptive policies ranges from -12% below to 9.7% above historical production. The drop in Lake Turkana levels measured between January 2015 and November 2018 ranges from 1.2 m (comparable to the historical 1.1 m) to just a few centimeters across the adaptive policies. Interestingly, the average magnitude of the flood pulse in late August, historically just over 500 m^3/s, is considerably improved by the entire ensemble of adaptive policies, which obtain a minimum of 720 m^3/s and a maximum of 1000 m^3/s. Overall, by considering the entire range of adaptive strategies we notice the potential to largely contain environmental alterations with a comparatively small to nonexistent loss in hydropower production.

4.4 Discussion and Conclusions

Sustainable dam planning has paved the way towards more socially and environmentally inclusive hydropower development that focuses on limiting dam-induced socio-environmental costs during regular operations. Yet, the initial filling phase of a dam can generate critical impacts by withholding in the reservoir a substantial fraction of the river's streamflow, significantly reducing downstream water availability. Hydrological variability can play a key role in magnifying or containing the stress of filling: if the filling occurs during a dry spell, the basin is further exposed to water shortage and intersectoral tensions. This is the case analysed in this paper, where an ongoing drought magnified the filling impacts of the Gibe III dam, thus inducing its association with the label of "most controversial dam in Africa" (The Economist, 2016a). In the proposed retrospective analysis, we demonstrate that consideration of climate oscillations could have informed a more favorable filling timing and operations, and significantly contained the associated social and environmental costs. Koysha

dam, located downstream of Gibe III, is again at risk of synchronizing its filling with an upcoming drought, further endangering the already precarious socio-environmental conditions of the Omo-Turkana Basin. While these quantitative results refer to this specific case study, the novel approach and tools we propose in this work can be generalized to other river basins. Currently, of the nearly 650 medium to large dams under construction in the world (Zarfl et al., 2015), 70% are being built in regions under the influence of the El Niño Southern Oscillation, the prevalent global interannual signal of climate variability (Ward et al., 2014a; Lee et al., 2018) (Figure 4.5). In these areas, teleconnection analysis has the potential to increase our predictive skills in anticipating hydrological variability in the medium-to-long term, which can be exploited to minimize filling impacts. operations during the critical initial filling phase.

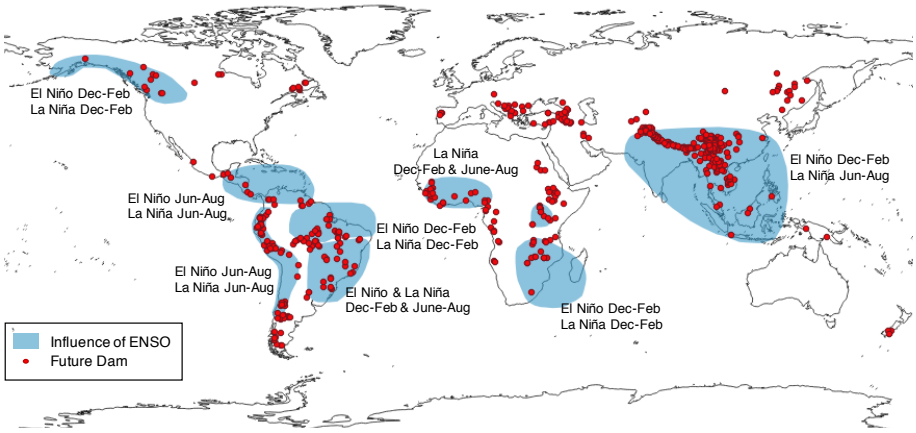

Figure 4.5: ***Future dams overlap regions with a strong ENSO influence.*** *The blue points indicate the locations of medium-to-large future hydropower reservoirs and dams, extracted from the FHReD database published in (Zarfl et al., 2015). Dam height is generally employed to discern between small, medium, and large dams, but in the absence of this information, we consider as medium-to-large the hydropower projects with an installed capacity greater than 150 MW, retaining a total of 642 dams of the over 3700 reported in the database. A red shade highlights the areas of the globe that are most affected by El Niño and La Niña oscillations (Lindsey, 2016). Over 70% of medium-to-large future dams are located in areas affected by the ENSO teleconnection.*

It is, however, important to consider that, while on the one hand filling a dam during a dry year can jeopardize water-related activities and overall basin stability, on the other hand, postponing dam filling will delay the project's energy generation, its expected economic returns, and the benefits of electrification to its patrons. Additionally, inferring a favorable filling timing by projecting past climatic trends in the future is associated with a level of uncertainty enhanced by

ongoing climate change trends. It is thus recommended to consider a portfolio of alternative renewable energy sources (e.g., solar, wind, geothermal, biomass, and tides), in addition to hydroelectricity, that can compensate for the delay in hydropower production possibly required by a sustainable filling strategy.

4.5 Methods

4.5.1 Omo-Turkana Basin Model

The model of the Omo-Turkana Basin relies on a combination of TOPKAPI-ETH, a spatially distributed hydrological model, with a dynamic, conceptual model of Gibe III, the lower Omo Valley, and lake Turkana.

TOPKAPI-ETH is a spatially distributed hydrological model that improves the original TOPKAPI rainfall-runoff model (Ciarapica and Todini, 2002). The model performs a spatio-temporal simulation of the main hydrological processes at the basin scale, accounting for runoff generation, routing, and evapotranspiration, including snow and glacier dynamics when necessary (see e.g., Paschalis et al., 2014, and reference therein). Spatial heterogeneity of the OTB basin is represented by discretizing the domain with a regular grid of 1 km^2 resolution, while the temporal dynamics are characterized at a daily time step. The model inputs are daily values of precipitation, temperature, and cloud cover; the model outputs are Gibe III inflows, lateral contributions in the Lower Omo valley (between Gibe III and lake Turkana), and the additional inflows to lake Turkana provided by the Turkwel and Kerio rivers in Kenya. Daily rainfall estimates are available from the TAMSAT archive with a 4km resolution for the African continent (Maidment et al., 2017), and satellite-based temperature records from MERRA-2 at a resolution of 0.5° x 0.625° (Rienecker et al., 2011).

The daily dynamics of Gibe III and lake Turkana are described by the mass balance of their water volumes, where the release volume of Gibe III is determined by the simulated filling policies, followed by a regime policy activated when the filling has completed (i.e., when the level of Gibe III reaches the normal operating level equal to 851 masl). Geomorphological and technical characteristics of Gibe III reservoir, dam, and power plant are published in the project's impact assessment (Ethiopian Electrical Power Company (EEPCO), 2009). Lake Turkana, instead, is an endorheic lake, and the only water output is due to evaporation. According to the daily time-step adopted in the model, the reach of the Omo river downstream from Gibe III is modelled as a plug-flow canal in which the velocity and direction of flow are constant. A transit lag time of $\mathtt{lag} = 18$ days from Gibe III to lake Turkana is estimated from the TOPKAPI-ETH simulations.

Different objective functions representing the main hydropower and environmental interests were formulated through a participatory process involving

key stakeholders active in the system, that participated in dedicated meetings called Negotiation Simulation Labs held during the DAFNE research project (http://dafne-project.eu/): maximization of hydropower production at Gibe III and minimization of the average daily squared distance between the simulated flow in the Omo delta and the annual hydrograph in natural conditions. Moreover, the maximization of the Gibe III level at the end of the filling transient is included in the design of the filling strategy to design solutions that, for a given hydropower and environmental performance, favor a fast rather than unnecessarily slow filling (see the Supplementary Information for the detailed mathematical formulation of these objectives). The presence of such clearly conflicting interests does not allow the design of a unique optimal solution, but rather a set of non-dominated (or Pareto optimal) solutions. A policy is defined as Pareto-optimal if no other solution gives a better value for one objective without degrading the performance in at least one other objective.

The optimal operation of Gibe III in regime conditions (after the filling has completed) is designed via Stochastic Dynamic Programming (SDP, Bellman, 1957) and is consistent across all policies. The selected tradeoff for the regime policy is the one located at the Pareto front knee of the multi-objective SDP optimization. Our proposed adaptive filling policy determines the dam release in a given day as a function of the cyclostatonary average streamflow for that day prior to dam construction. Specifically, the cyclostationary streamflow is scaled proportionally to the expected hydrological conditions for the incoming season according to the forecast of the Standardized Precipitation and Evaporation Index (SPEI) (see Methods). In this formulation, the three scaling factors associated with the three classes of SPEI (i.e., dry, normal, wet) are the decision variables of the filling optimization problem; we searched the optimal value of these factors with respect to hydropower production, environmental impact, and filling duration by using the self-adaptive Borg Multi-Objective Evolutionary Algorithm (Hadka and Reed, 2013). The Borg MOEA has been shown to be highly robust across a diverse suite of challenging multiobjective problems where it met or exceeded the performance of other state-of-the-art MOEAs (Hadka and Reed, 2012; Zatarain et al., 2016).

4.5.2 Historical filling strategy

The historical Gibe III filling strategy (i.e., sequence of dam inflows and releases during the first years) is not publicly available, and was thus reconstructed for the purpose of this study. We derived the sequence of dam releases by assuming that turbines were operated at full capacity (corresponding to maximum efficiency), and that the release from the dam was maintained constant within a season. The inflows were obtained from TOPKAPI-ETH hydrological simulations. The simulation of this filling strategy allows the reconstruction of Gibe

III and Lake Turkana levels (see Figure 4.2). While records of Lake Turkana levels are available in the Database for Hydrological Time Series of Inland Waters (DAHITI) (Schwatke et al., 2015), the observed Gibe III level was derived from Sentinel 2 satellite images classification. Sentinel 2 images are available for the area every 5 days at a 30 meters spatial resolution (Drusch et al., 2012). The images recorded in the same month are aggregated in the attempt of filtering cloud occlusion. We then performed a land cover classification of the composite images using a combination of NDWI (Gao, 1996) and NDVI (Rouse Jr et al., 1974) indexes, and derived an estimate of the reservoir surface area from the water pixels count (see Supplementary Figure A.1). Using the reservoir bathymetry we finally estimated the corresponding trajectory of the Gibe III level. The coefficient of determination of the simulated filling strategy with respect to the historical observations are equal to $R^2_{\mathrm{GibeIII}} = 0.9795$, $R^2_{\mathrm{Turkana}} = 0.9075$ computed at a daily time step.

4.5.3 Empirical derivation of climatic oscillations

The influence of climate oscillations on Ethiopian meteorology can be decomposed into three contributing phenomena associated with the three oceans (Lanckriet et al., 2015). The climatic oscillations shown in Figure 4.3 are therefore empirically derived by summing three single term Fourier series $h^i, i = 1, 2, 3$ of the form

$$h^i(x) = a0 + a1 * \cos(x * w) + b1 * \sin(x * w) \tag{4.1}$$

where $a0, a1, b1$ and w are the parameters to be calibrated, and x is the signal to be approximated. In particular, the first harmonic is specified as $h^1(p)$, where the signal p is the monthly timeseries of the annual cumulated precipitation in the OTB. For the second harmonic $h^2(p')$ the signal to be approximated is computed as the residual precipitation $p' = p - h^1(p)$ that is not captured in $h^1(p)$, and analogously, $h^3(p'')$ is calibrated on the second residual $p'' = p' - h^2(p')$. The resulting coefficient of determination is R^2 = 0.5273.

4.5.4 Performance of alternative timing of reservoir filling

We investigate the role of filling timing by assessing how the system would have responded to the filling stress if it had started in different years. To do so, we simulate the first 24 months of the filling subject to the hydrology of different years. System performance is then evaluated in terms of 4 indicators (see the Supplementary Information for the detailed mathematical formulation):

1. Mean annual hydropower production during the 24 months filling period;
2. Final Gibe III level at the end of the 24 months;

3. Final Turkana level drop relative to the initial lake level;

4. Flood Pulse defined as the maximum flow reaching the delta during the flood season of August-September averaged across the alternative filling years.

4.5.5 Seasonal forecasts

To develop season-ahead hydrological forecasts of water availability we use the Climate State Intelligence (CSI) framework (Giuliani et al., 2019), an extension of the Nino Index Phase Analysis (Zimmerman et al., 2016), which employs Artificial Intelligence tools to search relevant circulation patterns at the global scale that serve as predictors for meteorological anomalies at the local scale. The CSI framework is articulated in four steps:

1. *Phase distinction:* given a teleconnection signal, the associated teleconnection index is used to group the years in the time horizon into a specified number of phases, that are then evaluated individually. For instance, considering the El Niño Southern Oscillation (ENSO), the Multivariate Enso Index (MEI, Wolter and Timlin, 2011) is used to distinguish El Niño and La Niña years, uncovering possible asymmetries in the effect of a signal on the local scale, e.g., if in a given region El Niño years are associated with a wet spell, La Niña years are not necessarily associated with a dry spell.

2. *Univariate linear forecast:* For each phase of the climate signal, the procedure identifies relevant correlations between a gridded dataset of preseason SSTs and the local variable, retaining SST regions correlated at 95% significance level or above, assessed with the p-value test. Selected SST regions are then spatially aggregated via Principal Component Analysis (Jolliffe, 2002, PCA, see). As in previous applications (Zimmerman et al., 2016; Giuliani et al., 2019), only the first, most informative, PC is retained as a predictor for a linear forecast model of the local variable y:

$$\widehat{y}_t = \beta PC_{t-1} + \alpha \tag{4.2}$$

   A leave-one-out cross-validation is performed to calibrate model coefficients α and β.

3. *Test of Correlation Significance:* A Monte Carlo analysis is run to test the statistical significance of the obtained correlations by randomly shuffling the time series of the local variable to be predicted and repeating the above described steps with unshuffled SSTs and teleconnection index time series. SST grids must be present at the 95% significance level 90% of the time to be retained as predictors for the multivariate non-linear forecast.

4. *Multivariate non-linear forecast:* The most informative climate signals for the region of interest are then chosen based on their linear model accuracy and significance. A multivariate non-linear model (Extreme Learning Machine (Huang et al., 2006a)) is then cross-validated on the selected climate signals to produce a data-driven seasonal forecast of the local variable.

In this analysis, we considered 16 teleconnection signals at different time and spatial scales over the 21-year time horizon from 1998-2018 dictated by the precipitation data availability. We obtained Global Sea Surface Temperature anomalies from the NOAA's Extended Reconstructed SST (ERSST) Version 3b, a global monthly gridded dataset with a spatial resolution of 2.5 degrees available at `https://www.noaa.gov`. From the same source we retrieved the time series of teleconnection indexes. The local variable forecasted is the Standardized Precipitation and Evaporation Index (SPEI) drought index (Vicente-Serrano et al., 2010), which has proven to be more effective than the Standard Precipitation Index (SPI) to characterize hot and arid climates, where the evapotranspiration has a key role in depleting the soil moisture and becomes one of the main drivers of a drought (Lorenzo-Lacruz et al., 2010). SPEI substitutes the precipitation used for SPI computation with a net precipitation, by substracting the Potential Evapo-Transpiration (PET) estimated from temperature and latitude via the Thornthwaite method (Thornthwaite et al., 1948). In this work, the SPEI index with a 6 months cumulation span is used to characterize seasonal water availability in three classes according to a classification commonly used in the literature (Spinoni et al., 2014): *dry* (SPEI$<$-0.5), *normal* (-1$<$SPEI$<$1), and *wet* (SPEI$>$0.5). The 6 months time span was selected as frequently used to characterize medium-term hydrological conditions.

Phase specific accuracy of the univariate linear forecast models in cross-validation is reported in Supplementary Table A.1, along with corresponding statistical significance. Balancing accuracy and significance, we selected three teleconnection signals, namely the NAO, related to a climatic oscillation that originates in the Atlantic ocean, the PNA, which originates in the Pacific Ocean, and SEIO, which originates in the Indian Ocean. This choice aligns with the findings in Lanckriet et al. (2015) that demonstrates that three overlapping climatic oscillations, each originating in a different ocean, contribute to the Ethiopian climate. The first Principal Components related to these signals are the inputs of the multivariate Extreme Learning Machine forecast model, which is used to generate a 10-member forecast ensemble. The ensemble average is retained for classifying the upcoming season (see Supplementary Figure A.2).

Part II

A Posteriori Feature Representation for Multi-Objective Reinforcement Learning

The *a posteriori* approach to feature representation learning evaluates the optimality of a policy input set by assessing the performance of the policy conditioned upon it. In this chapter, the application of such an approach is the vehicle for a meta-analysis on the role of variable processing in conditioning a control policy. Inflow forecasts constitute a very intuitive asset for a reservoir control policy as they are a direct prediction of future system disturbance. However, long term hydrological forecasts are generally conditioned upon Sea Surface Temperatures (SSTs) in appropriate locations, processed through a multi-step modeling chain that derives first a precipitation forecast and, through a hydrological model, a reservoir inflow forecast (Cloke and Pappenberger, 2009; Sharma, 2000b; Block and Rajagopalan, 2007).

With the proposed Climate State Intelligence (CSI) framework, we design three alternative policy representations comprising, in addition to a basic input set, i) raw SST observations, ii) processed precipitation forecasts, and iii) highly-processed lake inflow forecasts for the Lake Como control problem. In an *a posteriori* comparison of policy performance it appears that the heavy processing needed for inflow forecast design introduces modeling bias that ultimately degrades its policy representation skills, compared to less processed counterparts.

This Part contains the following work:

- Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., Block, P., 2019. Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. Water Resources Research 55, 9133-9147.

5

Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations

Abstract[1]

Increasingly variable hydrologic regimes combined with more frequent and intense extreme events are challenging water systems management worldwide. These trends emphasize the need of accurate medium- to long-term predictions to timely prompt anticipatory operations. Despite in some locations global climate oscillations and particularly the El Niño Southern Oscillation (ENSO) may contribute to extending forecast lead times, in other regions there is no consensus on how ENSO can be detected and used as local conditions are also influenced by other concurrent climate signals. In this work, we introduce the Climate State Intelligence framework to capture the state of multiple global climate signals via artificial intelligence and improve seasonal forecasts. These forecasts are used as additional inputs for informing water system operations and their value is quantified as the corresponding gain in system performance. We apply the framework to the Lake Como basin, a regulated lake in northern Italy mainly operated for flood control and irrigation supply. Numerical results show

[1]This work has been published as: Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., Block, P., 2019. Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. Water Resources Research 55, 9133-9147.

the existence of notable teleconnection patterns dependent on both ENSO and the North Atlantic Oscillation over the Alpine region, which contribute in generating skilful seasonal precipitation and hydrologic forecasts. The use of this information for conditioning the lake operations produces an average 44% improvement in system performance with respect to a baseline solution not informed by any forecast, with this gain that further increases during extreme drought episodes. Our results also suggest that observed preseason SST anomalies appear more valuable than hydrologic-based seasonal forecasts, producing an average 59% improvement in system performance.

5.1 Introduction

Increasingly variable hydrologic regimes combined with more and more frequent and intense extreme events are challenging water systems management worldwide (Dai, 2011; Trenberth, 2011). Additional pressures generated by global trends in population growth and rising economic prosperity are expected to increase the demand for energy, food, and water (Padowski et al., 2015; Rodell et al., 2018). These variable and evolving conditions require further flexibility of water systems operations to activate early actions and decisions, possibly informed by accurate medium- to long-term predictions (Ziervogel et al., 2010; Benson, 2016).

Most existing water systems are currently operated with static rules conditioned on basic information systems including the day of the year and the storage, and, in few cases, also the previous day inflow (Hejazi et al., 2008b). Yet, the utility of forecasts in reservoir operations has long been recognized (e.g., Kelman et al., 1990; Kim and Palmer, 1997a; Faber and Stedinger, 2001), and recent studies illustrate promising applications using hydrologic forecasts spanning variable time scales (e.g., Anghileri et al., 2016; Turner et al., 2017; Lu et al., 2017; Nayak et al., 2018), possibly conditioned on global climate oscillations (e.g., Gelati et al., 2011; Giuliani and Castelletti, 2019; Libisch-Lehner et al., 2019). Hydrologic forecasts are typically generated via either dynamic, process-based climate models (see Yuan et al., 2015, and references therein) with outputs (e.g., precipitation and temperature) fed into hydrologic models (Cloke and Pappenberger, 2009), or via empirical, data-driven models (Sharma, 2000b; Block and Rajagopalan, 2007), which produce either meteorological forecasts to feed hydrologic models or directly predict future streamflows. However, neither approach has been shown to consistently outperform the other, with dynamic models often limited by resolutions and initialization procedures (Georgakakos et al., 2004) and empirical models constrained by short observational records and stationarity assumptions (Block and Goddard, 2012).

Both dynamic and empirical forecast models generally rely on variability

in sea surface temperature as the main source of predictability at seasonal or longer lead-times (Palmer and Hagedorn, 2006). However, the contribution of global climate oscillations to local predictability depends on the degree to which local conditions are affected by these global climatic anomalies. The El Niño Southern Oscillation (ENSO) is generally considered the dominant inter-annual signal of global climate variability (McPhaden et al., 2006; Ward et al., 2010, 2014b). ENSO is a coupled ocean-atmosphere phenomenon observed over the tropical Pacific Ocean with a 2-7 years period, with El Niño representing the oceanic component and Southern Oscillation the atmospheric one (Sarachik and Cane, 2010). In normal conditions the Walker circulation determines the atmospheric circulation over the tropical Pacific with trade winds blowing westward and moving warm moist air and warm surface water towards the western Pacific while keeping the central and eastern Pacific relatively cool. During an El Niño event, trade winds weaken or may even reverse, allowing the warm water from the western Pacific to move into the central and eastern tropical Pacific. Droughts in Indonesia and eastern Australia become far more common during El Niño events, while rainfall falls on the normally arid coasts of Perù and Equador. La Niña events are characterized by an intensification of the Walker circulation, with greater convection over the western Pacific and stronger trade winds, usually associated with exceptionally warm and wet conditions in the the western part of the tropical Pacific Ocean while the eastern part turns cool and dry. The ENSO teleconnection is therefore well defined in some locations, such as the United States (e.g., Kahya and Dracup, 1993; Hamlet and Lettenmaier, 1999; Grantz et al., 2007), western South America (e.g., Grimm and Tedeschi, 2009; Poveda et al., 2011), or Australia (e.g., Sharma et al., 2000; Chiew et al., 2003). Yet, little to no consensus exists on how ENSO influences other regions, including Europe, where local conditions depend on the concurrent state of other climate signals (Zanchettin et al., 2008; Steirou et al., 2017), particularly the North Atlantic Oscillation (NAO) (e.g., Kingston et al., 2006).

In this paper, we propose a novel framework called Climate State Intelligence (CSI) which aims to use artificial intelligence for producing seasonal hydrologic forecasts based on multiple global climate signals and assessing their value on operational decisions (Figure 5.1). The use of multiple climate signals ensures the portability of this framework to different geographic locations, including regions where traditional teleconnections are weak. The framework extends the Nino Index Phase Analysis (NIPA) originally proposed by Zimmerman et al. (2016) to enable capturing the concurrent state of multiple climate signals, such as ENSO and NAO. Adopting an empirical, data-driven approach based on multivariate Extreme Learning Machine models (Huang et al., 2006b), we use the detected teleconnections and other observed preseason sea surface temperature (SST) anomalies to forecast local meteorological variables on a sea-

sonal time scale. The resulting forecasts are subsequently transformed using a dynamic hydrologic model into streamflow predictions, which are used as additional inputs for informing water systems operations. Finally, we apply the Information Selection and Assessment (ISA) framework proposed by Giuliani et al. (2015) to compute the value of the generated hydrologic forecasts with respect to a baseline solution that does not use any forecast information.

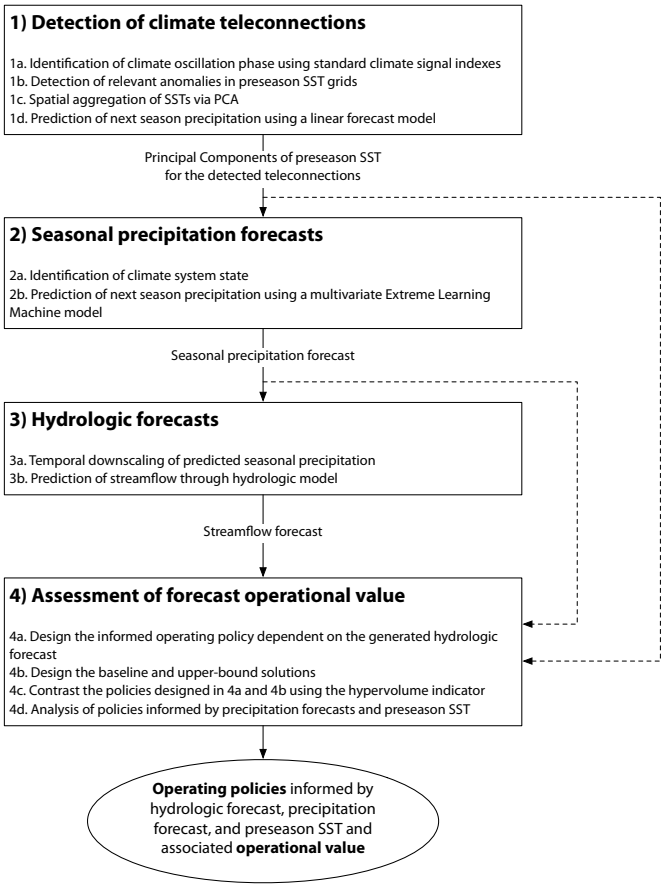

Figure 5.1: *Overview of the Climate State Intelligence (CSI) framework for the generation of seasonal hydrologic forecasts based on global climate oscillations and the assessment of their operational value. The dashed lines represent the operational value assessment of the outputs produced in the first two steps of the framework.*

However, while hydrologic (i.e., streamflow) forecasts might be an intuitive approach to improving water reservoir operations (e.g., Georgakakos and Graham, 2008; Block, 2011), they require a chain of models that also introduces modeling errors and forecast biases, which may diminish the estimated oper-

ational value. This prompts the following key questions: do we really need forecast models to make better decisions? Do hydrologic forecasts necessarily lead to superior reservoir operations? Can the original (observed) sources of predictability (e.g., preseason SST) be as valuable as the hydrologic forecast model outputs? To address these questions, we compare the operational value of the hydrologic forecast with two alternative pieces of information produced in the first two steps of our framework: (i) the seasonal meteorological forecasts, which only require the precipitation forecast model and eliminate the hydrologic model biases, and (ii) the observed preseason SSTs, which remove the entire forecast model chain.

We apply the CSI framework to the Lake Como basin, a regulated lake in northern Italy which is mainly operated for flood control and irrigation supply. The lake catchment, located in the Italian lake district, is characterized by a mixed snow- and rain- dominated hydrology (Giuliani et al., 2016a). Despite expected benefits to seasonal irrigated agriculture (Denaro et al., 2017b), the lack of recognized teleconnections in this region has historically limited the skill of seasonal forecasts (Li et al., 2017), precluding their integration into the lake regulation which, today, is simply informed by the day of the year and the lake level (Todini, 2014; Denaro et al., 2017b). A number of studies investigate the correlation between NAO phases and snow dynamics in the Alpine region (e.g., Beniston, 1997; Scherrer et al., 2004) and suggest that the negative phase of NAO is typically associated with above average temperature and precipitation in Southern Europe, leading to more liquid than solid winter precipitation and accelerated melting processes. Other studies, however, conclude that the role of NAO in the Alpine region is marginal (e.g., Durand et al., 2009; López-Moreno et al., 2011). A lack of consensus also exists regarding ENSO teleconnections; some studies suggest significant correlations in the region (e.g., Folland et al., 2009; Brandimarte et al., 2011), while others find marginal ENSO influence (e.g., Efthymiadis et al., 2007; Bartolini et al., 2009; Shaman, 2014).

In summary, this paper provides a three-fold contribution: (i) the application to the Lake Como system advances the understanding of hydroclimatic variability in the Alpine region and its dependency on global climate oscillations; (ii) the CSI framework introduces a formalized procedure to quantify the role of the propagation of error introduced by forecast models with respect to the forecast operational value; (iii) our framework provides an emblematic demonstration for the potential of Artificial Intelligence tools in supporting water management during extreme events (Rolnick et al., 2019).

5.2 Study site

Lake Como is a sub-alpine lake in the Italian lake district, northern Italy (Figure 5.2). It is the third-largest lake in Italy and reaches a maximum depth of 410 m, which makes it the fifth-deepest lake in Europe. The lake is shaped like an inverted 'Y' surrounded by mountains and primarily fed by the Adda River, which also serves as the only exit point of the lake. The lake catchment has an area of 4,733 km^2, with approximately 90% in Italy and 10% in Switzerland. The hydrologic regime is snow-rainfall dominated, characterized by relatively dry winters and summers and high peaks of flow in spring and autumn, due to snowmelt and rainfall, respectively.

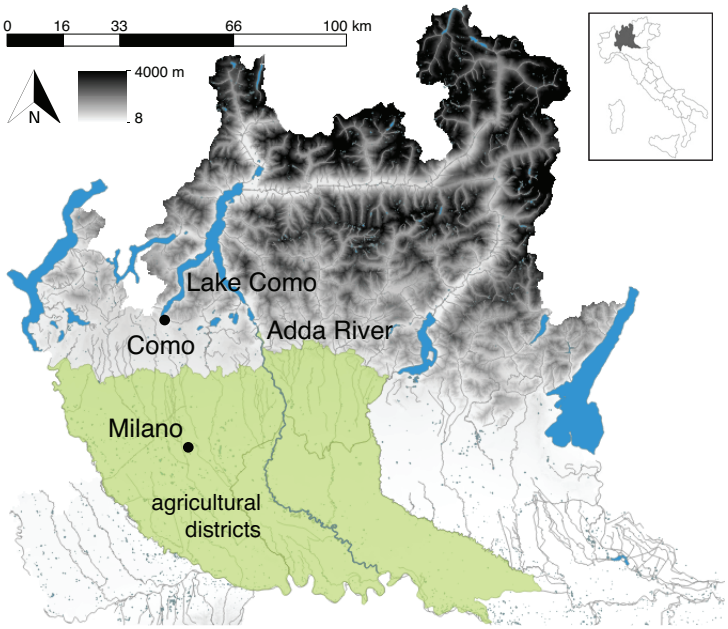

Figure 5.2: *Map of the Lake Como basin.*

The Lake Como system involves numerous economic activities and has been actively studied since the 1980s (e.g., Guariso et al., 1984, 1986). The existing regulation of the lake is driven by two primary, competing objectives: water supply, mainly for irrigation, and flood control in the city of Como, which sits at the lowest point on the lake shoreline. The agricultural districts downstream prefer to store snowmelt in the lake to satisfy the peak summer water demands, when the natural inflow is insufficient to meet irrigation requirements. Yet, storing such water increases the lake level and, consequently, the flood risk, which could instead be minimized by keeping the lake level as low as possible. On the basis of previous works (e.g., Castelletti et al., 2010a; Giuliani and

Castelletti, 2016; Giuliani et al., 2016a; Denaro et al., 2017b), these two objectives (both to be minimized) are formulated as follows:

- *Flood control* (J^F): the average annual number of flooding days in the simulation horizon.
- *Water supply deficit* (J^D): the daily average quadratic water deficit between lake releases and downstream water demands, subject to a minimum environmental flow constraint to ensure adequate environmental conditions in the Adda River. This quadratic formulation aims to penalize severe deficits in a single time step, while allowing for more frequent, small shortages (Hashimoto et al., 1982).

Further details about the Lake Como integrated model and policy design problem formulation are provided in Section S1 of the Supporting Information.

5.3 Methodology

Our Climate State Intelligence framework is composed of four steps as illustrated in Figure 5.1:

1. The detection of relevant climate teleconnections is performed by means of the Nino Index Phase Analysis (Zimmerman et al., 2016), which analyzes independently multiple climate signals, categorizes the available years looking at the climate oscillation phase and, for each phase, identifies preseason SST anomalies statistically significantly correlated with local conditions.
2. The seasonal precipitation forecasts are generated by capturing multiple climate signals using a nonlinear, multivariate Extreme Learning Machine model (Huang et al., 2006b), which is conditioned on the preseason SST for the relevant teleconnections detected in the previous step.
3. The hydrologic forecasts are produced through a temporal downscaling procedure of the seasonal precipitation forecast to feed a hydrologic model.
4. The forecast operational value is assessed by using the Information Selection and Assessment framework proposed by Giuliani et al. (2015).

Details about each step of the framework are reported in the next subsections.

5.3.1 Detection of climate teleconnections

The first step of our framework relies on the Nino Index Phase Analysis, a statistical framework originally developed by Zimmerman et al. (2016) for predicting seasonal precipitation conditioned on prior season atmospheric-oceanic

variables. In contrast to other studies that use the state of ENSO as a predictor in their models, this technique expresses ENSO as a physical influence on the "mean state" of the ocean-atmosphere system, in order to uncover possible asymmetric relationships (e.g., if in a given river basin El Niño is associated with dry conditions, La Niña not necessarily associates with wet conditions) that may be informative in a statistical prediction framework but are often overlooked. The NIPA methodology groups the available years into different phases (e.g., positive and negative if two phases are selected) according to the state of a climate signal (e.g., ENSO or NAO) as measured by its corresponding index. Subsequently phase-specific SST fields are identified and used as predictors in a seasonal forecast model. Each phase is then evaluated individually, thus constructing as many predictive models as the number of phases.

The first step of the modelling procedure consists of identifying the most significant SST predictor regions for each phase of the climate signal. For this purpose, correlation maps between the seasonal mean of local precipitation and preseason SST anomaly patterns are used, and correlated regions at the 95% significance level are identified for each phase. A Monte Carlo test is performed to ensure that SST grids are not randomly selected (i.e., SST grids must be present at the 95% significance level 90% of the time to be included). After identifying the SST predictor regions, a Principal Component Analysis (PCA, (see Joliffe, 2002)) is conducted on the entire predictor field and the first m resulting Principal Components (PCs) are retained as predictors in the forecast model. As in Zimmerman et al. (2016), we considered only the first Principal Component of preseason SST for each phase ($PC^1_{\tau-1}$), as it generally explains most of the variance in the selected SST predictor regions. The linear forecast model is defined as follows:

$$\hat{y}_\tau = \beta * PC^1_{\tau-1} + \alpha \tag{5.1}$$

where $\hat{y}_\tau$ is the predicted seasonal local precipitation, β the regression coefficient, and α the intercept. Given the low year-to-year persistence in precipitation data, a leave-one-out cross-validation procedure is then applied to the linear model to avoid model overfitting and predictive skill is measured with the Pearson correlation coefficient. Additional details on the NIPA framework are provided in Section S2 of the Supporting Information.

5.3.2 Seasonal precipitation forecasts

The second step of our framework constructs an empirical multivariate forecast model from the NIPA results. The teleconnections detected in the previous step are used for grouping the years into distinct climate states defined as combinations of the original phases of each climate signal. For example, the positive and negative phases of ENSO and NAO would combine to four climate states (i.e.,

positive-positive, positive-negative, negative-positive, and negative-negative). Including a neutral phase is also possible, although this results in nine climate states, and may create challenges due to the limited length of available data.

The multivariate forecast model is defined as a function of the first PC extracted from the preseason SST of each climate signal, along with a categorical flag indicating the climate state. We tested different data-driven model classes, including linear models, Artificial Neural Networks (ANNs), and Extreme Learning Machines (ELMs). These latter is a model belonging to the family of ANNs that by-pass the time-consuming gradient-based search of optimal ANN parameters by defining randomly parameterized hidden nodes, and subsequently optimizing only their output weights through a one-step matrix product (Huang et al., 2006b). Numerical results suggest that ELMs outperform the other models (for details, see Table S1 in the Supporting Information). The multivariate ELM forecast model is hence formulated as:

$$\hat{y}_\tau = \mathrm{ELM}(PC_{\tau-1}^{1,S1}, PC_{\tau-1}^{1,S2}, \gamma) = \sum_{i=1}^{N} w_i \psi_i(PC_{\tau-1}^{1,S1}, PC_{\tau-1}^{1,S2}, \gamma, \xi_i) \quad (5.2)$$

where $\hat{y}_\tau$ is the predicted seasonal local precipitation, $PC_{\tau-1}^{1,S1}$ and $PC_{\tau-1}^{1,S2}$ are the first PCs of preseason SST for two climate signals (e.g., ENSO and NAO), γ is a categorical climate-state flag, N is the number of nonlinear nodes in the hidden layers ψ_i, each using a sigmoidal activation function characterized by randomly generated (and thus not calibrated) parameters ξ_i, and w_i are the outputs weights to be calibrated.

5.3.3 Hydrologic forecasts

The third step of our framework generates hybrid hydrologic forecasts by transforming the seasonal precipitation forecasts produced by the ELM model into streamflow forecasts using a hydrologic model. This step requires temporal disaggregation of the predicted seasonal precipitation to an appropriate resolution for running the hydrologic model. Similarly to Souza Filho and Lall (2003), temporal disaggregation is performed by means of the k-Nearest Neighbor resampling method (Nowak et al., 2010). This data-driven method captures the observed variability, is consistent with the lag correlation structures in the observed data, and ensures mass conservation and continuity at the daily time scale (Rajagopalan et al., 1997). The temporal disaggregation relies on the computation of a seasonal proportion matrix P_t^τ, which distributes the seasonal precipitation volume over all days t within season τ. Then, k nearest neighbors of the seasonal precipitation are identified from the historical record. Finally, one of the neighbors is randomly selected according to a probability distribution that is proportional to the similarity of the historical season with the predicted

one.

The disaggregated precipitation forecasts are used as inputs to a hydrologic model. Specifically, we use the well known conceptual hydrologic model Hydrologiska Byrans Vattenbalansavdelning (HBV, see, Lindström et al., 1997), originally developed for operational flood forecasting in Sweden, which relies on four storage units, one for snow and the other three for different soil layers.

5.3.4 Assessment of forecast operational value

The last step of our framework assesses the operational value of the hydrologic forecasts, defined as the difference in system performance between an operating policy that uses the forecast information I_t (e.g., the hydrologic forecasts), and a baseline operating policy relying on more traditional information, such as the day of the year d_t and the lake level h_t, which in the case of Lake Como is sufficient to reproduce 85% of the variance of the observed release time series (Denaro et al., 2017b). According to the Information Selection and Assessment framework (Giuliani et al., 2015), this analysis also includes an upper-bound solution, designed assuming perfect foresight of future inflows, which allows estimating the potential maximum improvement of baseline operations. Additional details on the ISA framework are provided in Section S4 of the Supporting Information.

The optimal operating policies are computed by solving a multi-objective optimal control problem (Castelletti et al., 2008b) formulated as follows:

$$p^* = \arg\min_{p} \mathbf{J} = |J^F, J^D| \tag{5.3}$$

where the closed loop control policy p determines the release decision $u_t = p(d_t, h_t, I_t)$ at each time step t over the simulation horizon and the operating objectives are described in Section 5.2. Note that the problem in eq. (5.3) does not yield a unique optimal solution but a set of Pareto optimal solutions $\mathcal{P}^*$. The image in the objective space of the Pareto-optimal solutions is the Pareto front $\mathcal{F}$. To allow the direct use of hydrological forecasts as policy input, the problem in eq. (5.3) is solved by using the Evolutionary Multi-Objective Direct Policy Search method (Giuliani et al., 2016b), which implements a data-driven control strategy by integrating direct policy search, nonlinear approximating networks, and multi-objective evolutionary algorithms.

Given the Pareto optimal solutions of the problem in eq. (5.3), the operational value of the hydrologic forecast is quantified by the hypervolume indicator (HV), which captures both the convergence of the Pareto front conditioned on forecast information $\mathcal{F}$ to the ideal one assuming perfect foresight of future inflows $\mathcal{F}'$ as well as the representation of the full extent of tradeoffs in the objective space (Zitzler et al., 2003). The hypervolume measures the objective

space volume dominated ($\preceq$) by the considered set of solutions and HV is calculated as the hypervolume ratio between $\mathcal{F}$ and $\mathcal{F}'$, formally defined as follows (for a visual representation, see Figure S4a in the Supporting Information):

$$\begin{aligned} \mathrm{HV}(\mathcal{F}, \mathcal{F}') &= \frac{\int \alpha_{\mathcal{F}}(\mathbf{x})\,\mathrm{d}\mathbf{x}}{\int \alpha_{\mathcal{F}'}(\mathbf{x})\,\mathrm{d}\mathbf{x}} \quad \text{where} \\ \alpha_{\mathcal{F}}(\mathbf{x}) &= \begin{cases} 1 & \text{if } \exists \mathbf{x}' \in \mathcal{F} \text{ such that } \mathbf{x}' \preceq \mathbf{x} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{5.4}$$

This metric allows set-to-set evaluations, with HV assuming values between 0 to 1. Larger HV values imply superior Pareto optimal sets, with $\mathrm{HV} = 1$ assigned to the set of upper-bound solutions relying on perfect forecast information.

5.3.5 Data and experimental settings

The historical precipitation data used in this study are taken from Euro4M-APGD (i.e., Alpine Precipitation Gridded Dataset) developed by MeteoSwiss in the framework of the EURO4M (European Reanalysis and Observations for Monitoring deliverables) project. The data consists of a quality controlled, gridded dataset of daily precipitation from 1971 to 2008, which extends over the entire Alpine region and is based on measurements from high-resolution rain-gauge networks, encompassing more than 8,500 stations from Austria, Croatia, France, Germany, Italy, Slovenia, and Switzerland. For more information, see Isotta et al. (2014). Global Sea Surface Temperature anomalies were obtained from the NOAA's Extended Reconstructed SST (ERSST) Version 3b, a global monthly gridded dataset with a spatial resolution of 2.5 degrees.

We consider five candidate climate signals for improving the seasonal prediction of the precipitation in the Lake Como basin, including the El Niño Southern Oscillation, the North Atlantic Oscillation, the Pacific Decadal Oscillation, the Atlantic Multidecadal Oscillation, and the Indian Ocean Dipole. To distinguish the phases of different signals, we use the Multivariate ENSO Index (MEI) from NOAA, the Station-based Hurrell NAO Index from the National Center for Atmospheric Resource, along with the PDO Index, the AMO Index, and the Dipole Mode Index all from KNMI. The MEI is an index calculated based on the six main observed variables over the tropical Pacific: sea level pressure, zonal and meridional components of the surface wind, sea surface temperature, surface air temperature, and total cloudiness fraction of the sky (for details, see Wolter and Timlin, 1998). The NAO index is instead defined as the difference between the normalized mean winter (December-March) sea

level pressure anomalies at Lisbon (Portugal) and Stykkisholmur (Iceland) (Hurrell and Loon, 1997). The PDO Index is defined as the leading principal component of North Pacific monthly sea surface temperature variability (poleward of 20N for the 1900-93 period). The AMO Index is defined as the SST average over 0°-60°N, 0°-80°W minus the SST average over 60°S-60°N. The DMI Index is an indicator of the east-west temperature gradient across the tropical Indian Ocean, which is calculated as the difference of the Western Tropical Indian Ocean SST index and the South Eastern Tropical Indian Ocean SST index.

To solve the policy design problem via Evolutionary Multi-Objective Direct Policy Search, the policies are defined as Gaussian radial basis functions, which have been demonstrated to be effective in solving these types of multi-objective policy design problems (Giuliani et al., 2016b). The policy parameters are optimized using the self-adaptive Borg MOEA (Hadka and Reed, 2013), which has been shown to be highly robust in solving multi-objective optimal control problems (Zatarain et al., 2016). Each optimization was run for 2 million function evaluations over the simulation horizon 1996-2008, which was selected because it shows good variability in the local hydrological conditions including some intense droughts events. This time horizon is used for both the optimization and evaluation of the policy performance, which is hence likely overestimated as not tested on out-of-sample observations. However, the Lake Como inflows have already manifested a non-stationary trend over the last decades (see Giuliani et al., 2016d). The droughts of 2003-2005-2006, all used for the optimization of the policies as representative of extreme drought events, are therefore unprecedented and prevent using the 1971-1995 horizon for validating the policy performance as this latter is much wetter than the most recent years. To improve solution diversity and avoid dependence on randomness, the solution set from each formulation is the result of 20 random optimization trials. The final set of Pareto optimal policies for each experiment is defined as the set of non-dominated solutions from the results of all the optimization trials.

Finally, we provide a more tangible measure of the forecast operational value by converting the water supply deficit J^D into monetary values by using a spatially distributed agricultural model of the Muzza irrigation district, the largest district served solely by Lake Como releases. This model simulates soil-crop water balance (Facchi et al., 2004), crop growth stages as a function of the heat units accumulated (Neitsch et al., 2011), and final crop yield accounting for the effects of stresses due to insufficient water supply that may have occurred during the agricultural season (Steduto et al., 2009). Further details about the different model components are provided in Section S1 of the Supplementary Material.

Source code is available on Github: NIPA (https://github.com/mxgiuliani00/CSI), HBV model (https://github.com/mxgiuliani00/hbv), Lake Como simulation and

Table 5.1: *Pearson correlation coefficients between predicted and observed winter precipitation P_{JFM} with associated level of confidence for each phase of NAO and ENSO (predictions depend on preseason SST anomalies).*

Climate Signal	Phase	Pearson correlation	Level of confidence
NAO	negative	0.50	89.00%
	positive	0.52	69.70%
ENSO	negative	0.41	32.00%
	positive	0.51	96.20%

EMODPS implementation (https://github.com/mxgiuliani00/LakeComo).

5.4 Results

5.4.1 Detection of climate teleconnections

In the first step of the CSI framework (Figure 5.1), we run NIPA to detect the presence of potential teleconnections between multiple climate signals with the local precipitation in the Lake Como basin. We report here only results for ENSO and NAO related to wintertime precipitation (seasonal average precipitation in January, February, March, mainly as snow, which is highly correlated with spring inflows to the lake generated by the melting of the snow stored during the winter months on the mountains) which provide the best predictions of seasonal precipitation using the linear model in eq. (5.1). Results for other climate signals and other seasons are reported in the Supporting Information (see Table S2 and Table S3).

Wintertime predictions from preseason SST anomalies conditioned on the NAO negative phase and the ENSO positive phase independently result in high Pearson correlation values and high levels of confidence (Table 5.1). In particular, the prediction of winter precipitation obtained with the NAO negative phase has a correlation value of 0.50 with a level of confidence equal to 0.89; the positive phase of ENSO has a correlation value of 0.51 with a level of confidence equal to 0.96. These values of correlation are significantly higher than a standard regression model that is constructed without separating NAO and ENSO phases (Pearson correlation equal to 0.34), warranting the adoption of NIPA for the detection of climate teleconnections with asymmetric relationships with the local hydroclimatic processes.

Figure 5.3 illustrates the preseason SST grids selected by NIPA when exploring ENSO and NAO teleconnections (i.e., running a global search for preseason SST anomalies correlated with the seasonal precipitation in the Lake

Como basin splitting the data according to the positive and negative phase of each climate signal). The selected spatial extent of SST patterns is larger for the negative phase of NAO and the positive phase of ENSO than for the opposite phases, further indicating an asymmetric influence of the two climate signals during these phases. In particular, SST anomalies negatively correlated with winter precipitation during the negative phase of NAO are located in the western equatorial and North Pacific Ocean, in the eastern equatorial Atlantic Ocean, and in a small area in the central Indian Ocean. Positive correlations are evident in the western South Indian Ocean along the coasts of Madagascar. Conversely, for the positive phase of NAO, the selected SST are located in the central tropical Pacific Ocean, in a large area of the western Indian Ocean and along the African coasts of the Atlantic Ocean, with a small positively correlated region identified in the western tropical Pacific Ocean. Similarly, SST anomalies negatively correlated with winter precipitation in the positive ENSO phase are mainly located in the central tropical Pacific Ocean, along the coasts of central America, and the Atlantic Ocean along the coasts of central Africa, in the central North Atlantic Ocean, in the eastern Mediterranean Sea, and in large parts of the Indian Ocean. For the negative phase of ENSO, the correlation map illustrates an area of positively correlated SST anomalies in the southern Atlantic Ocean and a small region south of Madagascar. A numerical validation of these teleconnection patterns is reported in Section S7 of the Supporting Information using a complementary dataset.

Two interesting insights are evident from this part of the analysis (Table 5.1 and Figure 5.3): first, teleconnections from both NAO and ENSO are significant in the Lake Como basin but not equally active in the two phases considered; second, the SST grids characterized by high correlation with Lake Como basin precipitation are largely outside the areas of the Atlantic and Pacific Oceans traditionally considered for monitoring ENSO and NAO variability. These two aspects certainly contribute to the challenge of discovering teleconnection patterns in the Alpine region and the contrasting results reported in the literature, motivating the need for further process-based analysis to fully understand the physical drivers of these teleconnections.

5.4.2 Precipitation and streamflow forecasts

In the second step of the CSI framework (Figure 5.1), we generate seasonal precipitation forecasts using the Extreme Learning Machine forecast model described in Section 5.3.2 to capture the state of multiple climate signals and represent both their asymmetric relationships as well as their cross-influence with the local hydroclimatic processes (Huang et al., 1998; Mariotti et al., 2002; Matyasovszky, 2003). These model prediction depends on the couple of first Principal Components extracted from preseason SSTs for the concurrent ENSO and

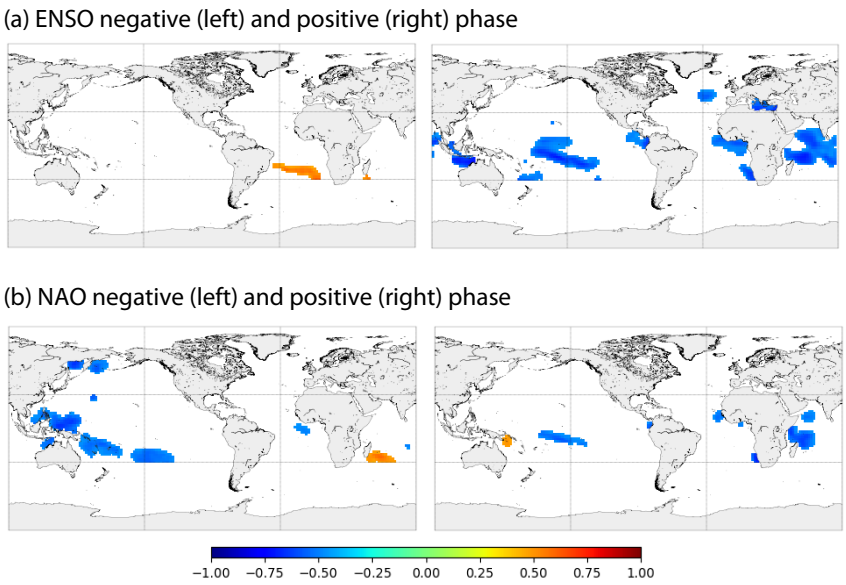

Figure 5.3: *Correlation maps between October, November, December SST anomalies and January, February, March precipitation in the Lake Como catchment for the two phases of ENSO (panel a) and NAO (panel b).*

NAO phases, where the average explained variance across signals and phases is 30.4% for the first PC and 19.7% for the second PC. The performance of all 3-month periods (e.g., JFM, FMA, MAM, etc.) predicted precipitation re-initialized at the beginning of each month is illustrated in Figure 5.4 and demonstrates high forecast accuracy: Pearson correlation coefficient between observed and predicted values is equal to 0.91 over the full dataset (i.e., calibration only), and to 0.81 in leave-one-out cross-validation (see Figure S6 in the Supporting Information).

The third step of the CSI framework (Figure 5.1) then transforms these seasonal precipitation forecasts into hydrologic forecasts. This step is performed by first disaggregating the seasonal precipitation forecast into a daily trajectory via kNN resampling to match the concentration time of the basin (~24 hours). Subsequently, the HBV model is run with the disaggregated precipitation forecasts and climatological temperature based on 1990-2003 observations. These simulations produce hydrologic forecasts of Lake Como inflows with a daily temporal resolution for a lead-time of 3 months, re-initialized at the beginning of each month as for the precipitation forecasts. The resulting hydrologic forecasts are cumulated over a lead-time of 51 days, a timeframe demonstrated by Denaro et al. (2017b) to be valuable for improving Lake Como operations.

The performance of the resulting streamflow forecasts illustrated in Fig-

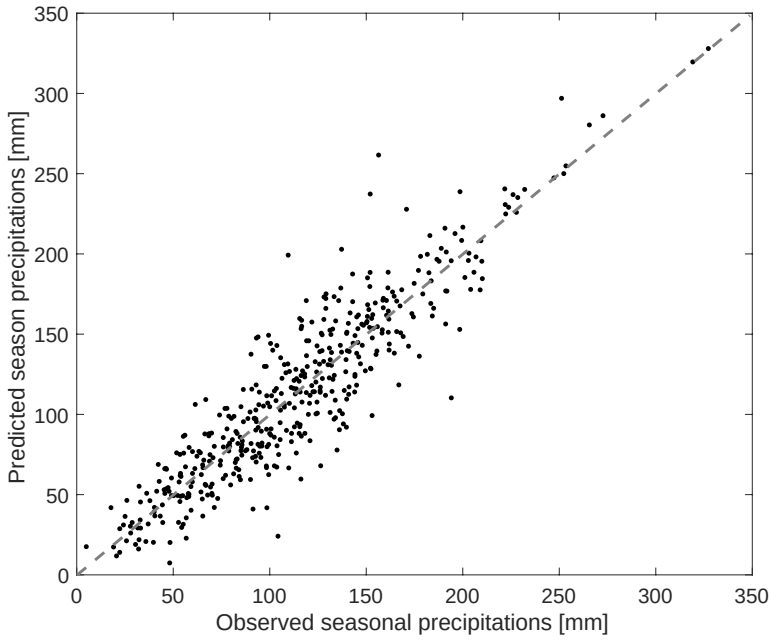

Figure 5.4: *Scatterplot between observed and predicted seasonal precipitation over the full dataset.*

ure 5.5 is inferior to the precipitation forecasts (Pearson correlation coefficient equal to 0.71 over the full dataset, results in cross-validation are reported in the Supplementary Information), which is likely attributable to the fact that HBV is run with climatological temperature rather than predicted temperature. Although temperature is less important than precipitation for generating seasonal hydrological forecasts, it however impacts on the daily streamflow dynamics simulated by the HBV model to match the 24 hours concentration time of the Lake Como catchment.

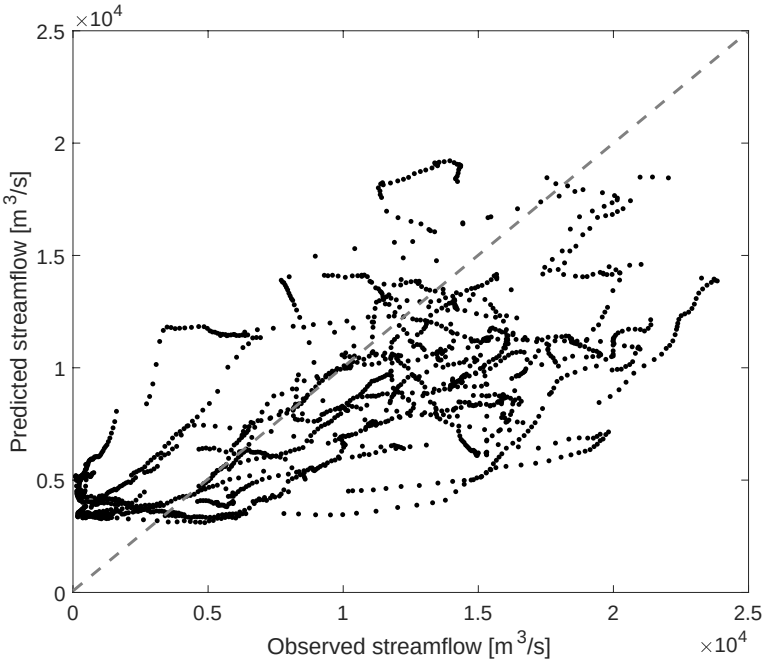

Figure 5.5: *Scatterplot between observed and predicted Lake Como daily inflows cumulated over a lead-time of 51 days over the full dataset.*

5.4.3 Hydrologic forecast operational value

In the final step of the CSI framework (Figure 5.1), the hydrologic forecasts presented in the previous section are used to inform Lake Como operations. The resulting informed operating policies' performance is illustrated in Figure 5.6 by the red circles, and is compared against an upper-bound solution designed assuming a perfect forecast (black circles) and a baseline solution corresponding to a traditional lake regulation conditioned on the day of the year and the lake level (gray circles). The figure also shows the performance of solutions in-

formed by the precipitation forecast (blue circles) and observed preseason SST (green circles) which will be discussed in the next section. The two axes of the figure represent the two operating objectives (to be minimized) and the arrows indicate the direction of increasing preference, with the best solution located in the bottom-left corner of the figure.

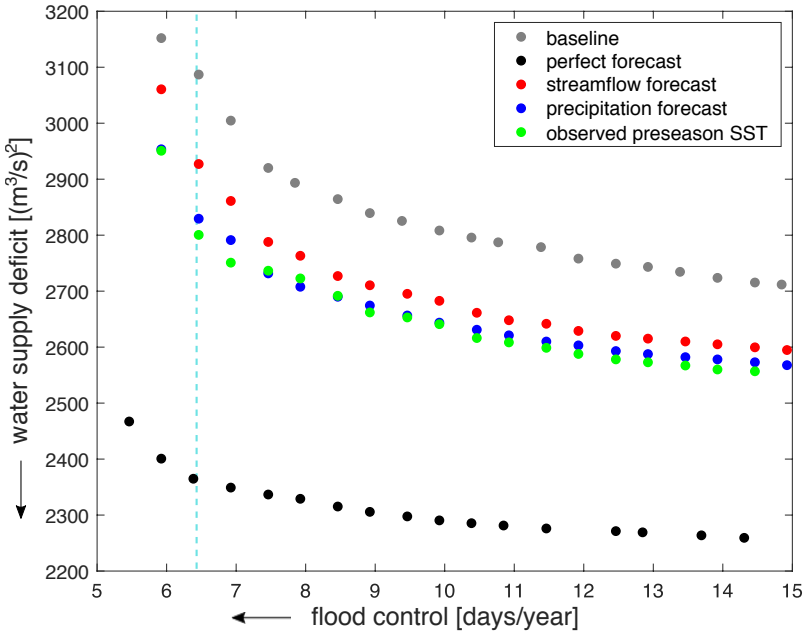

Figure 5.6: *Performance obtained by different Lake Como operating policies informed by streamflow forecasts (red circles), precipitation forecasts (blue circles), or observed preseason SST (green circles). The performance of these solutions is contrasted with the baseline operating policies (gray circles) and with policies informed by perfect forecast (black circles). The arrows indicate the direction of increasing preference for the two objectives and the cyan dashed line marks the performance of the historical lake regulation in terms of flood control.*

From the observation of the red, gray, and black Pareto fronts, a clear ranking can be made as the three sets of solutions not surprisingly do not intersect. The use of perfect forecasts outperforms other solutions as the (ideal) perfect knowledge of future inflows enables perfect decisions at each time step. The policies using the CSI forecasts, although inferior to the perfect-forecast solutions, are clearly superior to the baseline policies. This gain in performance is attributable to the information provided by the 51-day lead hydrologic forecasts, which result in a significant reduction in irrigation deficit. Conversely, this lead time doesn't positively contribute to improved flood control. Flood

Table 5.2: *Operational value of the hydrologic forecast, meteorological forecasts, and observed preseason SST in terms of hypervolume indicator* (HV).

Policies	HV	ΔHV	rel. ΔHV
baseline	0.32	-	-
streamflow forecast	0.46	0.14	44%
precipitation forecast	0.50	0.18	56%
observed preseason SST	0.51	0.19	59%
perfect forecast	1.00	0.68	212%

dynamics are on the order of hours to days, requiring much shorter lead times, whereas seasonal irrigation supply is more likely to benefit from seasonal forecasts. However, it is worth mentioning that the downward shift of the red Pareto front generated by the use of the CSI forecasts indirectly influence the performance in flood control as this new set of operating policies allows identifying better compromise altenatives. For example, if looking at the baseline solutions we assume that an acceptable value of J^D might be 2800 $(m^3/s)^2$, we can see that the informed operating policies attain similar levels of deficit by improving the performance on J^F from 12 to 6-8 flood days per year. The quantitative assessment of the operational value of the hydrologic forecasts is provided by the hypervolume indicator (HV, see eq. 5.4) reported in Table 5.2. The use of streamflow forecasts increases the baseline HV from 0.32 to 0.46, corresponding to a 44% gain in system performance, which is equivalent to about 20% of the maximum improvement attainable using a perfect forecast.

5.4.4 Operational value of precipitation forecast and observed preseason SST

In addition to evaluating the performance of solutions informed by the streamflow forecasts, we assess the performance of operating policies conditioned on seasonal precipitation forecasts (step 2 of the CSI framework) and observed preseason SST (step 1 of the CSI framework), illustrated by blue and green circles in Figure 5.6, respectively. Interestingly, the direct use of seasonal precipitation forecasts - skipping step 3 of CSI - results in a larger reduction of the irrigation supply deficit compared to the streamflow forecast-based solutions. Moreover, only using Principal Components of observed preseason SST values - skipping both steps 2 and 3 of CSI - further improves the policy performance.

Numerically, the use of streamflow forecasts improves over the baseline solutions by 44%; this gain increases to 56% and 59% when precipitation forecasts or observed preseason SST are used, respectively. These results suggest that

directly informing operational decisions with preseason SST observations selected depending on the climate state (i.e., combination of climate indexes and associated phases) as the main source of local predictability may produce equivalent or superior benefits than those from operational decisions relying on the full CSI modeling chain.

5.4.5 Analysis of the operating policies

To better understand the contribution of streamflow forecasts, precipitation forecasts, and observed preseason SST in improving Lake Como operations, we analyze the dynamic behavior of the system under operating policies that use distinct information. This analysis focuses on the solutions located along the cyan dashed line in Figure 5.6, which marks the performance of the historical lake regulation in terms of flood control. The rationale of this choice is to look at solutions that reduce the water supply deficit J^D without degrading the performance in J^F, as the historical regulation itself cannot be used as a reference since it also includes additional objectives not accounted for in our model (e.g., hydropower, navigation, fishing, tourism).

All simulated trajectories of the Lake Como level under each considered policy show a clear annual pattern, with the highest levels observed in late spring due to the snowmelt contribution (Figure 5.7a). In this period, maximizing the storage while avoiding floods is crucial for supporting the summer drawdown cycle driven by high irrigation demands. The policy conditioned on perfect forecast (black line) is able to maintain the highest level and delay the drawdown. The baseline solution (gray line), which has no information about future inflows, reaches the highest level at the beginning of May but subsequently the level is maintained about 10 cm below the perfect forecast trajectory to have space for buffering potential floods. A similar trajectory is followed by the policy informed by CSI streamflow forecasts (red line), although it is able to further delay the drawdown and to maintain a higher level than the baseline at the end of the agricultural season. Finally, the trajectories obtained under the policies informed by the precipitation forecast (blue line) and observed preseason SST (green line) are able to nearly match the perfect forecast trajectory until the middle of July, further reducing the water supply deficit.

This physical analysis can be translated into economic terms by estimating the agricultural profit for the farmers in the Muzza irrigation district served by Lake Como releases under these operating policies (Figure 5.7b). Results show the same ranking of solutions obtained in the space of the operating objective (Figure 5.6) and in the hypervolume indicator (Table 5.2). All the forecast-based policies increase farmers profit compared with the baseline, which falls shy of about 900,000 €/year with respect to the perfect forecast-based solution. Informing the lake operations with hydrologic forecast produces a 1.0% increase

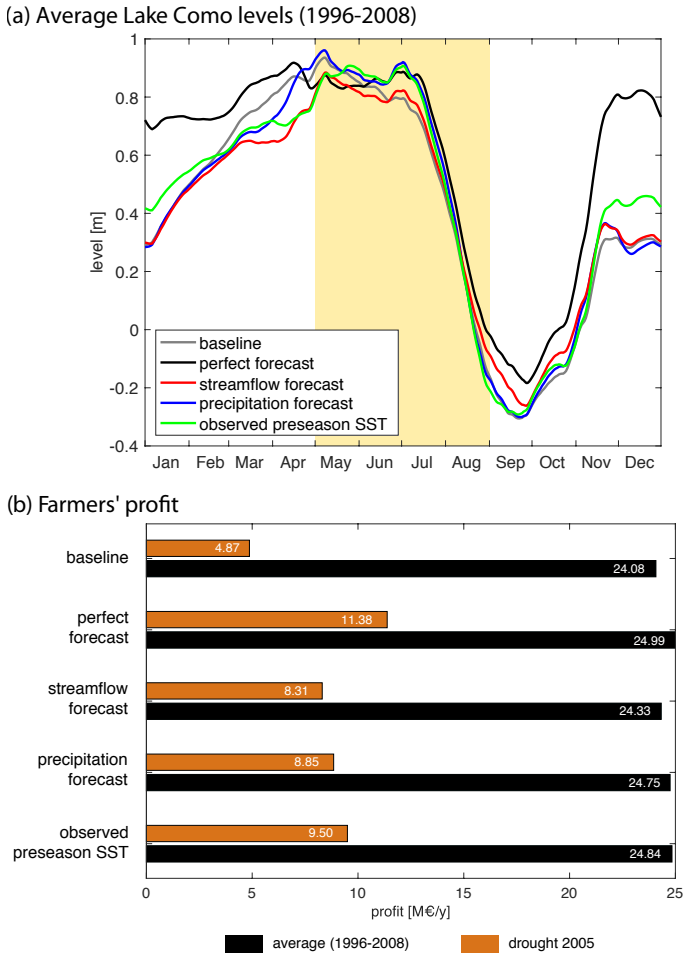

Figure 5.7: *Analysis of the average Lake Como levels (measured with respect to the Malgrate reference level at 197.37 m.a.s.l.) simulated under different selected operating policies (panel a) and corresponding profit of the farmers in the Muzza irrigation district (panel b). The yellow background in the top panel highlights the crop growing period.*

in average profits; this improves further when the precipitation forecast (+2.8%) or observed preseason SST (+3.2%) are used. Interestingly, these benefits are much larger when evaluated over extreme events, such as the drought recorded in 2005, when baseline annual profits dropped to 4.87 M€. This performance is about 20% of the 1996-2008 average baseline profit, and 6.5 M€ less than the expected profits in 2005 under the perfect forecast-based solution. In 2005, the value of the hydrologic forecast, precipitation forecast, and observed preseason SST also grows, producing a 71%, 82%, and 95% increase in farmer profits, respectively. These results suggest a large potential for directly using preseason SST anomalies in predicting and managing extreme droughts.

5.5 Conclusions

In this paper, we introduce the Climate State Intelligence framework to capture the concurrent state of multiple climate signals and produce seasonal forecasts for informing water reservoir operations. The framework is applied to Lake Como in the Italian lake district, a region where climate teleconnections are still not well understood or completely recognized.

Results indicate the potential of the CSI framework in extending the original Nino Index Phase Analysis to identify the combined ENSO–NAO climate state, which successfully uncovers notable teleconnection patterns within the Lake Como basin. These newly detected teleconnections are used to generate accurate seasonal precipitation forecasts using multivariate Extreme Learning Machines models, which subsequently feed a HBV model to produce skilful hydrologic forecasts of Lake Como inflows (Pearson correlation coefficients equal to 0.91 and 0.71 for precipitation and streamflow forecasts, respectively). Finally, the forecast operational value is quantified by estimating the difference in system performance between a baseline operating policy relying on traditional information (i.e., day of the year and lake level) with respect to an operating policy informed by the CSI forecasts. Numerical results show that streamflow forecasts produce a 44% improvement in the resulting Pareto optimal solutions. This gain transformed into farmers' profit is, on average, equal to about 250,000 €/year (1% of farmers' average profits), but increases to 3.44 M€ for the extreme drought recorded in 2005 (71% of farmers' profit in that year). Although we can expect this performance to degrade when simulated over out-of-sample data, we believe the policies informed by the CSI forecast will maintain their superiority with respect to the baseline solutions.

In addition, our results suggest that the CSI modeling chain provides intuitive information to improve water system operations but, at the same time, introduces modeling errors that negatively impact the final hydrologic forecast operational value. Overall, the direct use of observed preseason SST anomalies,

as identified in the step 1 of the CSI framework, provide more value than the precipitation or streamflow forecast in informing Lake Como operations, averaging a 3.2% gain in farmer profit and a 95% improvement for the 2005 drought as compared to the baseline solution.

These positive outcomes suggest a number of possible future research directions. Both the precipitation and streamflow forecasts produced by the CSI framework should be benchmarked against existing forecast systems, such as the ECMWF System5 (Johnson et al., 2019) and the European Flood Awareness System (Arnal et al., 2018). Additionally, the identified teleconnection patterns dependent on ENSO and NAO phases motivates a process-based investigation to better understand and clarify the underlying physical processes, perhaps by drawing on global climate model simulations. Extending the CSI framework to additional catchments in diverse hydroclimatic regimes with distinct management challenges is also of interest for better comparing the operational value of the multiple forecast approaches introduced in this paper. Finally, further improvements in system performance are likely achievable by properly combining local hydrologic information with global teleconnections, rather than considering these as alternative sources of information.

Exploring the utility for applying CSI to water reservoir operations under a changing climate is also warranted. In fact, the expected increase in frequency and intensity of extreme events, perhaps resembling the 2005 drought in the Lake Como basin, suggests that extending early actions and decision-making from short to medium and long lead-times by means of better forecasts may be increasingly valuable in the future (Turco et al., 2017).

Part III

Online Feature Representation for Multi-Objective Reinforcement Learning

In iterative *online* feature representation learning, policy search is interspersed with feature extraction routines to gradually refine the representation while learning the policy. The implementation of such a framework requires the selection of a feature extraction routine, a policy search method, and a strategy to interface the two. Direct Policy Search (DPS) is emerging as a promising policy search method for MO real-world control problems given its flexibility in problem and objectives formulation (Giuliani et al., 2016b). However, DPS traditionally defines the control policy within a prespecified rigid functional class that does not allow online changes in the policy input set.

In this Part, we first contribute NeuroEvolutionary Multi-Objective Direct Policy Search (NEMODPS), an original and highly flexible DPS technique that evolves the policy architecture and its parameterization simultaneously (Chapter 6). The findings of this work include that the implemented Pareto-dynamic architectural selection results in more robust policies when tested on unseen data. NEMODPS is firstly tested for a case study with a static input set, but the flexible architectural search can accommodate online changes in the feature representation dimensionality.

In the second work, NEMODPS is paired with a feature selection algorithm in a methodological contribution to iterative *online* feature representation learning (Chapter 7). The crucial novelty of the proposed framework, named Automatic Feature Selection-NEMODPS (AFS-NEMODPS), consists in supporting a tradeoff-dynamic representation learning that appears capable of significantly reducing conflicts between water users across different policy tradeoffs.

This Part contains the following works:

- Zaniolo, M., Giuliani, M., Castelletti, A., 2020b. Neuro-evolutionary direct policy search for multi-objective optimal control. IEEE transactions on neural networks and learning systems (under review);
- Zaniolo, M., Giuliani, M., Castelletti, A., 2020c. Dynamic retrieval of informative inputs for multi-sector reservoir policy design with diverse spatio-temporal objective scales. Environmental Modeling and Software (in preparation).

6

Neuro-Evolutionary Direct Policy Search for Multi-Objective Optimal Control

Abstract[1]

Direct Policy Search (DPS) is emerging as one of the most effective and widely applied Reinforcement Learning methods to design optimal control policies for Multi-Objective Markov Decision Processes (MOMDPs). Traditionally, DPS defines the control policy within a preselected functional class, and searches its optimal parameterization with respect to a given set of objectives. The functional class should be tailored to the problem at hand and its selection is crucial, as it determines the search space within which solutions can be found. In MOMDPs problems, a different objective tradeoff determines a different fitness landscape, requiring a tradeoff-dynamic functional class selection. Yet, in state-of-the-art applications, the policy class is generally selected a priori, and kept constant across the multidimensional objective space. In this work, we present a novel policy search routine called Neuro-Evolutionary Multi-Objective Direct Policy Search (NEMODPS), which extends the DPS problem formulation to conjunctively search the policy functional class and its parameterization in a hyperspace containing policy architectures and coefficients. NEMODPS begins with a population of minimally structured approximating networks and

[1] This work is currently under review as: Zaniolo, M., Giuliani, M., Castelletti, A., 2020b. Neuro-evolutionary direct policy search for multi-objective optimal control. IEEE transactions on neural networks and learning systems (under review).

progressively builds more sophisticated architectures by topological and parametrical mutation and crossover, and selection of the fittest individuals with respect to multiple objectives. We tested NEMODPS for the problem of designing the control policy of a multipurpose water system. Numerical results show that the tradeoff-dynamic structural and parametrical policy search of NEMODPS is consistent across multiple runs, and outperforms the solutions designed via traditional DPS with predefined policy topologies.

6.1 Introduction

The coexistence of multiple heterogeneous conflicting objectives is a major challenge to many complex real world control problems, which are often formalized as Multi-Objective Markov Decision Processes (MOMDPs). In these problems, the optimal solution is an ensemble of Pareto optimal policies covering the space of tradeoffs and compromises across different objectives. In the last decades, Multi-Objective Reinforcement Learning (MORL) established as solid approach to solve MOMDPs problems, but several open challenges remain in real world applications characterized by large continuous spaces that are too complex for a traditional optimal control formulation (for a review on MORL and open challenges see Liu et al. (2014) and references therein). Direct Policy Search (DPS) (Deisenroth et al., 2013) is emerging as one of the most popular MORL methods for solving complex MOMDPs problems, given its applicability to diverse tasks, scalability, and lack of restrictions in problem and objective formulation (Giuliani et al., 2016b). DPS defines the control policy within a given functional parameterization, and explores the policy parameters space by searching for the best solution with respect to a given set of objectives. So far, most of the DPS literature has focused on improving the search method (Heidrich-Meisner and Igel, 2008; Sigaud and Stulp, 2019), assuming that the subspace defined by the policy parameterization includes the optimal solution. This hypothesis, nevertheless, overlooks the impact that simplifications and mathematical assumptions in the problem formulation and the policy parameterization can have on the representation of the search space (Studley and Bull, 2007). Some DPS works apply a linear or piecewise linear policy parameterization, albeit conditioning the control decision on trivial monodimensional state vectors (Celeste and Billib, 2009; Abdolmaleki et al., 2016). A nonlinear multi-input multi-output function, such as an approximating network, provides a more flexible control policy shape (El-Fakdi et al., 2006; Rajeswaran et al., 2016; de Broissia and Sigaud, 2016). Yet, approximating networks require the specification of a topology, which is crucial to determine the network processing capability and training requirements. The a priori definition of the optimal network topology for a given problem requires a full knowledge of

the learning task that is generally unavailable, and, in practical applications, a topology is selected among few options via trials-and-errors, balancing the network approximation capacity, training costs, and overfitting tendency. Crucially, when multiple objectives are considered, the fitness landscape changes depending on the selected tradeoff, and the optimal network topology should be set accordingly. Yet, in state-of-the-art applications of DPS, a single policy class is selected to approximate solutions for every objectives tradeoffs.

This work contributes a novel policy search routine that addresses this challenge by evolving self-adaptive policy architectures responsive to changes in tradeoffs, namely, Neuro-Evolutionary Multi-Objective Direct Policy Search (NEMODPS). NEMODPS builds on a recent Reinforcement Learning branch called Neuro-Evolution (NE) (Stanley and Miikkulainen, 2003; Floreano et al., 2008), which employs Evolutionary Algorithms to generate optimal networks in terms of topologies and parameters. A well-known NE algorithm is the NeuroEvolution for Augmenting Topology (NEAT, Stanley and Miikkulainen, 2002), a Single-Objective (SO) technique which begins with a population of simple networks and progressively builds more sophisticated ones through a complexification process driven by parametrical and topological evolutionary operators. A topological niching scheme is featured to protect newly emerged architectures from premature disappearence. Several authors developed NEAT-inspired alternatives to adapt it to a variety of machine learning tasks, mainly for the fields of game playing and robotics (see e.g., Risi and Togelius, 2015, and references within). Among them, NEAT was tailored to problems characterized by highly discontinuous state-action mappings (RBF-NEAT, Kohl and Miikkulainen, 2008, SNAP-NEAT, Kohl and Miikkulainen, 2012, CA-NEAT, Nichele et al., 2018), little domain specific knowledge (Hausknecht et al., 2014), deceptive environments (Novelty Search, Risi et al., 2010), visual tasks (HyperNEAT, Gauci and Stanley, 2007, ES-Hyperneat, Risi et al., 2010), dynamic problems with moving optimum (DynNEAT, Krčah, 2012, SOMNE, Jiau and Huang, 2018), real-time adaptation of control policy (rtNEAT, Stanley et al., 2005, ICONE, Rempis and Pasemann, 2012), and compact policy representation (SUNA Vargas and Murata, 2017).

However, all the above algorithms address SO problems, and their application to a Multi-Objective (MO) problem requires its decomposition into several SO sub-tasks, each characterized by a scalarized monodimensional objective function. Sub-tasks are solved iteratively, each yielding one Pareto-approximate solution, causing a factorial growth of computational costs with the number of objectives, and suboptimality in the Pareto Front approximations in its convex regions (Vamplew et al., 2008). An attempt at developing an explicitly MO version of NEAT (MO-neuroevolution, Schrum and Miikkulainen, 2008) required to sacrifice several crucial NEAT operators, as they are supported by the

inherently single-objective niching scheme.

In this work we propose a MO generalization of the niching routine which allows to preserve all NEAT operators in a MO problem. We tested NEMODPS on a problem of designing a policy for a multipurpose water reservoir, typically featuring multiple conflicting objectives, a complex decision space, continuous domains, and a noisy input-output mapping. Currently, the state-of-the-art policy architectures for these problems are single-layer, fully connected Artificial Neural Networks (Zoppoli et al., 2002; Baglietto et al., 2010; Castelletti et al., 2013; Giuliani et al., 2016b).

NEMODPS implementation inherits NEAT basic structure for the dynamic search of efficient policy architectures, and the literature of NEAT refinements for problems presenting large decision spaces and noisy environments. In particular, inspiration came from the Evolutionary Acquisition of Neural Topologies (EANT) algorithm (Metzen et al., 2008) which addresses problems characterized by a large decision space, continuous domains, and a noisy environment by coordinating the search in a dual timescale, optimizing the network's connection weights on a small timescale (exploitation phase), and the network's structure on a larger timescale (exploration phase) in order to give newly created structures time to optimize their parameters. Other recent works dealing with noisy environments and complex decision spaces experimented with the activation functions of neurons. Applications to benchmark classification (Basirat and Roth, 2018) and regression problems (Hagg et al., 2017) demonstrate how heterogeneous networks characterized by a combination of activation functions can result in improved approximation capabilities, smaller networks with fewer training requirements, and a significantly reduced overfitting tendency when tested on noisy environments. Additionally, the niching routine is generalized for MO problems with a novel strategy, as to support the exploration of multidimensional tradeoffs in a single run of the algorithm.

In this application we run a benchmark analysis (Vamplew et al., 2011) comparing the policies produced by NEMODPS, NEAT, and traditional DPS, which demonstrates that the Pareto-dynamic structural and parametrical policy search of NEMODPS produces reliable policies, highly robust when tested on unseen data. Additionally, we perform a Pareto-dynamic convergence analysis of NEMODPS, and we analyze how the efficient architectures change in response to a change in the objective tradeoff, according to several metrics of structural analysis.

6.2 Methods

6.2.1 Problem formulation

In this work we consider a discrete-time continuous MOMDP defined as a tuple $< \mathcal{X}, \mathcal{U}, \mathcal{T}, \mathbf{G} >$where $\mathcal{X} \subset \mathbb{R}^{n_x}$ is the continuous state space, $\mathcal{U} \subset \mathbb{R}^{n_u}$ is the continuous action space, $\mathcal{T}(x_{t+1}|x_t, u_t)$ is the probabilistic transition function defining the transition density between state x_t and x_{t+1} under action u_t, $\mathbf{G}(x_t, u_t, x_{t+1}) = \left[G^1, \ldots, G^M\right]$ is a M-dimensional reward (or cost) function that specifies the vector of instantaneous rewards (costs) $\mathbf{g}_t = [g_t^1, \ldots, g_t^M]$ for each objective when state x_{t+1} is reached from state x_t by taking action u_t. Action u_t is extracted from a control policy π, $u_t = \pi(x_t, u_t)$, associated with a vector of expected returns $\mathbf{J}(\mathbf{1}) = \left[J^1(\pi), \ldots, J^M(\pi)\right]$ defined over the control horizon $[0, H]$ as:

$$J^m(\pi) = \mathbb{E}\left\{\sum_{t=0}^{H} (\gamma^m)^t g^m(t+1) | x_0 \sim \mu\right\} \tag{6.1}$$

where $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_M] \in [0, 1]$ is the vector of discount factors relative to each objective, and μ is the initial state distribution.

The solution of the RL problem defined above is the policy π^* that yields the optimal value of objective $\mathbf{J}$ (here considered as a cost, to be minimized) in its M dimensions:

$$\begin{aligned} \pi^* &= \arg\min_{\pi} \mathbf{J}(\pi, \mu) \\ &= \arg\min_{\pi} \left[J^1(\pi, \mu), \ldots, J^M(\pi, \mu)\right] \end{aligned} \tag{6.2}$$

In general, conflicts occur between different operating objectives, and it is thus not possible to define a single optimal policy, representing the optimum with respect to the M dimensions of $\mathbf{J}$. The solution of a MO problem is in general constituted by a set of non-dominated (or Pareto optimal) solutions $\mathcal{P}^* = \{\pi^* | \nexists \pi \prec \pi^*\}$, which maps onto the Pareto front $\mathcal{F}^* = \{\mathbf{J}(\cdot) | \pi^* \in \mathcal{P}^*\}$.

Remark. *Policy π dominates policy π', denoted by $\pi \prec \pi'$, if:* $\forall m \in \{1, \ldots, M\}$, $J^m(\pi) \leqslant J^m(\pi') \wedge \exists m \in \{1, \ldots, M\}, J^m(\pi) < J^m(\pi')$.

Algorithms designed for a SO optimization require to reformulate the MO problem as a series of SO sub-tasks, by combining the M objectives with a scalarization function $\Gamma : \mathbb{R}^M \to \mathbb{R}$. Traditionally, a convex combination of the objectives is applied using weights $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_M] \in \Lambda^{M-1}$, where Λ^{M-1} is the unit $(M-1)$-dimensional simplex (so that $\sum_{i=1}^{M} \lambda_i = 1$ and

$\lambda_i \geqslant 0 \quad \forall i$). For a SO control routine, problem (6.2) is hence reformulated as:

$$\begin{aligned} \pi^* = &\arg\min_{\pi} \mathbf{J}(\pi, \mu) = \\ &\Gamma\left([J^1(\pi, \mu), \ldots, J^M(\pi, \mu)]\right) \end{aligned} \tag{6.3}$$

The computational cost required by the solution of Problem (6.3) grows combinatorially with the number of objectives M (Giuliani et al., 2014a), and is defined by the following permutation:

$$S = \sum_{i=1}^{M} \frac{M!}{i!(M-i)!} + M \tag{6.4}$$

where S is the number of sub-tasks to be solved, equal to the number of Pareto approximate points produced. The idea is to explore the Pareto front by computing the M extreme solutions, obtained by setting to zero all weights except one, and some compromise solutions by relaxing the extremes and assigning the same weight to few objectives (Giuliani et al., 2014a).

Traditionally, the solution to Problem (6.2) is obtained by searching for the optimal action-value function $\mathbf{Q}^*(x_t, u_t)$, defined as the optimal cumulated future cost associated with each pair (u_t, x_t), as follows:

$$\begin{aligned} \mathbf{Q}^*(x_t, u_t) = \int_{\mathcal{X}} &[\mathbf{G}(x_t, u_t, x_{t+1}) + \\ &\boldsymbol{\gamma} \min_{u_{t+1} \in \mathcal{U}} \mathbf{Q}^*(x_{t+1}, u_{t+1})] \mathcal{T}(dx_{t+1} | x_t, u_t) \end{aligned} \tag{6.5}$$

The exact complete estimation of the value function in its M dimensions is however possible only for a limited class of problems, while it quickly becomes computationally intractable for problems characterized by high dimensional action or state spaces (i.e., curse of dimensionality Bellman, 1957) and objective space (i.e., curse of multiple objectives Powell, 2007). Moreover, any variable considered into the problem formulation must be explicitly modeled in order to compute the value function (i.e., curse of modeling Tsitsiklis and Van Roy, 1996).

In general, an approximated method is used when one or more curses prevent reaching an exact solution. The approximation can regard the action-value space (see e.g., Castelletti et al., 2011b, 2012), or the policy space, where the search for the optimal control policy is restricted to a prespecified parametric class of functions (Bertsekas, 2019). According to this second approach, the performance of a given policy π is evaluated by collecting the sequence of states and controls produced by applying the given policy over the horizon $[0, H]$. This sequence defines a trajectory τ employed in the calculation of the objective $\mathbf{J}(\pi) = \mathbb{E}[\mathbf{G}(\tau)|\pi]$.

Direct Policy Search belongs to this class, and according to the taxonomy of Policy Search methods proposed in (Deisenroth et al., 2013) configures as a stochastic, model-based and episode-based method. In particular, DPS approaches policy design as a problem of optimal functional parameterization, defining the control policy π_θ within a given function class, and then searching the parameters' space Θ to find the optimal parameterization $\theta^* \in \Theta$ with respect to the M-dimensional set of objectives $\mathbf{J}$. Hence, Problem (3.3) is reformulated as:

$$\pi_\theta^* = \arg\min_{\pi_\theta} \mathbf{J}(\pi_\theta, \tau) \tag{6.6}$$

Selecting an appropriate functional class for π_θ is critical, as DPS routines can find, at most, the best parameterization within the predefined class. In the absence of pre-existing knowledge of a (near-)optimal policy shape, highly flexible function classes (e.g., nonlinear approximating networks) are preferred (Zoppoli et al., 2002; Baglietto et al., 2010; Gong et al., 2015; Rajeswaran et al., 2016; Zhu and Jin, 2019; Dutta et al., 2015), in order not to restrict the search to a subspace of policies that, likely, does not contain the optimal one. Yet, optimizing the parameters of approximating networks requires searching high dimensional spaces, that map to a noisy and multidimensional objective space. MO Evolutionary Algorithms (MOEAs) are generally selected for this task given their demonstrated ability to efficiently handle performance uncertainties (Coello et al., 2007; Reed et al., 2013; Busa-Fekete et al., 2014). In state-of-the-art applications of DPS, an appropriate network dimension is selected by trials-and-errors, adjusting the number of neurons in a single-layer, fully connected, homogeneous network (Zoppoli et al., 2002; Baglietto et al., 2010; Castelletti et al., 2013; Xu and Jagannathan, 2014; Wang et al., 2015; Giuliani et al., 2016b). This architecture choice is motivated by theoretical results, which demonstrated that single- or multi-layer feedforward neural networks with continuous, non-constant, activation functions, could approximate any continuous bounded function to a desired accuracy, given enough nodes (Hornik, 1989). The nominal capacity of a neural network to absorb information is thus just limited by the number of its processing units, where numerous units imply large flexibility and approximation capacity. The network topology does not influence the theoretical expressiveness of a network; however, several studies show that, in practical applications, it significantly affects its training requirements, and approximation capacity. Firstly, fully connected networks offer high flexibility, but tend to force spurious connections that have no physical meaning, facilitating the overfitting of noise in training data (Srivastava et al., 2014; Liu et al., 2017; Jiang et al., 2017). Secondly, the depth (i.e., number of layers) of a neural network affects its behavior in solving high complexity learning tasks. While a shallow (single-layer) network provides a direct input-output mapping described by the single hidden layer, the global mapping provided by a deep (multilayer) net-

work is the result of the composition of several layers, a valuable asset in problems presenting regularities in the input-output mapping (Gauci and Stanley, 2007; Bianchini and Scarselli, 2014; Chang, 2015). Thirdly, comparative studies have demonstrated that the choice of nodes' activation functions plays a key role in determining convergence time and network accuracy (Kamruzzaman and Aziz, 2002; Laudani et al., 2015; Efe, 2008). An appropriate mix of activation functions generally reduces the number of processing units required for a task, and, accordingly, its training requirements and overfitting tendency (Hagg et al., 2017). Overall, these results indicate that in real-world applications, the network's topology plays a significant role in determining its suitability for a given task, and it should not be dismissed in DPS applications. Moreover, in MO problems, the multidimensional landscape defined by solutions mapped into corresponding value of objectives (i.e., fitness landscape) changes depending on the tradeoff. Every possible tradeoff combination originates a different sub-problem, and an efficient network topology should be set accordingly and tradeoff-dynamically.

6.2.2 Extending the scope of DPS

In this work, we extend the DPS problem formulation to search optimal policies in terms of architectures and relative parameterization Pareto dynamically. Accordingly, Problem (6.6) is reformulated as:

$$\pi^*_{\zeta(\theta)} = \arg\min_{\pi_{\zeta(\theta)}} \mathbf{J}(\pi_{\zeta(\theta)}, \tau) \tag{6.7}$$

where $\pi_{\zeta(\theta)}$ explicits the search for policy hyperparameters ζ defining a policy architecture as well as regular policy parameters θ, whose number and nature depend on the hyperparameters as in $\zeta(\theta)$. The policy search problem is thus expanded to conjunctively search architectural and parametrical spaces, enhancing DPS potential for single- and especially multi-objective problems.

6.2.3 NEAT

Problem (6.7) can be solved with Neuroevolution, a machine learning branch which employs evolutionary algorithms to automatically generate efficient artificial neural networks. NEAT (NeuroEvolution for Augmenting Topology, Stanley and Miikkulainen, 2002) is the first prominent neuroevolution algorithm, and the benchmark for this field. It begins with a population of simple networks and progressively builds more complex topologies through a complexification process. In every generation of the evolutionary progress, the performance of each individual is evaluated with respect to a fitness function, and the fittest individuals survive onto the next generation. New derivative networks are created based upon the surviving networks by applying evolutionary

operators (i.e., topological and parametrical mutation and crossover), to drive the search for efficient topologies and connection weights.

As the evolution proceeds and individuals complexify, increasingly sophisticated behaviors emerge. However, the addition of new structural elements with unoptimized coefficients is at first detrimental for an individual, and the usefulness of a topological innovation may become apparent only when given enough iterations to optimize. NEAT implements a niching scheme with the dual aim of protecting topological innovations from premature disappearance, and sustaining solution diversity. Topological innovation is protected by allowing individual competition only within niches of similar topologies. The population is partitioned into niches (or species), by evaluating a metric of topological distance δ between couples of individuals X_i and X_j:

$$\delta(X_i, X_j) = \frac{c_1 E_{i,j}}{NTE} + \frac{c_2 D_{i,j}}{NTE} + c_3 W_{i,j} \tag{6.8}$$

where $E_{i,j}$ is the difference in number of connections between X_i and X_j, $D_{i,j}$ is the difference in number of nodes, $W_{i,j}$ is the difference in average connection weights, $c1$, $c2$, $c3 \in [0, 1]$ express the relative importance of each factor, and NTE is the maximum Number of Topological Elements in the networks. Individual X_i is assigned to species s if:

$$\delta(X_i, X_{j,s}) < \delta^* \tag{6.9}$$

where δ^* is a predefined speciation threshold, and $X_{j,s}$ is the reference individual for the species, extracted randomly from species s at each generation. A new species is created if (6.9) is not verified for any existing one. Species are initialized by speciating the initial population.

Species compete among each other for their ability to reproduce, so that a larger offspring is assigned to well performing niches. However, a fitness sharing mechanism is introduced to penalize populous species and prevent them from taking over the entire population, thereby sustaining topological diversity.

In particular, a species' fitness is computed as the average shared fitness of its components. The Shared Fitness of individual X_i belonging to species s ($SF_{X_{i,s}}$) is determined by normalizing its fitness f_{X_i} to the species' numerosity n_s with the following:

$$SF_{X_{i,s}} = \frac{f_{X_i}}{n_s} \tag{6.10}$$

The allotted number of individuals n'_s to species s in the next generation is determined by its average shared fitness normalized by the population average $\overline{SF}$ and multiplied by population dimension N.

$$n'_s = \frac{\frac{1}{n_s} \sum_{i=1}^{n_s} SF_{X_{i,s}}}{\overline{SF}} \tag{6.11}$$

6.2.4 NEMODPS

Algorithm 1 NEMODPS meta-algorithm

```
    -Initialize-

    pop = initializePopulation(N)
    specThresh = δ
 5: speciesNumerosity* = σ
    probAddNode, probAddConnect, nExplore
    -Evolution-
    for gen = 1 to maxGenerations do
      -Evaluate individual fitness-
10:   for n in pop do
        fit = evaluateFitness(n)
        n = updateIndividualFitness(n, fit)
      end for
      -Species competition-
15:   pop' = {}
      for n in pop do
        fitSharingScore = countNonDominating(n, pop)
        n = updateFitSharingScore(n, fitSharingScore)
      end for
20:   for s = 1 to numSpecies do
        avgFit = averageFitSharingScore(pop,s)
        allottedInd = computeAllotInd(avgFit,popSize)
        elite = chooseSurvivingInd(pop,s)
        referenceInd(s) = extractReferenceInd(elite)
25:     add(pop', elite)
```

The implementation of NEMODPS inherits NEAT's structure, and the literature of NEAT improvements targeting complex control design problems, vast decision spaces, and noisy environments. Additionally, we propose an original strategy to extend the search to MO problems. The meta-algorithm for NEMODPS is reported in Algorithm (1). Below, we discuss the additional elements that differ from the original NEAT implementation.

First, NEMODPS assimilates the search dual timescale proposed in EANT (Metzen et al., 2008). Parametrical mutations occur in every generation to exploit existing structures. Topological innovations are injected every $nExplore$ generations (see Alg. 1), performing the exploration of the architectural hyperspace on a larger time scale (lines 33-40) with probability $probAddNode$ and $probAddConnect$.

```
        -Reproduction-
        for i = 1 to allottedInd do
          if rand < probabCrossover then
            [p1, p2] = chooseParents(elite)
30:         n′ = crossover(p1, p2)
          else
            n′ = chooseParent(elite)
          end if
          n′ = paramMutation(n′)
35:
          if isnull(remainder(gen ÷ nExplore)) then
            if rand < probAddNode then
              n′ = addNode(n′)

40:           end if
            if rand < probAddConnect then
              n′ = addConnect(n′)

            end if
45:       end if
          add(n′, pop′)

        end for
      end for
50:   -Speciation-

      for n′ in pop′ do
        n′ = assignSpecies(referenceInd, specThresh)
      end for
55:   numSpecies = countSpecies(pop′)
      if numSpecies > speciesNumerosity* then
        specThresh ++
      else if numSpecies < speciesNumerosity* then
        specThresh −−
60:   end if
      pop = pop′
    end for
```

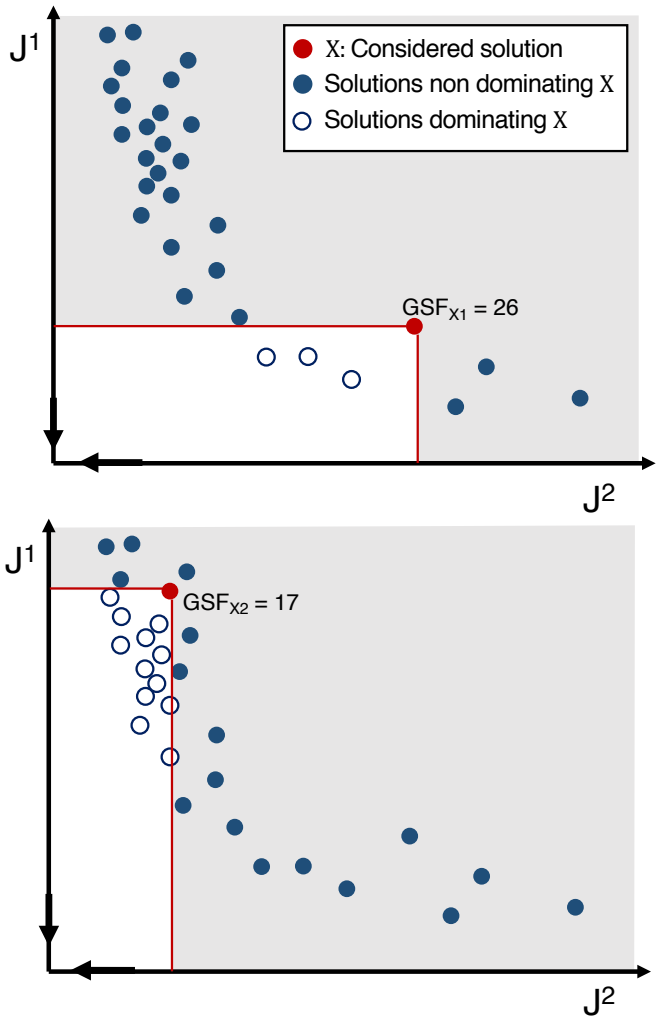

Figure 6.1: *Exemplification of the Generalized Shared Fitness computation for two individuals in a generic algorithmic iteration. In the top panel, the individual under evaluation is located in a relatively empty region of the objective space, and scores a value of GSF=26, equal to the non-dominating solutions (blue circles). The individual evaluated in the bottom panel is instead located in a crowded region of the objective space, scoring a lower GSF=17.*

Second, when new neurons are injected into a network (line 35), the activation function is randomly selected by the *addNode* operator among sigmoids and gaussians, allowing the generation of heterogeneous networks.

Third, the speciation strategy is modified to reduce the criticality of the speciation threshold choice. An excessively low speciation threshold produces too many species and an overly fragmented population with restricted interaction between individuals, and weakened selection pressure. On the contrary, an excessively high speciation threshold produces overly homogeneous populations, an unfavorable environment for new emerging topologies competing against well optimized structures. Moreover, the appropriate speciation threshold can vary significantly throughout the evolution as the population complexifies. Alternative to trying to guess a fair compromise for the selection threshold, some authors suggested to, instead, select an appropriate number of species to be maintained, and adjust the threshold accordingly during the evolution (Jang et al., 2009). NEMODPS implements the latter technique, initializing a threshold for speciation $\mathtt{specThresh} = \delta$, and an appropriate number of species to be maintained relatively constant during the search, $\mathtt{speciesNumerosity}^* = \sigma$ (lines 3-4). In every generation, $\mathtt{specThresh}$ is incremented if the number of species is above $\mathtt{speciesNumerosity}^*$, and lowered if inferior (lines 49-53).

Lastly, NEAT supports SO optimization, and its application to a MO problems requires the iterated solution of several SO tasks with a scalarized monodimensional objective as in Problem (6.3). In a previous attempt to define a multiobjective neuroevolution routine, named MO-neuroevolution, the Non-dominated Sorting Genetic Algorithm II (NSGAII, Schoenauer, 2000) was embedded in NEAT to perform the selection of the fittest individuals within niches in a multidimensional objective space (Schrum and Miikkulainen, 2008). The niching scheme supported by the Shared Fitness defined in eq. (6.10), however, does not seamlessly generalize to MO problems, given the difficulty to compare fitnesses with respect to multiple objectives, and therefore the MO-neuroevolution implementation sacrificed the speciation and fitness sharing operators. In NEMODPS, we employ NSGAII for intra-species competition (line 21), and we contribute an original definition of the fitness sharing operator for MO problems, thus restoring the speciation operator accordingly.

The Generalized Shared Fitness of individual X_i in species s, $GSF(X_{i,s})$, assigns a score to X_i equal to the number of individuals $X_j, j \neq i$ that are not dominating X_i.

$$GSF(X_{i,s}) = \sum_{j \in [1,\dots,N]: j \neq i} d_j; \quad d_j = \begin{cases} 0 & \text{if } X_j \prec X_i \\ 1 & \text{else} \end{cases} \tag{6.12}$$

where N is the total number of individuals in the population. The top score achievable is $GSF(X_i, s) = N - 1$, attained by individuals populating the best

current approximation of the Pareto front. Solutions close to the approximate Pareto front are assigned good scores if they are located in sparsely populated regions, and lower scores if they are located in crowded areas, as they are more likely to be (semi-)dominated. An example of GSF computation for 2 individuals in a 2-objectives problem is presented in Fig. 6.1. Species grow or shrink depending on whether the average generalized shared fitness of their individuals is above or below the population average (lines 19-20), in accordance with the NEAT implementation in eq. (6.11). Species competition is thus based on a relative individual ranking, a strategy that is often featured in MOEAs, and has been demonstrated to handle performance uncertainties more effectively than relying on the estimation of absolute performance (Busa-Fekete et al., 2014). Additionally, in this formulation, the fitness sharing operator penalizes individuals' proximity in the objectives space, rather than in the topological space as originally conceived in NEAT. This transition is encouraged by several authors, who have observed that topological diversity does not necessarily induce a behavioral diversity of solutions for every task (Moriguchi and Honiden, 2010; Lehman and Stanley, 2010). This observation is key in MO problems: if a certain sector of the Pareto front can be approximated with a trivial solution, a broad set of topologies will succeed in reaching a high performance. By rewarding topological diversity, solutions will quickly concentrate in the trivial region, resulting in a topologically diverse population, but a poor approximation of the Pareto front, which instead should be the ultimate goal of MO policy search. With the proposed generalized fitness sharing, species are encouraged to achieve solution diversity intended as a good exploration of the tradeoffs in the Pareto front, rewarding ensembles that are well performing, and that occupy relatively empty and non-dominated regions of the objective space.

6.2.5 Metrics of Structural Analysis

As argued in Section 6.2.2, the learning behavior of a network largely depends on its topology, therefore, topological analysis of Pareto-approximate networks could provide useful insights into the learning task. Neuro-optimized topologies are generally irregular, presenting sparse connections, hidden layers of different sizes, and heterogeneity in the activation functions. In order to characterize their topology, we use three metrics of structural analysis that capture critical network features, allowing us to compare and contrast different topologies.

The first metric, namely the Preference for Deep Learning (PDL), is measured as the ratio between number of hidden layers (L) and hidden nodes (H) in a structure.

$$\mathrm{PDL} = \frac{\mathrm{L}}{\mathrm{H}} \tag{6.13}$$

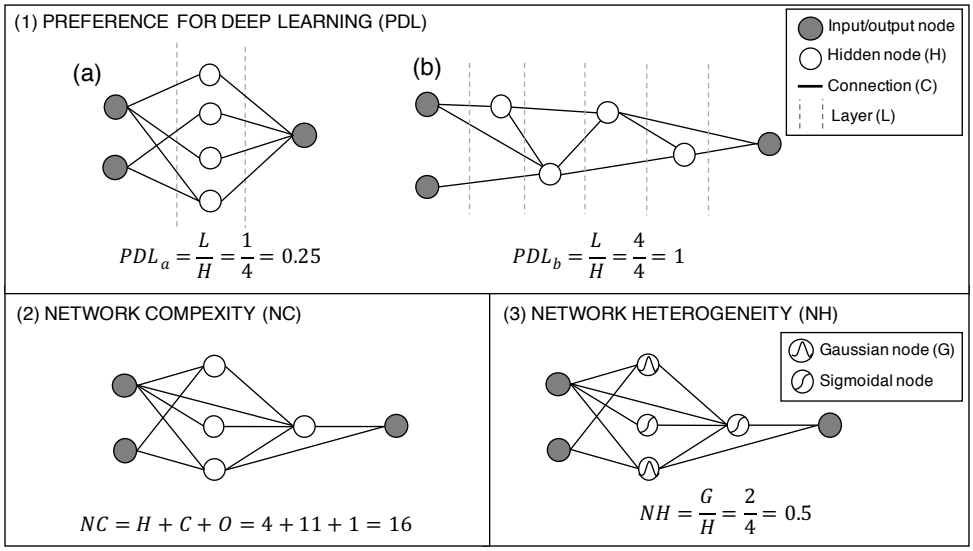

Figure 6.2: *Metrics of structural analysis. Panel 1 presents the metric Preference for Deep Learning with two examples representative of opposite configurations. Panels 2 and 3 report examples of computation of the Network Complexity, and Network Heterogeneity metrics, respectively.*

PDL $\in (0, 1]$, tends to zero when hidden nodes are organized in one or few very populated layers, and assumes value one when there are as many layers as nodes. The first panel of Fig. 6.2 represent two examples of networks with 4 hidden nodes but a different distribution into layers. Network (a) has one hidden layer, while network (b) is distributed into 4, resulting in different values of PDL.

The second metric is a measure of Network Complexity (NC), defined as the total number of parameters, namely connection weights and node biases, needed for its description. The sum of the number of connections (C), hidden nodes (H), and output nodes (O), determines the Network Complexity as follows:

$$\mathrm{NC} = \mathrm{C} + \mathrm{H} + \mathrm{O} \tag{6.14}$$

Fig. 6.2 reports one example of NC computation in Panel (2), where connections and nodes present different line width as a visual representation of varying connection weights and bias. High NC values are representative of complex networks, likely to reproduce sophisticated behaviors.

Lastly, the third considered metric of structural assessment is a measure of Network Heterogeneity, computed as the ratio of Gaussian nodes (GN) to the

total number of hidden nodes.

$$NH = \frac{GN}{H} \tag{6.15}$$

By definition, $NH \in [0, 1]$ where NH = 0 indicates a homogeneous network comprising only sigmoidal nodes, and NH = 1 indicates a homogeneous Gaussian network. An even mixing of Gaussian and sigmoidal activation functions is verified for NH = 0.5, as in the example of Panel (3) in Fig. 6.2.

6.3 Case Study

NEMODPS is tested for a problem of designing the optimal control of a multipurpose water resources system. Typical features of these problems are large decision spaces, presence of noise, and multiple conflicting objectives.

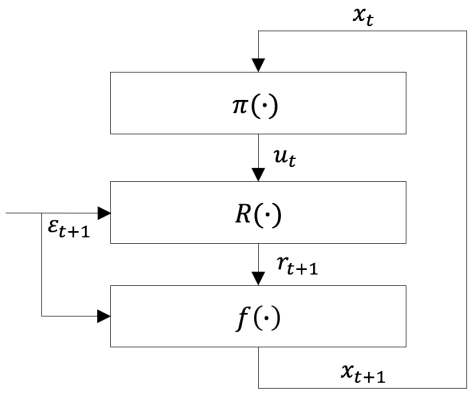

Figure 6.3: *Control scheme of the considered case study. A non-linear release function* $R(\cdot)$ *converts the policy output (i.e., release decision)* u_t *into a feasible release* r_{t+1}*, which determines the state transition according to function* $f(\cdot)$*. Stochastic disturbances (i.e., lake inflow)* ε_{t+1} *affect release and state transition of the system.*

In this application, we design the control policy of Lake Como, a multipurpose regulated lake situated in the southern Alpine belt (Italy). The main tributary, and only emissary of the lake is the Adda river, whose waters are withdrawn downstream of the lake to irrigate four agricultural districts. The southwestern branch of Lake Como constitutes a dead end, and exposes the city of Como to flooding events.

The control scheme relative to this problem is reported in Fig. 6.3. The system is modeled as a discrete-time, periodic, non-linear, stochastic process defined by a scalar state variable x_t (i.e., storage), a control variable u_t represent-

ing the release decision from the dam gates, stochastic disturbances ε_{t+1} (net reservoir inflow), and a state-transition function $f(\cdot)$: $x_{t+1} = x_t - r_{t+1} + \varepsilon_{t+1}$ where the effective release r_{t+1} coincides with the release decision u_t corrected, where appropriate, with a non-linear release function $R_t(x_t, \varepsilon_{t+1})$ determining the minimum and maximum releases feasible for the time interval $[t, t+1)$ to respect physical and legal constraints.

The Adda River is described by a plug-flow model, which simulates the routing of the lake releases from the lake outlet to the intake of the irrigation canals. The adopted time step is 1 day, and the system is periodic with period T = 365 days (Giuliani et al., 2016c).

The lake regulation has two conflicting aims of minimizing flood risk on the lake shores, and supplying water to downstream users by storing spring snowmelt-driven inflow peak and releasing throughout summer when the irrigation demand is highest. On the basis of previous works (Castelletti et al., 2010b), these two objectives are defined as:
Flooding: the average number of annual flood days, defined as days in which the lake level h_t is above the flood threshold $\bar{h} = 1.24$ m, i.e.:

$$J^{flood} = \frac{1}{N_y} \sum_{t=0}^{H-1} g_{t+1}^{flood}; \quad g_{t+1}^{flood} = \begin{cases} 1 & \text{if } h_{t+1} \geqslant \bar{h} \\ 0 & \text{if } h_{t+1} < \bar{h} \end{cases} \tag{6.16}$$

where N_y is the number of years in the simulation horizon. Irrigation: the daily average squared water deficit with respect to the daily downstream demand w_t, subject to the minimum flow constraint $q^{MEF} = 5\ m^3/s$ to guarantee environmental stakes. The quadratic formulation is selected with the aim of penalizing severe deficits in a single time step, while allowing for more frequent, small shortages. i.e.,

$$J^{irr} = \frac{1}{H} \sum_{t=0}^{H-1} (max(w_t - (r_{t+1} - q^{MEF}), 0))^2 \tag{6.17}$$

We hereby assume the considered simulation horizon $\mathcal{H}$ is sufficiently long to not require the addition of a penalty function to the final state. Furthermore, the release decision is not conditioned on time beyond its position within the annual cycle, and thus the decision at the end of the time horizon is no different than during the equivalent period of all previous years.

6.4 Computational Experiment

The problem of finding a set of Pareto approximate control policies for the Lake Como system was solved via three policy search methods, NEMODPS, NEAT,

and traditional DPS, respectively. In these experiments, the designed optimal control policies provide the control u_t as a function of a three-dimensional input set I_t comprising the state of the system (i.e., the current reservoir storage) and two transformations of the time index t with sine and cosine, to embed time-variability and cyclostationarity in the control policy $I_t = |x_t, \sin(2\pi * t/T), \cos(2\pi * t/T)|$.

NEMODPS solves Problem (6.7), specified for the case study as:

$$\pi^*_{\zeta(\theta)} = \arg\min_{\pi_{\zeta(\theta)}} [J^{flood}, J^{irr}] \tag{6.18}$$

where J^{flood} and J^{irr} are formulated as in eq. (6.16) and (6.17), respectively. NEMODPS was run for 10 independently initialized and randomized seeds. Each seed comprises a Number of Function Evaluations (NFE) equal to 600 thousands. Individuals of the initial population consist of one hidden, one output node, and 4 connections, for a total of 6 parameters. Connections link inputs to the hidden node, and the hidden node to the output. Evolved individuals feature different complexities, spanning from 10 to 31 parameters across the 10 runs.

NEAT solves a SO version of Problem (6.18) where the two objectives are aggregated using a weighted mean:

$$\pi^*_{\zeta(\theta)} = \arg\min_{\pi_{\zeta(\theta)}} [\lambda_1 J^{flood} + \lambda_2 J^{irr}] \tag{6.19}$$

Problem (6.19) is iteratively solved for 15 uniformly sampled combinations of $[\lambda_1, \lambda_2]$ in [0,1]. NEAT thus demanded the same computational effort of NEMODPS multiplied by the 15 tradeoff combinations considered.

Finally, the application of traditional DPS solves Problem (6.18) searching only the policy parameters $\theta \in \Theta$ for a pre-defined functional class:

$$\pi^*_{\theta} = \arg\min_{\pi_{\theta}} [J^{flood}, J^{irr}] \tag{6.20}$$

DPS requires the specification of a search algorithm, and of a policy structure. As search algorithm we selected the ϵ-NSGAII MOEA (Kollat and Reed, 2005), which demonstrated consistently high levels of performance on an extensive diagnostic benchmarking for challenging MO problems (Zatarain et al., 2016). ϵ-NSGAII extends the original NSGAII by including epsilon dominance archiving, adaptive population sizing, and time continuation (for a detailed description of the algorithm, please refer to Kollat and Reed, 2005, 2006). Concerning the policy structure, a single-layer, fully connected, homogeneous network was selected, as in state-of-the-art applications (Castelletti et al., 2013; Giuliani et al., 2016b). The experiment was repeated for differently sized networks, from 1 node (corresponding to 6 parameters), to 6 nodes (31 parameters), covering an

interval of parameters which contains the range delimited by optimized NEMODPS networks. These networks were populated homogeneously with sigmoidal activation functions, generating common Artificial Neural Networks (ANN), and with Gaussian functions, generating Gaussian Perceptrons (GP).

Because Problem (6.20) only searches the parameters' space, in contrast to Problems (6.18) and (6.19) which search the hyperspace comprising networks' parameters and topologies, the number of function evaluations had to be adjusted to ensure a fair comparison across methodologies. By inspecting the search progression in NEMODPS, it was determined that, on average, the structures populating NEMODPS Pareto fronts remained fairly constant for the last 300 thousands evaluations. As a result, each DPS experiment was run for 10 seeds, and for NFE = 300,000.

The above policy search experiments were performed on a 10 years calibration horizon 1997-2006 comprising a mix of wet and dry years. Optimal policies were then tested on three validation chunks: an extended 20-years validation from 1977-1996, and combinations of extreme dry years (1949, 1962, 1990, 1994, 2007), and wet years (1951, 1960, 1977, 2008, 2014) selected by searching the driest and wettest years from the available historical record of inflows to Lake Como (1947-2014), discarding the calibration years.

6.5 Numerical results

6.5.1 Benchmark analysis

The first experiment we present is a benchmark analysis, contrasting the performance of Pareto-approximate control policies produced via NEMODPS, NEAT, and state-of-the-art DPS. Figure 6.4 reports the solutions' performance for the calibration in panel (a) and validation periods in panels (b), (c), and (d), with respect to the two objectives of irrigation deficit (J^{irr}, vertical axis) and flood days (J^{flood}, horizontal axis), both to be minimized. Marker size is proportional to network dimension (i.e., number of parameters, or topological elements). The solutions displayed in this figure are the non-dominated solutions resulting from merging the Pareto front approximations of independent repetitions of the three policy search routines. State-of-the-art ANN and GP networks (pink and green diamonds, respectively) of every size (from 1 to 6 hidden nodes) obtain the best calibration results, outperforming NEMODPS (blue circles), and NEAT (black triangles). However, when tested on unseen validation datasets, their performance significantly deteriorates. Benchmark DPS architectures thus demonstrate a tendency to overfit noise patterns in training data, which enables the attainment impressive calibration results, but without effectively producing superior policies when compared to other policy search routines. On the contrary, NEMODPS control policies offer a much more stable

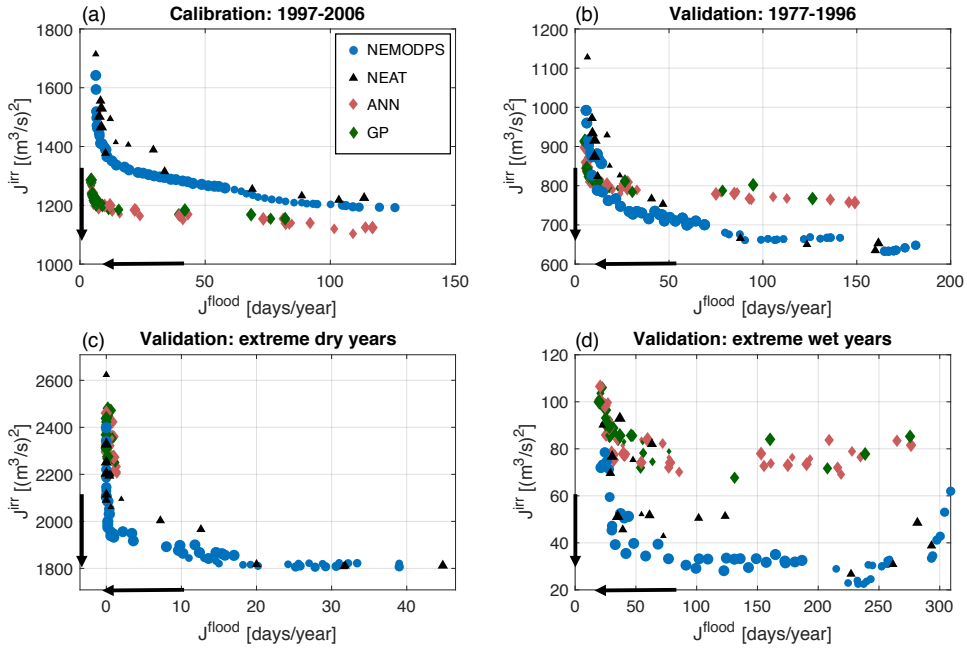

Figure 6.4: *Comparison of the control policies' performance designed via NEMODPS (blue circles), NEAT (black triangles), and traditional DPS with fixed structures ANN and GP networks (pink and green diamonds). Policies are evaluated over a 10 years calibration period (panel (a)), a 20 years validation horizon of recorded inflows (panel (b)), and two 5 years extreme validation horizons (extreme dry in panel (c), and extreme wet in panel (d)).*

validation/calibration ratio and consistently outperform benchmark DPS on all three validation datasets. Moreover, NEMODPS consistently offers an exhaustive exploration of the Pareto front, with very limited gaps even when tested on validation datasets. Conversely, solutions produced by fixed structure DPS tend to concentrate in restricted portions of the frontier, (e.g., panel (c)).

NEAT policies almost overlap with NEMODPS solutions in the extremes of the Pareto front; however, the central region of the front is poorly characterized, presenting large gaps, and dominated solutions. Remarkably, selecting evenly spaced set of weights to aggregate the two objectives does not guarantee a uniform distribution of NEAT solutions in the Pareto front. The non-uniformity in the distribution of solutions could be a consequence of concavities in the real unknown Pareto front, which are impossible to capture with a convex combination of objectives.

Intuitively, the higher reliability of neuro-evolved policies in contrast to traditional pre-defined structures against a suite of diverse validation experiments can be explained by the fact that each topological element of neuro-optimized networks was established as the result of a genetic selection. Consequently, the added value of every element is tangible, otherwise simpler networks, with lower calibration requirements, would have prevailed. On the contrary, by prespecifying a network structure, any superfluous element populating the network (e.g., connections with no physical sense) will contribute to overfitting to the noise patterns, ultimately undermining the network generalization capability.

6.5.2 Trends in policies architectural features

The following analysis is aimed at exploring the NEMODPS topology selection in more detail, by uncovering possible trends and regularities in the architectural features of the Pareto-approximate solutions produced by the 10 independent runs of NEMODPS. This analysis is supported by the three structural metrics defined in Sec. 6.2.5, computed for every solution, and plotted against their performance with respect to J^{flood} in Fig. 6.5. The flood objective is used as a proxy to represent the solution tradeoff, as, for a given seed, lower J^{flood} values correspond to higher J^{irr} values.

The first panel of Fig. 6.5 displays the Preference for Deep Learning (PDL), defined in eq. (6.13). Each line represents one of the 10 independent runs of NEMODPS. By inspecting the lines ensemble, a clear trend is visible: as J^{flood} increases, (corresponding to moving the tradeoff in favor of good J^{irr} performance) the values of PDL tend to increase as well, eventually reaching 1 in all the iterations.

The second panel of Fig. 6.5 shows the values of NC with respect to increasing values of J^{flood}. A visible trend persists in all 10 runs, indicating that

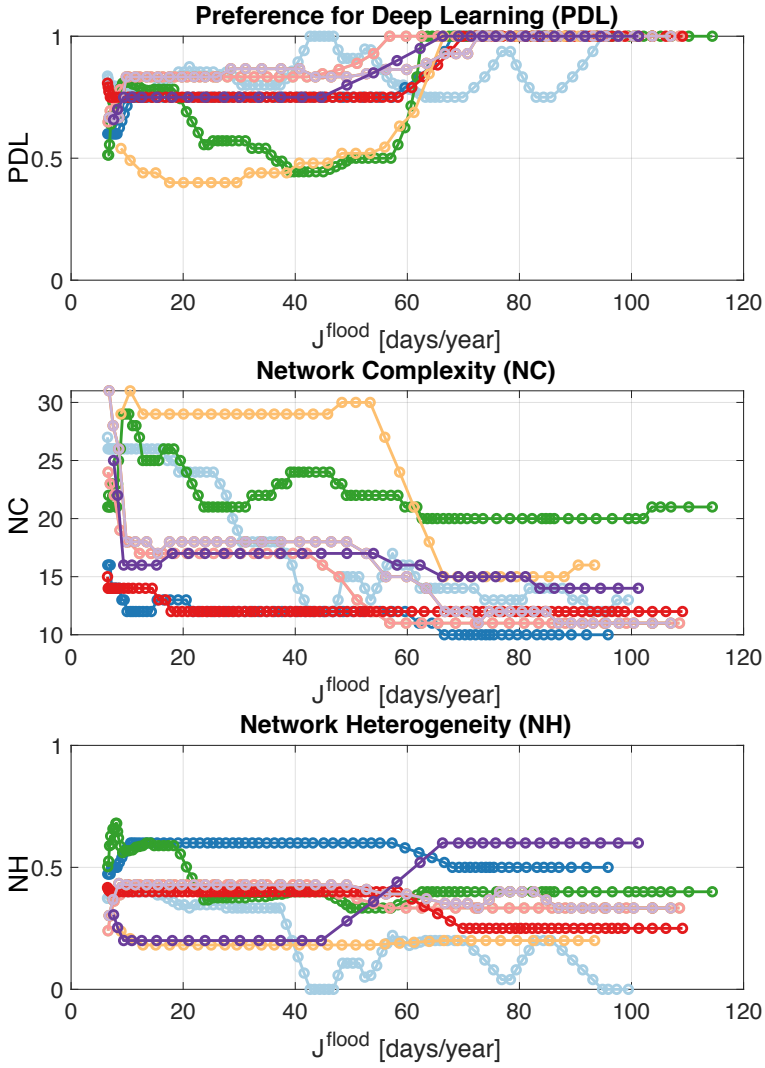

Figure 6.5: *Pareto dependent structural analysis of optimal solutions resulting from 10 independent runs of NEMODPS, represented by different line colors. The three metrics employed for structural analysis are Preference for Deep Learning (top panel), Network Complexity (middle panel), and Network Heterogeneity (bottom panel).*

efficient architectures tend to simplify, on average, for high values of J^{flood}. Also the range of complexities covered by the solutions is sensible to a change in tradeoff. Low flood solutions display high variability in NC across different seeds, spanning from 12 to 41 parameters below 20 flood days. On the other end of the tradeoff curve, instead, solutions are confined within the 10 to 15 parameters range except for one seed stabilizing on 20 parameters. The last indicator of Network Heterogeneity (eq. 6.15) does not present any visible trend in response to the change of the J^{flood} objective. However, except for very few cases, Pareto-approximate networks select heterogeneous configurations comprising a mix of sigmoidal and Gaussian functions. In this mix, generally, sigmoidal functions constitute the greater portion (verified for $NH < 0.5$).

In summary, different runs of NEMODPS evolve independently to reach a coherency in the architecture of Pareto-approximate networks, indicating rationality in the network generation. The optimization routine, moreover, responds to changes in tradeoff by consistently adapting the solution topology, confirming that multi-objectives problems should be approached with a Pareto-dynamic selection of optimal architectures.

6.5.3 Convergence analysis from a multi-objective perspective

The last experiment is aimed at verifying the convergence of the solutions produced via NEMODPS across its independent runs from a multi-objectives perspective. First and second panel of Fig. 6.6 represent the minimum value of the two objectives, J^{flood} and J^{irr}, respectively, throughout the search until the maximum generation is reached. These two objectives present a remarkably different behavior: the best value of $J^{flood} = 6.3$ is consistently found at an early stage of the search by every algorithmic iteration, represented by differently colored lines, indicating that policies that minimize flood days (irrespective of their J^{irr} value) are relatively trivial to obtain. Conversely, the quest for an optimal irrigation deficit performance appears much more complex, given the slower progression towards low values of J^{irr}. As opposed to J^{flood}, the best J^{irr} solution obtained at the end of the search differs for every iteration, however, the final solutions place within a range of 5.19% with respect to the lowest, indicating a contained inter-seed variation even in regions of the Pareto front that appear more difficult to approximate. Notably, a marked difference in computational effort required by different objectives poses an additional challenge to the use of SO policy search routines, as it complicates the selection of the set of weights employed in the objective aggregations. An example of this is visible in Fig. 6.4, where an evenly spaced set of aggregation weights for NEAT produces clusters of solutions scoring low values of J^{flood} and gaps in the Pareto front.

The bottom panel of Fig.6.6 reports the hypervolume indicator HV (Zitzler

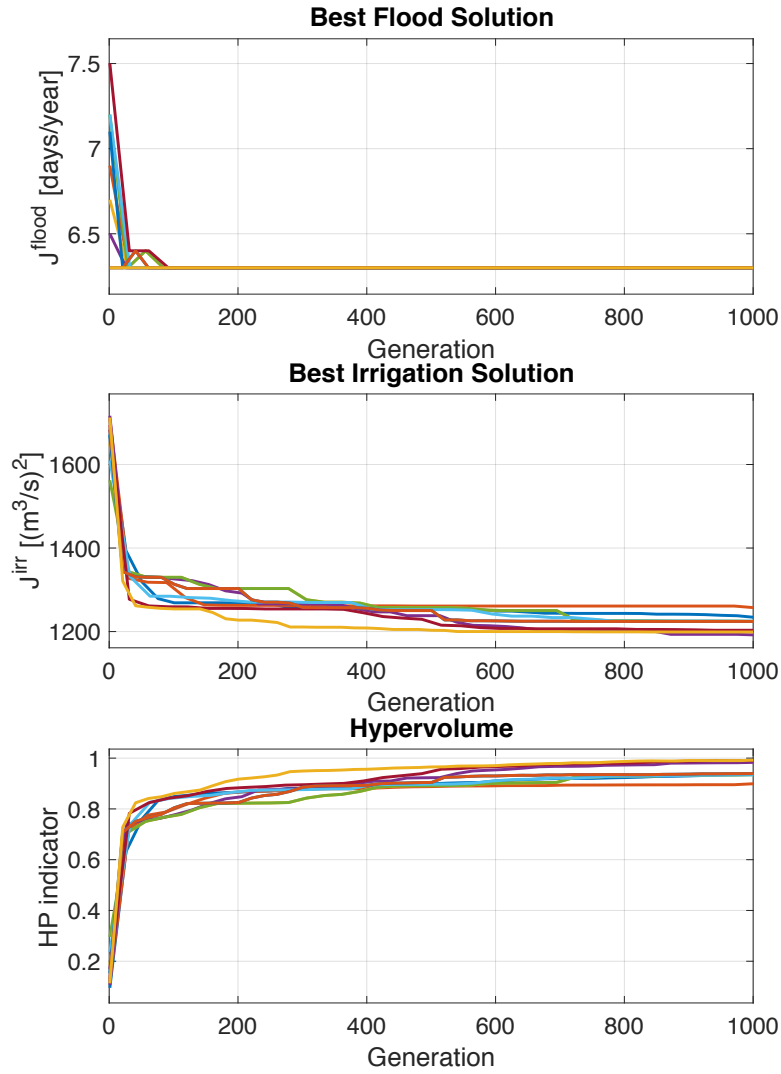

Figure 6.6: *Analysis of solution convergence with respect to multiple objectives. Each line represents the behavior of one of the 10 runs of NEMODPS. First and second panels report, respectively, the best value of the Flood and Irrigation objectives in the population, across the 1000 generations of the evolution. The third panel represents the value of the Hypervolume indicator during the evolution.*

et al., 2003) scored by each seed during the search progression. HV accounts for both convergence and diversity of an approximate set of solutions $\mathcal{F}$ capturing the behavior in intermediate regions of the front, with respect to the best known approximation Pareto optimal set $\mathcal{F}^*$, constituted by the front resulted from the combination of the 10 seeds approximation. The hypervolume measures the volume of objective space Y dominated ($\preceq$) by the considered approximate set, with HV formally defined as:

$$\begin{aligned} \mathrm{HV}(\mathcal{F}, \mathcal{F}^*) &= \frac{\int \alpha_{\mathcal{F}}(\mathbf{y})\mathrm{d}\mathbf{y}}{\int \alpha_{\mathcal{F}^*}(\mathbf{y})\mathrm{d}\mathbf{y}} \quad \text{where} \\ \alpha_{\mathcal{F}}(\mathbf{y}) &= \begin{cases} 1 & \text{if } \exists \mathbf{y}' \in \mathcal{F} \text{ such that } \mathbf{y}' \preceq \mathbf{y} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{6.21}$$

Its generational growth somewhat mirrors the search for the best irrigation solution, and by the end of the search, the worst solution covers over 90.5% of $\mathcal{F}^*$, remarking a satisfying convergence and a limited dependency of NEMODPS solutions on initial conditions. Lastly, the NFE assigned to the evolution appear more than sufficient to reach convergence, given that the Hypervolume indicator does not significantly improve for any seed in the second half of the search.

6.6 Conclusions

In state-of-the-art applications of Direct Policy Search, the control policy is a priori defined as a fully-connected, single-layer, homogeneous neural network, independent of the problem characteristics or the objectives tradeoffs. This choice is motivated by theoretical results that assert the universal approximation capabilities of a wide range of network architectures. Many real-world applications, however, demonstrate a key role of topology in determining a network's approximation skills and training requirements. Our results show that traditional DPS with such predefined policy topology is prone to overfitting in noisy environments, and does not offer enough flexibility in MO problems, where different tradeoffs should be associated with different network architectures. By embedding NeuroEvolutionary (NE) techniques into the DPS framework, we extend the DPS problem to search a hyperspace containing control policy architectures and parameters. Yet, existing NE techniques, most notably NEAT and NEAT-inspired alternatives, are tailored to SO problems, and demonstrate a limited capacity to produce a high-quality approximation of the Pareto front in terms of solutions distribution and performance, while also requiring a substantially higher computational effort when compared to MO routines. This work contributes NEMODPS, a novel policy search algorithm which features the structure of the neuroevolutionary benchmark NEAT,

several NEAT improvements proposed in the literature, and an original strategy to extend the routine to MO problems, exploring a multidimensional objective space in a single run of the algorithm. Numerical results show significant consistency in topological features of networks optimized across independent runs of NEMODPS, suggesting that the generated control policy architecture is rational and depends on the characteristics in the fitness landscape. Moreover, a change in objective tradeoff corresponds to a change in fitness landscape, and the Pareto-approximate topologies adjust accordingly. Finally, neuro-generated control policies demonstrate the ability to handle noisy environments featuring remarkable reliability, and generalization potential with respect to benchmark fixed-structure DPS solutions when tested on a suite of diverse validation experiments.

7

Dynamic retrieval of informative inputs for multi-sector reservoir policy design with diverse spatio-temporal objective scales

Abstract[1]

Advances in monitoring and forecasting water availability at various time and spatial scales offer a cost-effective opportunity to enhance water systems' flexibility and resilience by anticipating hydrological extremes. Currently, most reservoirs' control rules are conditioned upon basic feature representations, i.e., time index, reservoir storage, previous day's inflow. However, the advantage of enriching a policy's feature representation is generally undisputed. Numerous candidate hydro-meteorological variables and forecasts may potentially be included in operation design, and the best input set for a given problem is not always evident. Additionally, in multi-purpose systems characterized by multiple demands with varying temporal scales, the most appropriate information set might change according to the objective tradeoff. In this work, we contribute a novel feature representation learning approach that links a feature selection routine with a multi-objective Direct Policy Search framework in order to retrieve the best policy input set online (i.e., while learning the pol-

[1]This work is currently in preparation as: Zaniolo, M., Giuliani, M., Castelletti, A., 2020c. Neuro-evolutionary direct policy search for multi- objective optimal control. Environmental Modeling and Software (in preparation).

icy) and tradeoff-dynamically. The selected policy search routine is the Neuro-Evolutionary Multi-Objective Direct Policy Search (NEMODPS) which generates flexible policy shapes adaptive to changes in the input set. This approach is demonstrated on the case study of Lake Como (Italy), where the operating objectives are highly heterogeneous in their dynamics (fast and slow) and vulnerabilities (wet and dry extremes). We show how varying objectives, and tradeoffs therein, benefit from a different feature representation, ultimately yielding remarkable results in terms of conflict mitigation between different users.

7.1 Introduction

Complex real world decision making problems are often approached via Reinfocement Learning (RL) techniques (Sutton et al., 1998). Recent advances in RL are expanding the boundaries of its domain beyond the native robotics and industrial applications to healthcare, (Shortreed et al., 2011), education (Mandel et al., 2014), and several water resources control applications, such as scheduling of water allocation networks (Abolpour et al., 2007), water-energy systems (Giudici et al., 2019), and the control of water reservoirs (e.g., Castelletti et al., 2011b). Real-world applications pose two main challenges to RL, namely, 1) Defining an informative and compact set of features to condition the control policy, i.e., *Learning a Feature Representation*; 2) The coexistence of conflicting operating targets, i.e., *Multiple Objectives*.

Learning a Feature Representation: Typically, a RL agent performs actions on a dynamic environment and collects a reward signal as the environment transits into a new state. By repeated interactions, the agent aims to learn an action policy, or control policy, that maximizes its cumulated reward. The success of a RL agent is determined by the subset of features conditioning its control policy that allow the agent to interact with its environment (Bu et al., 2008). Yet, in real-world RL applications, the environment is a sparse and heterogeneous entity comprising a number of processes and dynamics whose relevance for the agent's task is often unclear. For instance, considering the control of a water infrastructure operating in a hydrological basin, the observable environment comprises the water available in the basin under multiple forms, e.g., reservoirs, snow, groundwater, and topsoil humidity. Such water availability is influenced by the atmospheric conditions in the watershed, including temperature, winds, relative humidity, cloud cover, and precipitation (Turner and Galelli, 2016; Denaro et al., 2017a). In addition to observable features, all their possible manipulations can also be considered, including spatial and temporal aggregations, indices, models, and forecasts. Yet, likely, only a subset of this information set is relevant for the agent's task, and considering redundant or irrelevant information in the problem formulation will exponentially increase

the policy learning computational time, and degrade its performance. On the other hand, however, the control policy of a water reservoir is typicaly conditioned upon very limited information systems comprising time index, reservoir storage, and sometimes the previous day's inflow (Hejazi et al., 2008a), and the potential of including additional information to enhance the system flexibility and resilience is generally undisputed and long recognized (e.g., Stedinger et al., 1984; Kim and Palmer, 1997b; Maurer and Lettenmaier, 2004; Zhao et al., 2014; Giuliani et al., 2015; Zaniolo et al., 2019). The problem of deriving a compact and informative set of features to condition the control policy is called feature representation learning, and will be the main focus of this work.

Multiple Objectives: The coexistence of Multiple Objectives (MO) is a major challenge to many real world control problems. The result of a MO control problem is not a single optimal policy, but an ensemble of Pareto-efficient policies that explores the tradeoff space of conflicts and compromises between objectives. In the operation of multi-purpose water reservoirs, common operating targets, e.g., flood protection and irrigation supply, can be vastly heterogeneous in their dynamics and vulnerabilities. Flood events are caused by the onset of fast and intense wet meteorological extremes, while irrigation supply failures are the result of a prolonged period of water shortage caused by slow-developing dry hydrological extremes, i.e., droughts (Quinn et al., 2019). The tradeoff space between these two extreme strategies is populated by an ensemble of policies that balance the relative preference between these opposite control targets likely relying on a complex mixture of indicators. One policy input set is thus inadequate in multipurpose systems, and feature representation should be searched dynamically, to appropriately characterize the entire set of alternative control behaviors.

In this work, we consider the control problem of a multi-purpose water reservoir for which deriving an appropriate feature representation is not an obvious task. The control problem will be solved offline, via simulation on a virtual environment and using an available dataset of precollected environmental signals as candidate policy inputs. While the issue of deriving an appropriate feature representation was identified since the early applications of RL (Dominey, 1995; Sutton et al., 1998), the literature on this topic has been overwhelmingly focusing on single-objective applications, and still remains a vastly unexplored issue in MO RL problems (Liu et al., 2014). Here, we contribute a novel automatic feature representation learning routine, namely the Automatic Feature Selection for NeuroEvolutionary Multi-Objective Direct Policy Search (AFS-NEMODPS), specifically addressing feature representation learning for MO control problems.

In the literature, feature representation learning is generally tackled by pairing Feature Extraction algorithms with Policy Search methods (Liu et al., 2015;

Lesort et al., 2018). Feature extraction techniques are an ensemble of data processing tools used to transform the original dataset into a more compact, but still informative, feature set. Three main approaches can be identified when pairing feature extraction and policy search for feature representation learning, namely *a priori*, *a posteriori*, and *online*. In the *a priori* approach, the feature extraction step is antecedent and independent from the policy search step, and the dimensionality of the dataset of candidate features is reduced on the basis of intrinsic properties of the controlled system (Morimoto et al., 2008; Nouri and Littman, 2010; Zaniolo et al., 2019). In general, a priori feature representation is advisable whenever there is sufficient knowledge of the task to confidently devise an appropriate dimension reduction. This very low computationally demanding approach, in fact, does not offer any guarantees on the optimality of the chosen representation. The *a posteriori* approach evaluates the suitability of a feature representation by assessing the performance of the policy conditioned upon it. Multiple policies are designed with alternative feature representations, and the desired representation is identified as the one generating the best performing policy (Gaudel and Sebag, 2010; Giuliani et al., 2019). Both *a priori*, and *a posteriori* approaches in general rely on heavy expert-based manual engineering in defining potentially appropriate feature representations to implement or test (Bengio et al., 2013). The *online* approach is, instead, fully automatic and is preferred when the dataset of candidate policy input is large, and the pre-existing understanding of the task is not adequate for an appropriate input screening. This method interleaves feature extraction phases throughout the policy search process, using progressively refined representations to support policy learning. In *online* routines, typically, feature representation is progressively refined by extracting features that approximate the state space (Curran et al., 2016; Alvernaz and Togelius, 2017), state-transition space (Assael et al., 2015; Van Hoof et al., 2016), or the reward trajectory (Munk et al., 2016; Oh et al., 2017) of the policy learned thus far (for a comprehensive review, see Lesort et al., 2018). The progressively updated representation is employed to refine the control policy in a feedback loop between feature extraction and policy search.

AFS-NEMODPS is the novel *online* feature representation learning routine introduced in this work. Starting from a minimal feature representation and policy architecture, AFS-NEMODPS progressively builds a more informed control policy by complexifying its input set and processing capacity. The implementation of an *online* technique requires the selection of a i) feature extraction routine, ii) a policy search routine, and iii) a strategy to interface the two. Below, we comment on what routines are frequently chosen in the literature of *online* methods, and we motivate the choices made in building AFS-NEMODPS.

i *Feature extraction routine*: most *online* applications reduce the dimensionality of the policy representation by projecting the initial feature space into

a lower dimensional latent space that preserves information content albeit losing physical interpretability. Experiences in representation learning make wide use of these projection techniques, e.g., via auto-encoders (Goroshin et al., 2015; Watter et al., 2015; Van Hoof et al., 2016; Alvernaz and Togelius, 2017), or PCA (Karakovskiy and Togelius, 2012; Curran et al., 2016). However, such an approach compresses the dataset of candidate policy inputs, but does not guarantee that any candidate feature is actually excluded from the problem formulation. In principle, all the variables constituting the initial dataset can contribute to the compressed representation, even to a negligible extent (Loscalzo et al., 2015). As a result, while the control policy can actually benefit from a lower-dimensional representation, the actual problem size remains unchanged. In an operational setting, this implies that all features must be retrieved continuously. Alternatively, Feature Selection methods are a subset of the feature extraction techniques that reduces the dataset size by identifying a subset of the initial features. Some authors suggest the use of feature selection routines, rather than information encoders, for representation learning (e.g., Loscalzo et al., 2012, 2015), in order to effectively restrict the number of candidate variables included in the problem formulation. The representation obtained through variable selection, moreover, highlights relevant policy drivers, is easily interpretable, and can thus generate insights on the task at hand. Within Feature Selection techniques, the iterative online framework can accommodate simple correlation-based selection, as well as highly non-linear randomized models (e.g., extra-trees Castelletti et al., 2010b). Information theoretic applications suggest the use of entropy based metrics that estimate the mutual information content of two variables (Hachiya and Sugiyama, 2010). Here, we use a normalized version of the conditional mutual information, namely, the Symmetric Uncertainty (Singh et al., 2014).

ii *Policy Search*: The application of exact RL methods for policy search becomes computationally intractable for non-trivial applications characterized by high-dimensional state spaces (curse of dimensionality, Bellman, 1957), the presence of multiple conflicting control objectives (curse of multiple objectives, Powell, 2007), and the need to explicitly model every variable included in policy representation (curse of modeling, Tsitsiklis and Van Roy, 1996). Direct Policy Search (DPS) is an approximated RL method that has emerged as a promising alternative to mitigate the three curses of exact methods, thus scaling well to complex real world control problems (Giuliani et al., 2016b). DPS defines the control policy within a prespecified class of functions and solves a problem of optimal functional parameterization with respect to the problem's objectives (Salazar et al., 2017; Quinn et al., 2018; Giuliani et al., 2019; Quinn et al., 2019). Flexible universal ap-

proximators such as Neural Networks (NNs) are generally employed to parameterize the control policy in order not to restrict the parametrical search to a small functional subspace that may not contain skillful solutions (Giuliani et al., 2014b, 2018). The architecture of a NN employed for policy design includes as many input nodes as the number of features in the policy representation, and as many output nodes as the actions to be taken on the environment. The internal NN complexity, (i.e., number of hidden nodes, connections, and layers) is instead arbitrarily adjusted by the modeler, generally by manual trials and errors, based on the problem complexity. Given this rigid prespecified policy structure, DPS techniques do not support dynamic changes in the dimensionality of the policy feature representation. In fact, previous efforts in coupling an Evolutionary Multi-Objective Direct Policy Search (EMODPS) with feature selection, restart the policy search from scratch every time the feature representation is updated with a new input (Giuliani et al., 2015; Denaro et al., 2017a).

A promising alternative that obviates policy rigidity is represented by NeuroEvolution (NE), a set of techniques that employ evolutionary algorithms to evolve neural networks structures and parameters. By pairing NE with DPS, it is possible to derive policy search routines that support online changes in policy architecture. Popular NE algorithms (e.g., NEAT Stanley and Miikkulainen, 2002) are, however, strictly applicable to single-objectives problems. The here-employed NeuroEvolutionary Multi-Objective Direct Policy Search (NEMODPS), is the first NE routine specifically designed to solve MO problems in one algorithmic iteration (Zaniolo et al., 2020).

iii *Interfacing strategy*: in many applications, the selection of relevant features is performed via supervised learning using as target the state, state-transition, state-value spaces, or the reward trajectory produced by the policy learned thus far (for a review, see Lesort et al., 2018). Reward-based selection is generally recognized as more effective in identifying task-oriented policy representations (Loscalzo et al., 2015), however, in MO RL problems, the coexistence of multiple reward signals complicates the reward-based selection process. In AFS-NEMODPS, we propose a novel interfacing strategy that is both task-tailored, and suitable for MO problems. In particular, we use as reference a deterministic Perfect Control Policy (PCP) that assumes full knowledge of future system disturbance. For a given state, we contrast the actions extracted from the PCP to those extracted from the policy under design. We assume that the difference in actions is due to the information gap in the policies representations, and thus surrogates the information that the designed policy would require to meet the PCP performance. The trajectory of action residuals is used as an interfacing strategy, and employed as target for feature selection. Such a target can be considered task-relevant, as

it is a proxy of the policy information deficiency for a given task. Additionally, it can be applied to MO problems by contrasting each Pareto efficient policy with the corresponding perfect counterpart supporting a tradeoff dynamic feature selection.

To summarize, AFS-NEMODPS combines feature selection, neuroevolution and an original interfacing strategy. The choices made in the selection and development of the building tools of AFS-NEMODS target the overarching goal of designing the first multi-objective feature representation learning routine that automatically specifies an optimal policy representations for each tradeoff. This framework is tested for the real-world case study of the multi-purpose Lake Como, operated for the two conflicting and heterogeneous objectives of flood protection and irrigation supply. The flood objective is characterized by a fast dynamic and vulnerability towards wet extremes, while the seasonal irrigation supply is characterized by a slow dynamic and vulnerability towards dry extremes. Results show that a tradeoff-specific policy representation is capable of significantly reducing conflicts between water users, and that varying objectives, and tradeoffs therein, benefit from different information.

This paper is organized as follows. The next section presents the methods of this work, by formulating the Problem of MO feature representation for RL 7.2.1, presenting the methodological Framework 7.2.2, and expanding on the key concepts and tools employed in the methodology, including NEMODPS 7.2.3. Section 7.3 is dedicated to the presentation of the case study and experimental settings. Results are discussed in Section 7.4, and in the following Section 7.5 we draw conclusions and introduce some discussion points.

7.2 Methods

7.2.1 Problem formulation

In this work, we consider a discrete-time continuous Multi-Objective Markov Decision Process (MOMDP) defined as a tuple $< \mathcal{X}, \mathcal{U}, \mathcal{T}, \mathbf{G} >$ where $\mathcal{X} \subset \mathbb{R}^{n_x}$ is the continuous state space, $\mathcal{U} \subset \mathbb{R}^{n_u}$ is the continuous action space, $\mathcal{T}(x_{t+1}|x_t, u_t, \varepsilon_{t+1})$ is the transition function between x_t and x_{t+1} under action u_t, and subject to disturbance ε_{t+1}, $\mathbf{G}(x_t, u_t, x_{t+1}) = \left[G^1, \ldots, G^M\right]$ is a M-dimensional reward (or cost) function that specifies the vector of instantaneous rewards (costs) $\mathbf{g_t} = [g_t^1, \ldots, g_t^M]$ for each objective when state x_{t+1} is reached from state x_t by taking action u_t. Action u_t is extracted from a control policy $\pi(x_t)$, $u_t = \pi(x_t, u_t)$, associated with a vector of expected returns

$\mathbf{J}(\pi) = \left[J^1(\pi), \ldots, J^M(\pi)\right]$ defined over the control horizon $[0, H]$ as:

$$J^m(\pi) = \mathbb{E}\left\{\sum_{t=0}^{H} (\gamma^m)^t g^m(t+1) | x_0\right\} \tag{7.1}$$

where $\boldsymbol{\gamma} = [\gamma^1, \ldots, \gamma^M] \in [0, 1]$ is the vector of discount factors relative to each objective.

The solution of the RL problem defined above is the policy $\pi^*(x_t)$ that yields the optimal value of objective $\mathbf{J}$ (here considered as a cost, to be minimized) in its M dimensions:

$$\pi^* = \arg\min_{\pi} \mathbf{J}(\pi, \mu) = \arg\min_{\pi} \left[J^1(\pi, \mu), \ldots, J^M(\pi, \mu)\right] \tag{7.2}$$

In general, in MO problems, conflicts occur between different operating objectives, and it is thus not possible to define a single optimal policy, representing the optimum with respect to the M dimensions of $\mathbf{J}$. The solution of a MO problem is in general constituted by a set of non-dominated (or Pareto efficient) solutions $\boldsymbol{\pi}^* = \{\pi^* | \nexists \pi \prec \pi^*\}$, which maps onto the Pareto front $\mathcal{F}^* = \{\mathbf{J}(\cdot) | \pi^* \in \boldsymbol{\pi}^*\}$.

The control policy $\pi^*(x_t)$ is searched in a n_x dimensional state space, that can in principle be very large for real-world systems, resulting in a computationally intractable MOMDP. We are interested in finding a low-dimensional feature representation space $\mathcal{X}^{LD}$, subset of the original state $\mathcal{X}$ that retains its task-relevant information, thus generating an approximately unaltered control policy:

$$\mathcal{X}^{LD} \in \mathcal{X} \;\mid\; \pi^{LD}(x_t^{LD}) \approx \pi(x_t) \tag{7.3}$$

where $\mathcal{X}^{LD}$ is defined in a low dimensional $\mathbb{R}^{n_x^{LD}}$ making $\pi^{LD}(x_t^{LD})$ computationally tractable.

We solve Problem 7.3 via DPS, which approaches policy design as an optimal functional parameterization, defining the control policy within a given function class π_θ^{LD}, and then searching the parameters' space Θ to find the optimal parameterization $\theta^* \in \Theta$ with respect to the M-dimensional vector of objectives $\mathbf{J}$.

Hence, Problem (7.3) can be reformulated for a given $\mathcal{X}^{LD}$ as:

$$\pi_\theta^{LD*} = \arg\min_{\pi_\theta^{LD}} \mathbf{J}(\pi_\theta^{LD}(x^{LD})) \tag{7.4}$$

An appropriate $\mathcal{X}^{LD}$ is however in general not known, and in this application is searched online while learning the policy, requiring the functional class to adapt to changes in the policy input set. We thus extend the DPS problem

formulation to the dynamical search of policy architecture and their relative parameterization $\zeta(\theta)$. Accordingly, Problem (7.4) is reformulated as:

$$\pi^*_{\zeta(\theta)} = \arg\min_{\pi_{\zeta(\theta)}} \mathbf{J}(\pi_{\zeta(\theta)}(x^{LD})) \tag{7.5}$$

where $\pi_{\zeta(\theta)}$ explicits the search for policy hyperparameters ζ defining a policy architecture (including number of policy inputs and internal network complexity), as well as regular policy parameters θ, whose number and nature depend on the hyperparameters as in $\zeta(\theta)$. The policy search problem is thus expanded to conjunctively search a reduced feature representation, policy architecture, and parameters.

7.2.2 Framework

In this section, we present the flowchart of the proposed AFS-NEMODS framework employed to approach Problem 7.5, reported in Figure 7.1. In the search for the optimal policy representation $\mathcal{X}^{LD}$, we start from a minimal policy input set and gradually enrich it with additional information.

As the procedure begins, in round R1, we initialize a population of simple neural networks, with a basic input set $\mathcal{X}^{R1}$, a minimal architecture, and random weights. This ensemble is the input to the Policy Search building block that employs NEMODPS. For a given input set, NEMODPS evolves policies' architecture and parameters in a MO problem (more details in the dedicated Subsection 7.2.3). The output of this step is an ensemble of Pareto efficient control policies $\boldsymbol{\pi}^{R1}(x_t^{R1})$, where each policy is specified with a tailored architecture, resulting in an architecturally heterogeneous population.

In the first round, the flowchart proceeds to the building block named Compute Residuals. In this step, we contrast the controls produced by each Pareto efficient policy, with the controls given by a Perfect Control Policy (PCP) (more details in the dedicated Section 7.2.4) extracting the trajectories of control residuals e_t.

Then, in the Feature selection step, we search a dataset of candidate policy inputs to identify the input that shares the most information content with a given residual trajectory, via the computation of the Symmetric Uncertainty metric. Because the trajectory of residuals is computed independently for each efficient policy, the inputs selected are policy-specific, and may vary across the tradeoff space.

Each efficient policy is then updated by including the selected feature in the input set, with a single input-output connection and a randomly initialized weight. The population of policies is now heterogeneous in its feature representation. In the second round R2, this population is further evolved via

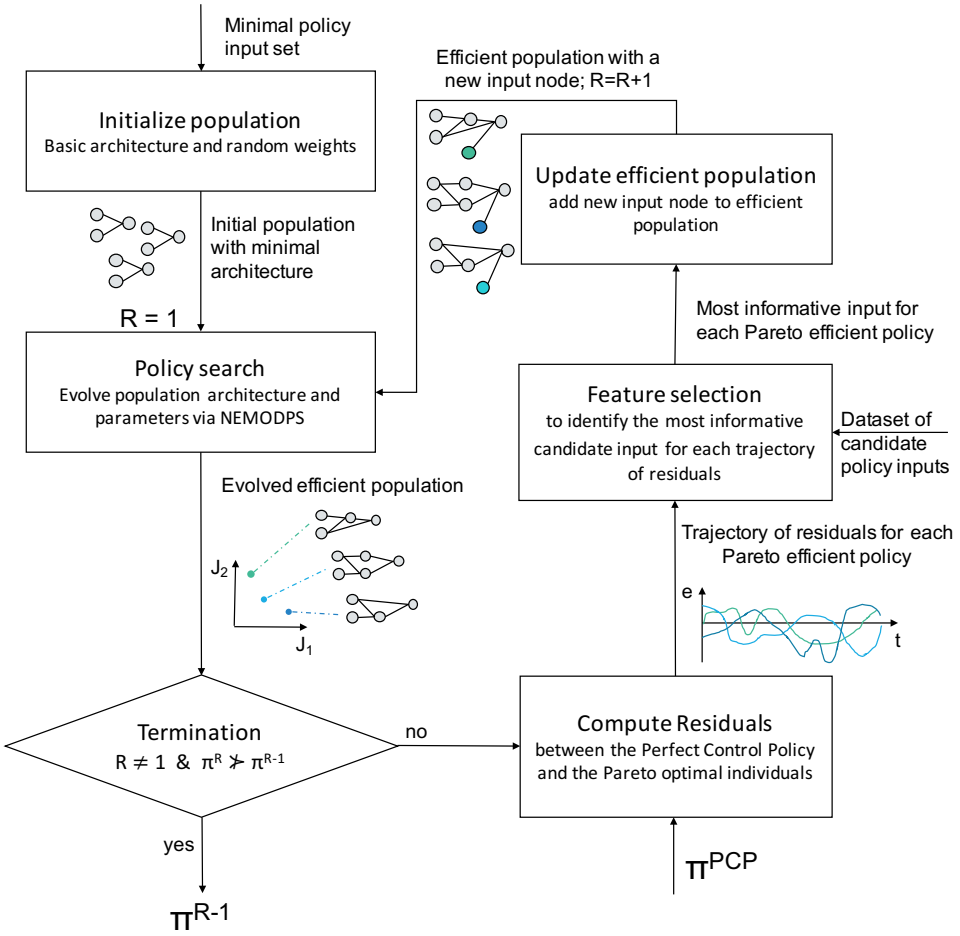

Figure 7.1: *AFS-NEMODPS flowchart. By looping through the building blocks of this flowchart, the procedure complexifies the initial population in terms of feature representation and policy architecture.*

NEMODPS. Individuals will appropriately complexify their architecture by genetic evolution to adapt to the newly inserted input, and learn how to make use of its information content. Neuro-evolutionary competition will further filter feature representation, causing only the fittest representations to survive in the efficient policies of round R2.

AFS-NEMODPS proceeds analogously until the Termination check is positive, namely when the efficient Pareto set at Round R does not significantly dominate the Pareto set in the previous round: $\boldsymbol{\pi}^{R} \not\prec \boldsymbol{\pi}^{R-1}$. More details on the termination criterion are presented in Section 7.2.5. Upon termination, we retain as efficient solutions the Pareto set generated at the previous round $R-1$, as it achieves virtually the same performance as round R with a simpler representation.

7.2.3 NEMODPS

In this section, we give an overview of the main components of the selected policy search routine. NEMODPS (Zaniolo et al., 2020) builds on a recent Reinforcement Learning branch called Neuro-Evolution (NE) (Stanley and Miikkulainen, 2003; Floreano et al., 2008), which employs Evolutionary Algorithms to generate optimal networks in terms of topologies and parameters. NEMODPS inherits the main features of a NE benchmark, NEAT (Stanley and Miikkulainen, 2002), and the subsequent literature of NEAT improvements targeting complex control problems, vast decision spaces, and noisy environments. Additionally, NEMODPS contains an original strategy to extend the search scope to MO algorithms.

Key elements of NEMODPS are (a) a process of evolutionary complexification, (b) the use of parametrical and topological operators, and (c) an architecture-based competition scheme that sustains solution diversity and avoids premature convergence.

(a) NEMODPS begins with a population of uniform simple networks with randomly initialized weights, and gradually builds more complex architectures by applying topological operators. As the evolution proceeds, increasingly sophisticated behaviors emerge, and the more promising ones survive in the following generations.

(b) Classic parametric evolutionary operators (i.e., mutation and crossover) are interpreted from a topological perspective. The topological mutation operator performs a randomized addition of a node (sigmoidal or Gaussian) or a connection to an individual, and assigns a unique historical marker to the new element. Topological crossover takes place among two individuals by aligning genes with the same historical markers and assigning the offspring a mix of the parents' genes, favoring those coming from the fittest parent.

Parametrical evolution is instead performed via a classical mutation operator. NEMODPS coordinates the search in a dual timescale, optimizing the network's parameters on a small timescale, i.e., every generation (exploitation phase), and the network's structure on a larger timescale, every few generations (exploration phase).

(c) The addition of a new topological element with randomized coefficients can be, at first, detrimental for an individual. As a result, simpler architectures tend to prevail over a newly complexified one, even if the topological innovation would prove beneficial if given enough iterations to optimize. NEMODPS thus implements a niching scheme to protect topological innovations from premature disappearance. In particular, the population is partitioned into niches (or species) of similar topologies by evaluating a metric of topological distance. A dynamically adjusted threshold controls the creation of new species as the population complexifies. As a result of the niching scheme, individual competition is only allowed within niches of similar topologies thereby protecting topological diversity. Species compete among each other for their ability to reproduce, so that a larger offspring is assigned to well performing ones. A fitness sharing mechanism penalizes populous species preventing them from taking over the entire population causing loss of topological diversity and premature convergence. NEMODPS generalizes the fitness sharing strategy for MO problems, rewarding species that perform well in a Pareto sense, while penalizing populous species and individuals in close proximity in the objectives space in order to encourage the exploration of the entire tradeoff space.

7.2.4 Extraction of optimal decision from a Perfect Control Policy

The perfect control policy π^{PCP} is designed by solving Problem 7.2 under the hypothesis of deterministic knowledge of the trajectory ε_1^H of external drivers over the entire evaluation horizon H at any given time step, and can be solved via Deterministic Dynamic Programming (DDP). Note that the hypothesis of designing the policy conditioned upon the future disturbance trajectory ε_1^H is not equivalent to conditioning it on the entire, unreduced, state space $\mathcal{X}_t$. Rather, it corresponds to considering the entirety of the present and future system states, in all its deterministic, stochastic, and pure white noise components, perfectly distilled in a single feature, which would never be available for real online operations. Such a deterministic policy should be considered as a mathematical object that constitutes the reference for improving a basic policy design, but cannot be realistically implemented in a real-world system. For our purposes, we consider the PCP as conditioned upon the minimal input set and the future disturbance trajectory $\pi_t^{PCP}(x_t^{MIN}, \varepsilon_1^H)$. In order to extract the trajectory

of policy residuals, we compare the actions extracted from the π^{PCP} with those extracted from each efficient policy at a given round R, π^R referring to the same minimal state trajectory produced by the simulation of π^R. The difference in actions is assumed to be due to the information gap between the the policy under design, and the perfectly informed policy. For well designed systems, as the one considered in this study, operating with perfect knowledge of future disturbance allows to contain conflicts between objectives, resulting in a sharp-cornered PCP Pareto front. The corner solution represents the less conflictual solution and can be employed as target reference. In case a PCP front presents more than one corner solution with a substantial tradeoff, each efficient policy at a given round R, π^R can be associated the closest PCP policy. In particular, policy proximity can be evaluated in a normalised objective space with respect to maximum and minimum values assumed by the objectives at a given round.

7.2.5 Termination criterion

The framework terminates for rounds $R > 1$ when the efficient Pareto set at Round R does not significantly dominate the Pareto set in the previous round: $\pi^R \not\succ \pi^{R-1}$. Several metrics could in principle be used to express dominance in a Pareto sense. Here, as suggested in (Giuliani et al., 2015), we use the hypervolume indicator (HV), which captures both the convergence of the Pareto front under examination $\mathcal{F}$ to the optimal one $\mathcal{F}^*$, as well as the representation of the full extent of tradeoffs in the objective space. The hypervolume metric allows set-to-set evaluations, measuring the volume of objective space Y dominated ($\preceq$) by the considered approximate set. HV assumes values between 0 to 1, where Pareto fronts with higher HV are considered better. The indicator is formally defined as:

$$\begin{aligned} HV(\mathcal{F}, \mathcal{F}^*) &= \frac{\int \alpha_{\mathcal{F}}(\mathbf{y})\,d\mathbf{y}}{\int \alpha_{\mathcal{F}^*}(\mathbf{y})\,d\mathbf{y}} \quad \text{where} \\ \alpha_{\mathcal{F}}(\mathbf{y}) &= \begin{cases} 1 & \text{if } \exists \mathbf{y}' \in \mathcal{F} \text{ such that } \mathbf{y}' \preceq \mathbf{y} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{7.6}$$

7.3 Case Study and Data

We consider the control problem of Lake Como, a multipurpose regulated lake located in the southern Alpine belt, Italy (Fig. 7.2). The main tributary, and only emissary of the lake is the Adda river, whose sublacual reach originates in the southeastern branch of Lake Como, crosses the Po valley, and eventually serves as a tributary to the Po river downstream. In its course, part of its waters are withdrawn to irrigate four agricultural districts. The southwestern branch of

Lake Como constitutes a dead end, and exposes the city of Como to flooding events. The Lake Como basin hydrological regime is snow-rainfall dominated, characterized by scarce winter and summer inflows, a large snowmelt peak in late spring, and a secondary rainfall peak in autumn.

The system is modeled as a discrete-time, periodic, non-linear, stochastic process defined by a state variable x_t, a control variable u_t representing the release decision from the dam gates, stochastic disturbances ε_{t+1} (net reservoir inflow), and a state-transition function $f(\cdot)$: $x_{t+1} = f(x_t, r_{t+1}, \varepsilon_{t+1})$ where the effective release r_{t+1} coincides with the release decision u_t corrected, where appropriate, with a non-linear release function $R_t(x_t, \varepsilon_{t+1})$ determining the minimum and maximum releases feasible for the time interval $[t, t+1)$ to respect physical and legal constraints. The Adda River is described by a plug-flow model, which simulates the routing of the lake releases from the lake outlet to the intake of the irrigation canals. The adopted time step is 1 day, and the system is periodic with period T = 365 days (Giuliani et al., 2016c).

Figure 7.2: *Digital Elevation Model of the Lombardy region, Italy. The highlighted elements are Lake Como and its Alpine basin, the sublacual Adda River, the city of Como and the downstream irrigation districts.*

The lake regulation has two conflicting aims of supplying water to downstream users by storing spring snowmelt peak, and minimizing flood risk on the lake shores by maintaining the lake level as low as possible. On the basis of previous works (Castelletti et al., 2010b), these two objectives are defined as:
Flooding: the average number of annual flood days, defined as days in which

the lake level h_t is above the flood threshold $\bar{h} = 1.24$ m, i.e.,

$$J^{flood} = \frac{1}{N_y} \sum_{t=0}^{H-1} g_{t+1}^{flood}; \quad g_{t+1}^{flood} = \begin{cases} 1 & \text{if } h_{t+1} \geqslant \bar{h} \\ 0 & \text{if } h_{t+1} < \bar{h} \end{cases} \tag{7.7}$$

where N_y is the number of years in the simulation horizon. Irrigation: the daily average squared water deficit with respect to the daily downstream demand w_t, subject to the minimum flow constraint $q^{MEF} = 5$ m^3/s to guarantee environmental stakes. The quadratic formulation is selected with the aim of penalizing severe deficits in a single time step, while allowing for more frequent, small shortages. i.e.,

$$J^{irr} = \frac{1}{H} \sum_{t=0}^{H-1} (max(w_t - (r_{t+1} - q^{MEF}), 0))^2 \tag{7.8}$$

We assume the considered simulation horizon $\mathcal{H}$ is sufficiently long to not require the addition of a penalty function to the final state. Additionally, since the policy is conditioned upon a cyclic representation of time, the decision towards the end of the time horizon is based on similar conditions taking place in previous years. For this application, we used Lake Como inflow data for a 10-year simulation horizon from 1997 to 2007. This time span contains a diverse range of hydrological conditions, including average and extreme years, from the 2005 record drought to the late 2000 high inflow pulses. This time horizon is used for both optimization and evaluation of the policy performance, thus ensuring that the PCP represents the optimum in terms of system operations.

The set of candidate policy inputs employed in this analysis includes perfect forecasts of the lake inflow computed over the historical timeseries at different lead times, ranging from one day to over 6 months (Table 7.1). The forecasts are provided in terms of cumulative values of inflows as well as inflow anomalies with respect to the inflow cyclostationary mean. As per previous works, the use of perfect forecasts removes possible modeling biases in forecasts construction (Zhao et al., 2011; Denaro et al., 2017a).

7.3.1 Experimental Settings

In these experiments, the minimal policy input set comprises the current reservoir storage and two transformations of the time index t with sine and cosine, to embed time-variability and cyclostationarity in the control policy $X_t^{min} = |x_t, sin(t), cos(t)|$.

AFS-NEMODPS was run for 20 independently initialized and randomized seeds. As termination criterion, we set a threshold for HV improvement of 5%.

Days ahead	Feature name	
	Cumulated inflow	Inflow Anomaly
1	C1	A1
2	C2	A2
3	C3	A3
5	C4	A5
7	C7	A7
14	C14	A14
21	C21	A21
28	C28	A28
51	C51	A51
62	C62	A62
75	C75	A75
90	C90	A90
120	C120	A120
145	C145	A145
200	C200	A200

Table 7.1: *Dataset of candidate policy inputs comprising perfect inflow forecasts in terms of cumulated inflows and anomalies at various lead times.*

In each seed, the termination criterion is met at the 4th round, which is responsible for no tangible advancement in the Pareto front, (lower than 5%), therefore, we retain as efficient solutions those generated at round 3. At each round, NEMODPS is run for a Number of Function Evaluations (NFE) equal to 600 thousands, with populations of 600 individuals. When feature selection identifies new policy inputs, these are connected to the efficient population of the previous round through an input-output connection. This set of individuals is then replicated until the population dimension is equal to 600 with randomized connection weights, and constitutes the initial population of the new round of NEMODPS optimization.

7.4 Results

Figure 7.3 reports the Pareto fronts resulting from 3 optimization rounds of AFS-NEMODPS with respect to the two objectives of Irrigation deficit (vertical axis) and Flood days (horizontal axis), both to be minimized as indicated by the arrows. The black square in the bottom left corner of the graph represents the ideal performance of the PCP. In accordance with other studies on the same water system, we find that the conflicts between irrigation and flood objectives

in Lake Como disappear under the assumption of perfect knowledge of future inflow (Denaro et al., 2017a). A control policy with full foresight is able to guarantee a sufficient flood pool to buffer the peak inflow and avoid overflow when physically possible, while storing in the lake any excess of water to be used for irrigation purposes during the dry season. Therefore, the deterministic solution of this MO problem does not yield a Pareto front of efficient solutions, but collapses to a single optimal point into the objective space. However, in the absence of a perfect future foresight, we expect that the addition of tailored information can reduce conflicts between water users.

The first round of optimization, conditioned upon basic information only, produces the Pareto front of white circles that lays in the top right portion of the objective space in Figure 7.3a, showing a sharp conflict between the two operating objectives. Additionally, a concavity can be recognized in the central region of the Pareto front, for values of the Flood objective between 20 to 80. Concave regions of the front are usually regarded as disadvantageous tradeoffs, as one objectives degrades more than proportionally to the second objective's improvement. The normalized HV indicator (panel b) relative to round R1 scores 0.142, indicating a large space for improvement between PCP and R1.

In the second optimization round, indicated with colored triangles, the more informed policies significantly outperform R1, scoring an over 3-fold increase in the HV metric. The shape color corresponds to the new feature added to the policy representation, and divides the R2 front in two, around its middle and in correspondence to the persisting concavity in the Pareto front. If we analyze the information selected, results may seem unexpected: flood-inclined policies do not select short term predictions of fast inflow peaks, but long forecasts lead times (75 days). Vice versa, irrigation-inclined policies select, in comparison, shorter lead times (62 days) instead of preferring season-long look-ahead. This behavior can be explained from the point of view of conflict mitigation. A minimally-represented flood-inclined policy has, in fact, already developed a solid strategy to prevent floods when physically possible, namely, keeping a low lake level for the most part of the year to always count on a buffer pool to accommodate incoming inflow peaks. This strategy is valid from a lakeshore protection perspective, yet, comes at a remarkable price in terms of irrigation supply. Such a policy, therefore, does not require any additional information on upcoming inflow peaks, as the lake is virtually always ready to buffer them. On the contrary, it can significantly benefit from a longer term information on how to improve irrigation supply while still remaining strongly flood risk-averse, thereby alleviating water supply deficits downstream, and mitigating conflicts between water users. In fact, by comparing flood conservative policies of R1 and R2 (left region of the Pareto fronts), we notice that the added information has the effect of improving the policies in the direction of a significantly lower

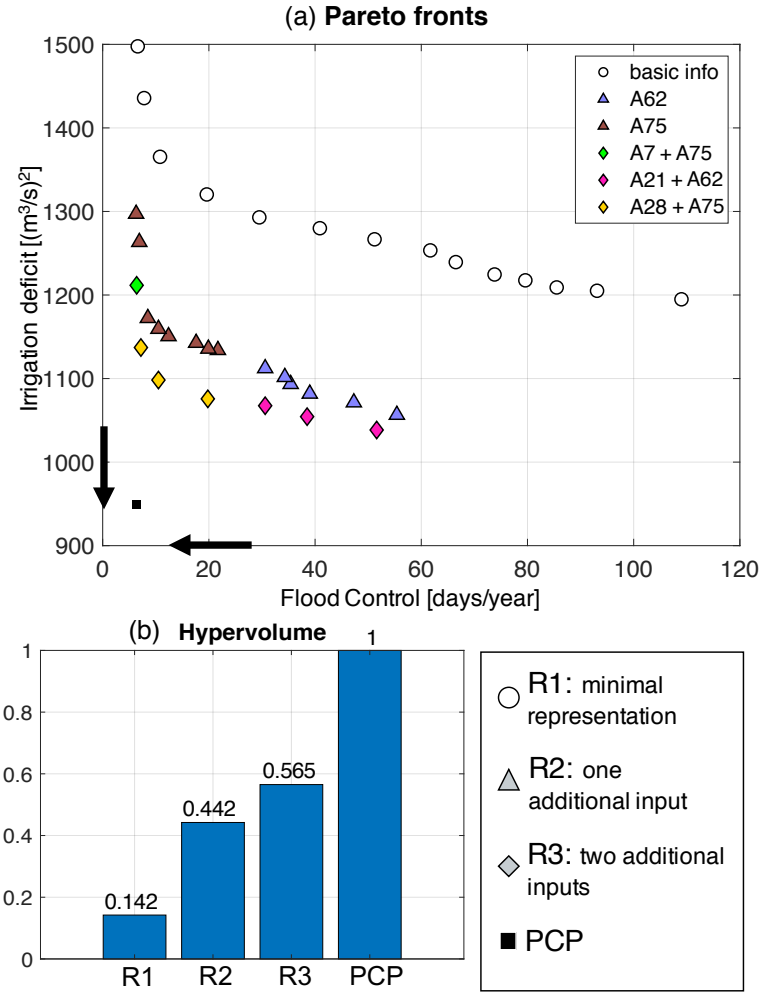

Figure 7.3: *Panel (a): Performance obtained by different Lake Como operating policies with respect to the two cost objectives of irrigation deficit (vertical axis) and Flood days (horizontal axis). The black square indicates the ideal performance of the PCP, white circles the performance of efficient policies designed at round R1, triangles refer to policies at round R2, and diamonds at round R3. For rounds R2 and R3, the shape color is associated with the information added to the feature representation. Panel (b) shows the improvements in the Hypervolume indicator across different rounds, normalized to the value of hypervolume scored by the PCP.*

irrigation deficit, at no cost for the flood objective. The long lead time information selected by flood oriented policies is thus employed to minimize objectives conflicts, rather than further improve the flood objective. The other half of the Pareto front selects a shorter lead time, which allows policies to move both in the direction of a reduced flood and irrigation damage.

The third optimization round includes a second additional information source as a policy input, generating further improvement in the HV indicator. The Pareto front of round R3 not only dominates the fronts of the previous rounds, but also resolves their concavity generating a fully convex front, where it is possible to identify a knee. Contrary to the previous round, the front shift between R2 and R3 is mainly horizontal, i.e., contributing to a Flood objective improvement rather than an irrigation improvement. Accordingly, the policy inputs selected in this round have a much shorter lead time, between 1 and 4 weeks. The solutions that at this round select the longer lead time, 4 weeks, are those showing a diagonal improvement that unfolds in both objective directions.

In Figure 7.4 we explore how added information is employed by progressively informed policies for a given tradeoff. This analysis focuses on the solutions located along the lilac vertical line in panel (a), corresponding to an average of 6.3 flood days per year. This tradeoff was chosen in order to compare the 4 Pareto fronts only in terms of the irrigation objective, for a given flood performance. A common cyclostationary behavior emerges for different policy representations in panel (b). The lake recharges in May, in correspondence to the onset of the irrigation season, reaches a level peak around late June, followed by an emptying phase lasting for the entire irrigation season until September/October, when abundant rains cause a new level increase. In the PCP, perfect future foresight informs the policy on the exact onset of inflow peaks, allowing to timely generate an adequate flood pool to contain them, while keeping, on average, a high lake level that ensures water availability to supply downstream irrigation demand. Whenever the full trajectory of future disturbance is not available, policies have to be more conservative towards flood events, thereby keeping a lower lake level to buffer possible incoming inflow peaks, at the expense of irrigation availability. This behavior is sharper in the minimally informed round R1 (red line), while more informed policies can confidently maintain a fuller lake during the summer, resulting in a smaller water deficit downstream, without damaging the flood objective. Cyclostationary behaviors outside the irrigation season are fairly divergent, however, the system's winter downstream demand is almost negligible with respect to summer demand, thereby not contributing significantly to the irrigation objective performance.

Lastly, in Figure 7.5, we analyze how a refinement in policy representation operationally modifies lake regulation towards conflict mitigation. The shaded area in panels (a), (b), and (c) delimits the ensemble of lake level trajectories asso-

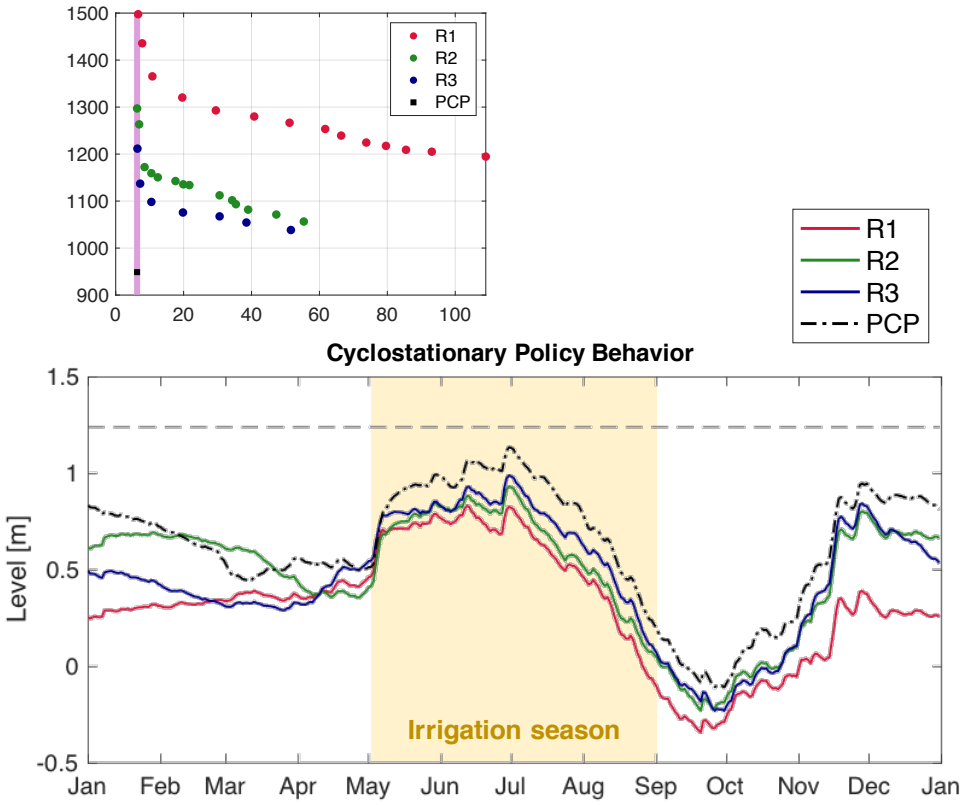

Figure 7.4: *Cyclostationary behavior of efficient policies across different optimization rounds. The investigated policies are aligned along the lilac line in the Pareto front of panel (a) and yield an average number of flood days equal to 6.3, and different values with respect to the irrigation objective. In panel (b), their cyclostationary behavior is shown, and contrasted with the perfect control policy.*

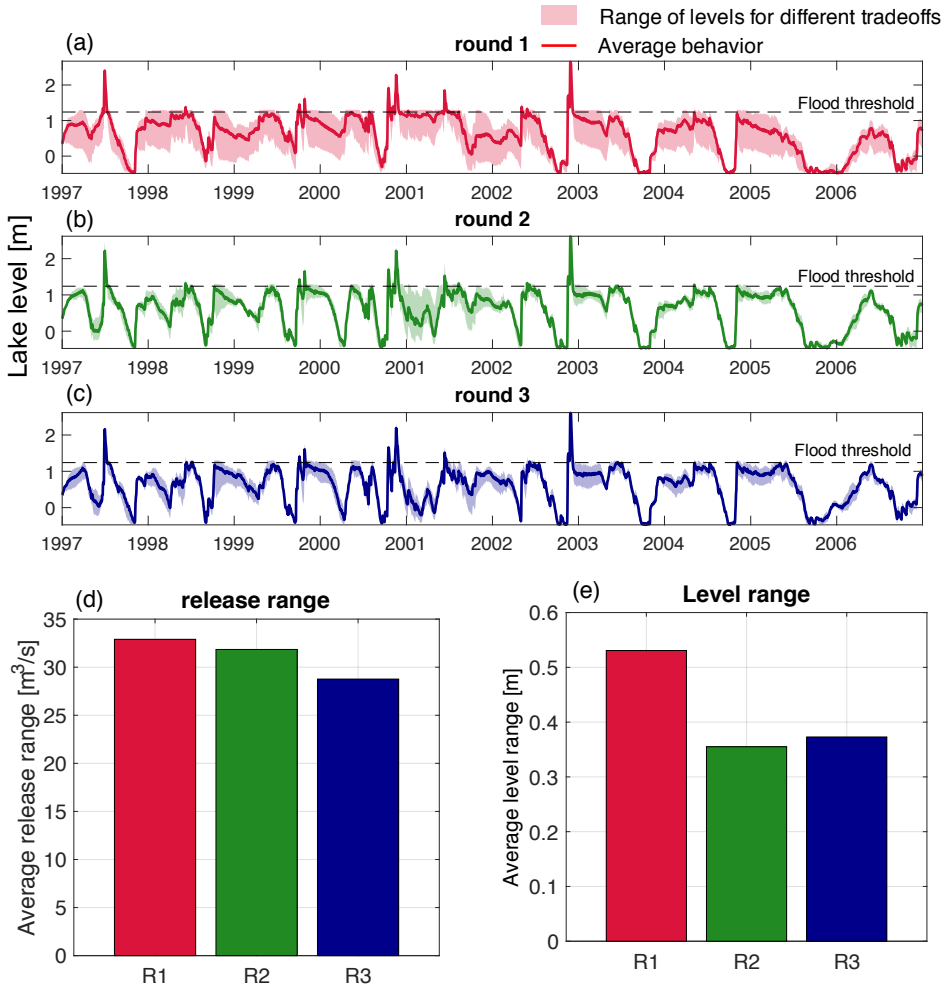

Figure 7.5: *Conflict mitigation. Panels (a), (b), and (c) report the range of lake levels yielded by all the Pareto efficient policies designed at the given optimization round across different tradeoffs. The average round-specific release range is quantified in the barplot of panel (d), while the lake level range is shown in panel (e).*

ciated with the set of Pareto efficient policies produced in a given round, while the central bold line represents the average behavior. The width of the shaded area indicates the range of variability in operations spanned by the efficient policies, where a thick area indicates that different tradeoffs are associated with diverse operations, and a narrow area suggests similar operations even across opposing tradeoffs. The plots show a visible narrowing in the operational variability from the first round to the following ones. Operationally, this translates into a mitigated conflict between water users, as different interests tend to converge towards a common efficient policy. This convergence is quantified in the barplots showing the average daily range in levels (panel e) and releases (panel d) associated with different policies in the Pareto set resulting from a given round. The addition of information in the policy representation shows a consistent reduction in release variability. Level variability significantly drops from round R1, where Lake Como is operated at an average difference of more than 53 cm for different tradeoffs, to about 35 cm in round R2. R3 shows a slight increase in variability that is however below 2 cm, and can be considered negligible.

7.5 Conclusions

While Feature Representation Learning is widely addressed in the literature, previous works do not specifically consider how defining one single policy representation to characterize the entire tradeoff space of multi-purpose systems might be inadequate. The coexistence of fast and slow process dynamics, and different vulnerabilities requires the search of a tradeoff tailored policy representation. In this work, we propose AFS-NEMODPS, a novel framework for automatic, tradeoff-dynamic feature representation and policy learning. AFS-NEMODPS starts with a population of minimal policies and gradually complexifies their feature representation by selecting variables that surrogate the policy information deficit, measured by comparison to a perfect control policy. Policies' architectures are adjusted accordingly, in order to accomodate new inputs and support more complex behaviors. We apply AFS-NEMODPS to the case study of Lake Como, characterized by conflicting heterogeneous objectives, and we use a dataset of deterministic inflow forecasts at different lead times as candidate policy inputs.

Results show that different objective tradeoffs benefit from different information sets with unexpected, but insightful, outcomes. Flood-conservative policies select forecasts with long lead times, thereby improving irrigation performance without increasing flood failures. Irrigation-inclined policies select, in comparison, shorter lead times achieving better flood and irrigation results. Not only do we notice a trend in the information selected for different tradeoffs, but also across successive selection rounds. The first forecast included in

the representation at the second round counts on a forecast with over 2 months lead time, and produces the largest improvement in the direction of a lower water supply deficit, and only partially, flood mitigation. In round three, lead times are shorter than a month, enhancing primarily flood mitigation skills.

Overall, the search for a tradeoff-specific feature representation demonstrates the potential to significantly enhance the water system overall reliability, resilience towards both dry and wet extremes, while reducing conflicts across conflicting water uses.

7.6 Discussion: the role of competition in feature representation

The results presented in this chapter are being refined at the time of writing this thesis. One aspect in particular seems to be worth further investigation i.e., the information selection process. In AFS-NEMODPS, the features eventually composing the efficient representations undergo a two-fold selection. First, a filtering process identifies promising features based on their correlation measured in Symmetric Uncertainty with the policy information gap. These features are included in the representation of the previous efficient individuals. Second, this newly initialized population, heterogeneous in its policy representation, is further evolved via NEMODPS and the fittest representations prevail through evolutionary competition.

By contrasting the features selected in the filtering step and those surviving after the competition, we can better comprehend the relative role of these selections. In Figure 7.6, panels (a) and (c) show the features selected in the filtering step for the Pareto-efficient individuals produced by the 20 independent runs of ASF-NEMODPS in rounds R2 and R3. The subset of representations surviving the competition is reported in panels (b) and (d) for the Pareto efficient individuals of the 20 runs. Green-colored bars correspond to features that are also included in the final Pareto front derived from the combination of all the independently optimized seeds.

In round R2, the filtering step identifies a larger number of features, with some clearly prevailing for the frequency in which they are selected, i.e., Cumulation and Anomalies at 28 and 51 day lead times. Interestingly, the competition-based selection does not yield the same relative frequencies: 62 and 75 day-ahead anomalies consistently overcome competition across multiple seeds, as well as in the final assembled Pareto front. In round R3, the frequency difference between filter- and competition-based selections is not as sharp as in round 2, but still notable.

These results seem to suggest that the filter-based selection frequency was not a reliable indicator of the representation that was eventually established

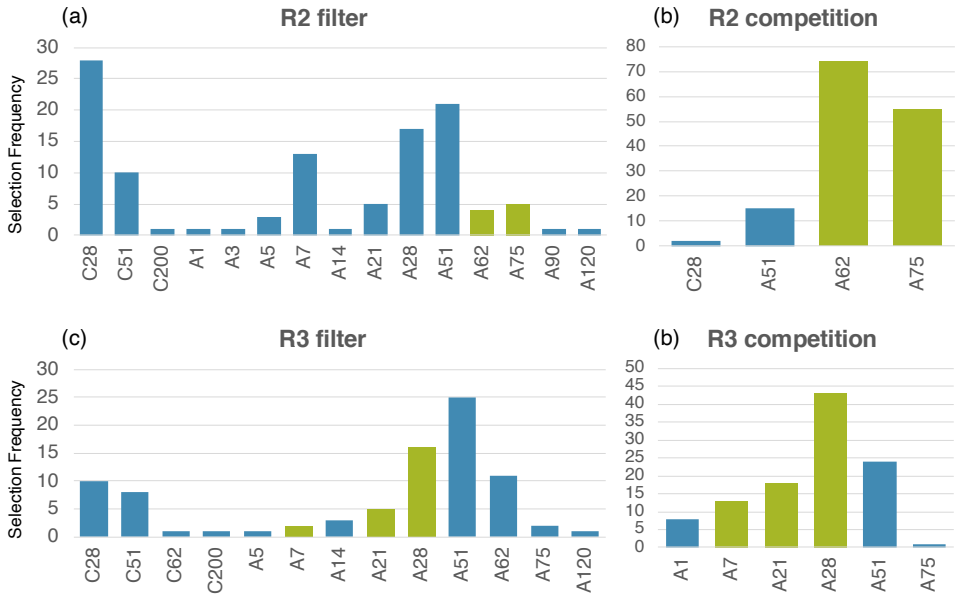

Figure 7.6: *Selection frequency of candidate variables in round R2 (panels (a) and (b)), and round R3 (panels (c) and (d)) during the filtering- and competition- based selection steps, respectively, for the 20 independent runs of AFS-NEMODPS. Green colored bars correspond to variables that also appear in the feature representations of the resulting Pareto front that assembles the results of the single runs.*

as efficient, raising questions around the trustworthiness of the filtering process. A first hypothesis is that the filtering step is at least helpful in reducing the pool of policy inputs to a smaller set of stronger candidates, thereby reducing the computational burden of a dataset-wide competition. The second hypothesis, however, is that the filtering process is completely blind to the competition skills of a variable, and is actually harmful to the feature representation learning as it might exclude good candidates. While presenting the results of this Chapter we operated under the first assumption, but we aim at testing the second hypothesis in upcoming experiments, where instead of a filter-based selection based on information gap, we will operate a randomized extraction of variables from the dataset of candidate policy inputs. Given the dataset dimension we will also increase the number of randomizes seeds and use larger NEMODPS populations.

8

Conclusions and future research

This thesis contributes to the literature of feature representation learning for multi-objective RL problems, specifically tackling the control of water resources systems, a challenging decision problem characterized by non-linearities, strong disturbances, possible alternative problem framings, and multiple objectives. By conjunctively addressing the fields of water resources systems, feature selection, and multi-objective control, we propose methodological innovations in automatic selection of information and policy architecture in complex decision problems. Novel methods are tested on real-world case studies of water resources conflicts in contentious river basins.

Across the six papers presented in this thesis, we advocate the need to more carefully consider the breadth of available information sources to inform reservoir control policies, especially in contexts where climate and societal change are exacerbating water conflicts. The existing literature of water resources control has rarely and sparely addressed the topic of selecting appropriate policy inputs, and this thesis aims at filling this gap by calling attention to this promising line of research, demonstrating its potential, and structuring the conversation around a proposed taxonomy comprising *a priori*, *a posteriori*, and *online* approaches.

Among the novel contributions of the thesis we include:

- the design and application of an unconventional feature representation centered around a basin-specific analysis of the drought state,

- the reconstruction of climate oscillations to expand the concept of feature representation to a problem of reservoir planning and filling,

- a meta-analysis on the role of variable processing in determining a feature's representations skills,

- a fully automated and tradeoff-dynamic framework for the online selection of policy architectures and inputs, targeting multisectoral conflict mitigation.

Among the most valuable outcomes of this collection of works, we highlight that i) learning an appropriate feature representation goes beyond the control policy performance improvement, but also generates insights on the learned task by identifying relevant policy drivers and system vulnerabilities. ii) The benefit of certain representations are especially noticeable in critical situations. For instance the benefits of including the drought index of Chapter 3, and the long term forecasts of Chapter 5 are mainly noticeable in dry years. iii) The value of hydrological forecasts in informing reservoir operations, a very common choice to enhance water system resilience, might be afflicted by modeling distorsions, and a less processed alternative could bring more benefit to the control policy (see Chapter 5). iv) The concepts of feature representations can be valuable beyond strictly control problems to applications of planning and filling (Chapter 4). Lastly, v) in Chapter 7, we see how guiding representation learning towards conflict mitigation determines a cross-sectoral information selection that tends to improve the performance in the weaker objective.

The common thread in these outcomes is that, in an operational context, the quest for an *optimal* feature representation might be an elusive concept. Instead of chasing optimality per se, we found several opportunities to mitigate recurrent water systems failures, that catalyze the largest concentrated impacts related to water management. As a consequence, our efforts were mostly concentrated at mitigating damages of drought emergencies, of critical phases in reservoir development (i.e., construction and filling), and social tensions deriving from conflicts between different users and their demands.

The here proposed contributions are intended to lay the foundation of several new promising research directions that can be further investigated. Among them,

- NEMODPS should go through an extensive diagnostic analysis that can shed more light on its potential. A direction of investigation could be the reliability and consistency of the policy architecture selection in a series of experiments targeted to test architectural convergence. For example, we could operate randomized restarts by extracting efficient individuals from an evolving population to be placed in a different randomly initialized population to investigate architectural convergence. Orders-of-magnitudes larger populations could also be used for the same purpose.

- An additional direction of investigation for NEMODPS can test its scalability to more complex control problem, for instance comprising multiple control decisions and many-objectives. A many-objective application of NEMODPS with 4 operating objectives is currently in the works.
- Lastly, new architectural and parametrical operators can be included in NEMODPS to enhance its effectiveness in exploring the architectural-parameteric hyperspace. Additional architectural operators can include new activation functions like ReLu, linear, or step functions, or the removal of existing nodes and connections. New parametrical operators can be included to target the investigation of a solution's proximity, or, alternatively, of unexplored regions of the parameters space (see, e.g., Hadka and Reed, 2013).
- In regards to feature representation learning, more challenging applications can be tested, including the control of multi-reservoir, and many-objectives systems. A more challenging application of AFS-NEMODPS is currently in the works. In particular, the framework is applied to the control of Gibe III dam in southern Ethiopia, considering three operating objectives and two control decisions, namely reservoir release and irrigation abstraction. Additionally, the dataset of candidate policy inputs is in this case represented by real forecasts developed by searching relevant anomalies in gridded global climatic datasets of sea-surface temperature, sea-level pressure and geopotential height. A novel strategy is proposed for situations in which water users' conflict persist even in the assumption of perfect future knowledge of the system disturbance.
- We plan to further investigate the efficacy of preliminary filtering promising candidate information to complement the evolutionary competition in determining efficient feature representations, as motivated in the discussion of Chapter 7.
- Unprecedented advances in monitoring technologies allow to characterize hydrological processes at finer spatial and temporal resolutions from multiple sources, e.g., ground observations, remote sensing, and crowdsourcing, offering new opportunities to enhance feature representation. As the volume of available information expands, however, new challenges emerge regarding how to extract valuable and compact signals in large heterogeneous datasets comprising a multiplicity of spatial and temporal resolutions, variables types, accuracies, and observability. To address these new challenges, water resources management can take advantage of state-of-the-art techiniques developed in the booming fields of big data analysis, image recognition, and computer vision in order to navigate such large and heterogeneous datasets extract relevant information.

- Feature representation in water resources management should not be considered a crystalized concept, but should dynamically adapt in response to variations in the ever evolving boundary conditions that coupled human-natural systems are exposed to. In particular, we should monitor variations in socio-economic drivers e.g., a water user experiencing unprecedented and more frequent failures; climatic drivers, i.e., an increased likelihood of extreme events, or modification in their nature; and physical drivers, e.g., when new water demands or infrastructures come online in the system. When one or more of these drivers change, the previous feature representation may not be adequate to represent the new system conditions, and should be updated accordingly. Yet, a critical challenge yet to address is to determine *when* and *how* to timely update the feature representation by means of appropriate triggers.

- Dynamicity in feature representation can be also introduced in the short term, by switching to alternative representations in response to specific system conditions. For instance, the appropriate representation to control a water system during a drought may give more space to key drought drivers that won't be needed in average or wet conditions. Therefore, instead of designing a single control policy with a single feature representation that is expected to be optimal in every possible hydrological conditions, a pool of alternative policies may be designed for a given hydrological condition (e.g., wet, normal, and dry), and activated in turns given the current or forecasted system state. Again, proper triggers should be defined to effectively and timely switch between policies, for instance basin-specific drougth indices in real time observation or in forecast mode, or teleconnection-based indices.

- Beyond single basin studies, a spatially explicit large scale comparative analysis of relevant policy drivers across different basins may uncover interesting trends in information relevance. Expected findings include the identification of correlation between relevant policy drivers and basin characteristics such as climate, basin size, morphology, and water uses.

- Lastly, more effort should be put into disseminating findings beyond the theoretical research domain. The potential synergies between theory and practice in the identification and retrieval of relevant information are evident: a well designed real-world monitoring system is essential to feature representation, as a feature can be recognized as valuable only if it is observed. At the same time, feature representation learning can guide the development and maintainance of monitoring systems by informing targeted investments towards the retrieval of most valuable information. Such virtuous collaboration between theory and practice can only be achieved

with a deliberate dissemination effort, and a continuous engagment of decision makers and stakeholders in contentious water systems.

Appendix A

Supporting Information for: When timing matters - misdesigned dam filling impacts hydropower sustainability

A.0.1 System Model

The conceptual model of the Omo-Turkana Basin comprises Gibe III reservoir (GIII), lake Turkana (T), and the Omo river stretch connecting them. A topological scheme representing the system is shown in Supplementary Figure A.4.

The dynamics of reservoir and lake storage s_t is modeled as:

$$s_{t+1}^{GIII} = s_t^{GIII} + q_{t+1}^{GIII} - e_t^{GIII} A_t^{GIII} - r_{t+1}^{GIII} \tag{A.1a}$$

$$s_{t+1}^{T} = s_t^{T} + r_{t+1-lag}^{GIII} + q_{t+1}^{lateral} + q_{t+1}^{Turkwel} + q_{t+1}^{Kerio} - e_t^{GIII} A_t^{GIII} \tag{A.1b}$$

where s_t^i is the storage to the i-th water body (i = GIII, T), e_t^i is the cyclostationary daily evaporation rate, and $A_t^i = a(s_t^i)$ is the reservoir surface estimated from s_t^i given the biunivocal surface-level-storage relation. In the adopted nota-

tion, the time subscript of a variable indicates the instant when its value is deterministically known. The release from Gibe III is defined as $r_{t+1}^{GIII} = f(s_t^{GIII}, u_t^{GIII}, q_{t+1}^{GIII})$ where f describes the nonlinear relation between the release decision u_t^{GIII} and the actual release Soncini-Sessa et al. (2007). The actual release at the end of the time interval is generally equal to the release decision unless physical constraints prohibit it (i.e., if the prescribed release lies outside the minimum and maximum allowable releases, for instance as a result of insufficient water volume in the reservoir, or exceedence of reservoir storage capacity). Lake Turkana, instead, is an endorheic lake, and the only water output is due to evaporation. The lake total inflow is given by the sum of GIbe III release (r_{t+1}^{GIII}), flow of Turkwel ($q_{t+1}^{Turkwel}$) and Kerio rivers (q_{t+1}^{Kerio}), and the lateral contributions in the lower Omo valley ($q_{t+1}^{lateral}$). According to the daily time-step adopted in the model, the river reaches are modelled as plug-flow canals in which the velocity and direction of flow are constant. A transit lag time of $lag = 18$ days is estimated from streamflow data, and corresponds to the average time employed by the water to transit between these two points.

Filling and regime policies are optimized with respect to the following objective functions, defined over the filling horizon HF from January 2015 to October 2018:

- Hydropower Production J^{Hyd} (to be maximized):

$$J^{Hyd} = \frac{1}{Nyears}\left[\sum_{t=1}^{HF} g_t^{Hyd}\right] \tag{A.2a}$$

where

$$g_t^{Hyd} = \eta\sigma\gamma\bar{h}_t^{GIII} q_t^{turb}; \tag{A.2b}$$

with $Nyears$ being the number of years in the simulation horizon, η the turbine efficiency, σ the gravitational acceleration, γ the water density, $\bar{h}_t^{GIII}$ the net hydraulic head, and q_t^{turb} the turbinated flow.

- Environmental deviation J^{Env} (to be minimized):

$$J^{Env} = \frac{1}{HF}\sum_{t=0}^{HF-1} g_t^{Env} \tag{A.3a}$$

where

$$g_t^{Env} = \left(q_t^{natural} - q_t^{delta}\right)^2 \tag{A.3b}$$

with $q_t^{natural}$ representing the cyclostationary trajectory of natural inflow regime in the Omo Delta estimated from streamflow data prior Gibe

III construction, and q_t^{delta} the streamflow reaching the delta for the policy under evaluation. The objective is formulated as the average daily squared distance between these two trajectories over the simulation horizon, and is aimed at preserving natural flow conditions in correspondence to the Omo delta, and consequently a natural inflow pattern in lake Turkana. According to the adopted squared formulation, smaller, and more frequent deviations are preferred to large concentrated deviations with respect to the target.

- Final Gibe III level J^{Fill} (to be maximized and considered in the optimization of filling policies only):

$$J^{Fill} = h_H F^{GIII} \tag{A.4}$$

where h_H^{GIII} is the reservoir's level at the end of the filling horizon HF.

In the investigation of the role of filling timing, we simulated the first two years of the reconstructed filling strategy assuming to start the reservoir filling in different years. Historically, 2 years were necessary to achieve the normal operating level and, subsequently, the reservoir level remained reasonably constant apart from contained seasonal fluctuations.

To assess the resulting system performance, we used 4 evaluation indicators assessed in a 2-years filling horizon H_{2y} formulated as follows.

1. Annual average hydropower production (i^{Hyd}):

$$i^{Hyd} = \frac{1}{Nyears}\left[\sum_{t=1}^{\mathbb{H}} \eta g \gamma \bar{h}_t^{GIII} q_{t+1}^{turb}\right] \tag{A.5}$$

where $Nyears$=2 is the number of years in the simulation horizon, η=0.34 is the turbine efficiency, g = 9.81 m/s^2 is the gravitational acceleration, γ = 1000 kg/m^3 is the water density, $\bar{h}_t^{GIII}$ is the net hydraulic head, and q_t^{turb} is the turbinated flow.

2. Final Gibe III level (i^{GIII_L}), evaluated at the end of the filling transient H_{2y} in meters above sea level [masl]:

$$i^{GIII\_L} = h_{\mathbb{H}}^{GIII} \tag{A.6}$$

3. Final Turkana level drop (i^{T_Ld}), evaluated in terms of absolute level drop referred to the initial lake level h_0^T:

$$i^{T\_Ld} = h_{\mathbb{H}}^{T} - h_0^T \tag{A.7}$$

4. Flood Pulse (i^{FP}), defined as the average annual flood pulse magnitude computed as the maximum inflow reaching the delta during the flood season of August-September:

$$i^{FP} = \frac{1}{Nyears}\left[\sum_{y=1}^{Nyears} \max_{\tau\in[Aug,Sept]} q_{\tau,y}^{delta}\right] \tag{A.8}$$

where he flow in the Omo delta is given by $q_{t+1}^{delta} = r_{t+1-lag}^{GIII} + q_{t+1}^{lateral}$.

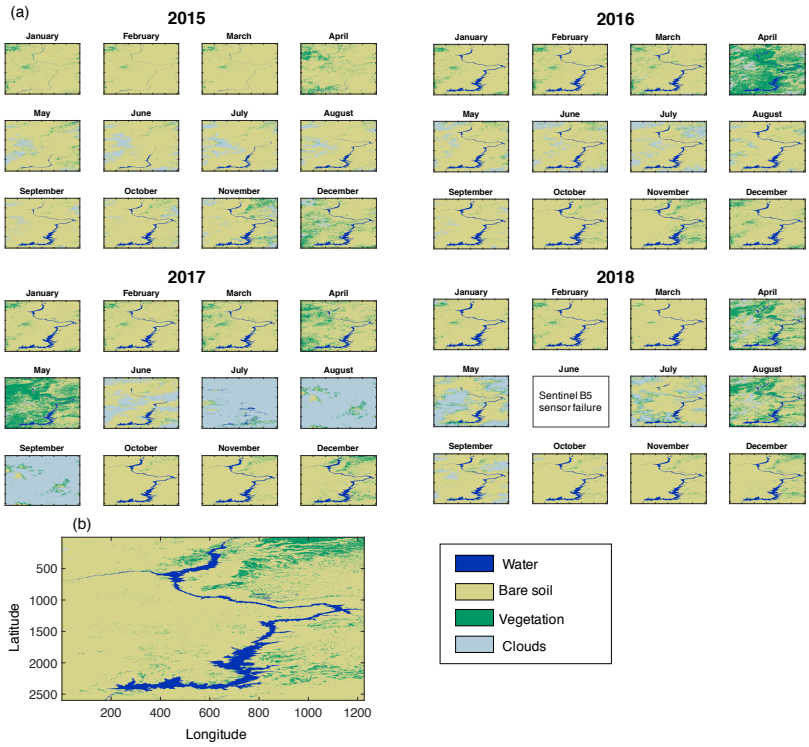

Figure A.1: *Classification of water, vegetation, bare soil, and clouds at Gibe III reservoir location from Sentinel 2 satellite imagery. For the months of June to September 2017, and May to July 2018, the cloud cover was too persistent to allow composing a cloud free image; the values of adjacent months were thus interpolated. Additionally, a sensor failure temporarily interrupted Sentinel data collection via B5 sensor in the month of June 2018.*

Signal	Crossvalidation Accuracy		Significance Score (%)		Location	Period
	negative	positive	negative	positive		
AMM	0.64	0.81	88.3	84.2	Atlantic	5-10 years
AMO	0.62	0.67	99.0	71.5	Atlantic	Multi-decadal
NTA	0.58	0.57	97.4	67.0	Atlantic	2-4 years
CAR	0.57	0.56	96.2	82.4	Atlantic	5-10 years
NAO	0.71	0.47	95.7	71.0	Atlantic	2-4 years
BEST	0.17	0.46	91.7	73.9	Pacific	2-4 years
MEI	0.59	0.47	63.8	91.3	Pacific	2-4 years
WP	0.39	0.61	33.3	99.9	Pacific	1-3 years
PMM	0.52	0.41	91.0	90.3	Pacific	Decadal
PNA	0.58	0.53	98.8	60.7	Pacific	2-5years
PDO	0.55	0.27	88.4	88.8	Pacific	Decadal
EPNP	0.31	0.63	66.8	94.2	Pacific	3-6 years
NP	0.44	0.52	93.1	93.6	Pacific	3-6 years
WIO	0.62	0.52	99.9	81.3	Indian	2-5years
SEIO	0.80	0.51	98.7	55.2	Indian	Decadal
DMI	0.61	0.55	98.5	97.0	Indian	2-5 years

Table A.1: *Accuracy of the phase specific, univariate linear forecast models and associated significance score for the 16 tested climate signals. Model accuracy in crossvalidation is measured via the Pearson correlation coefficient. The Significance Score corresponds to the percentage of Montecarlo random shuffling trials that identify a smaller number of significantly correlated SST grid points with respect to unshuffled data.*

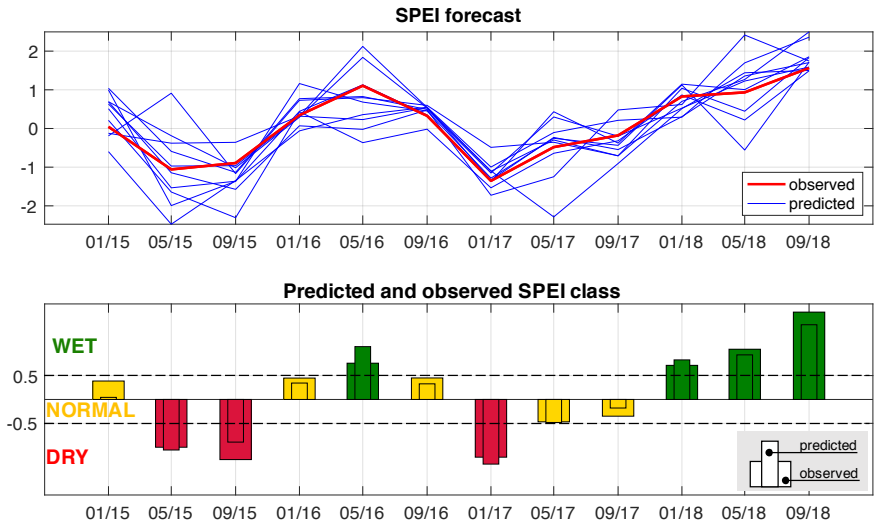

Figure A.2: *SPEI forecast. Top panel reports the ensemble of teleconnection-based seasonal forecast for the Standardized Precipitation and Evaporation Index. The ensemble average is then used to classify the SPEI forecast in dry, normal, and wet conditions; this classification is correct for all the seasons.*

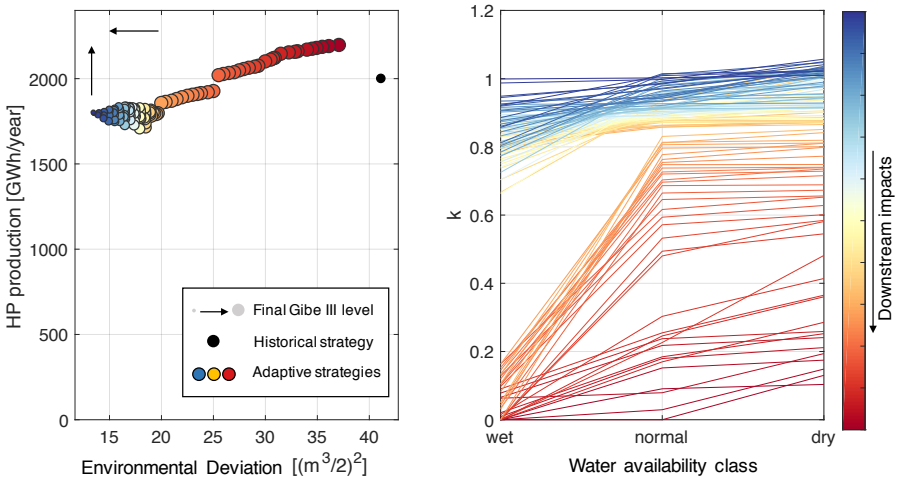

Figure A.3: *Optimal adaptive filling policies. Left panel reports policies performance in the objective space, in terms of Hydropower Production (vertical axis, to be maximized), Downstream Impacts (horizontal axis, to be minimized), and Final Gibe III level (circle size, to be maximized). The historical policy (black circle) attains a comparable performance with respect to the designed policies in terms of hydropower production and Gibe III final level, but is associated to the highest downstream impacts demonstrating that adaptive policies have the potential to contain downstream alterations without impacting filling efficiency. The right panel reports the optimal values of the scaling factors associated to the three SPEI classes, which are lower for wet seasons than for dry seasons, indicating that the adaptive filling policies will release a larger water volumes when a dryer than average season is expected to avoid magnifying drought impacts on downstream activities, and will impound a larger fraction of inflows in case of a wet spell as more abundant basin wide precipitations can support downstream activities.*

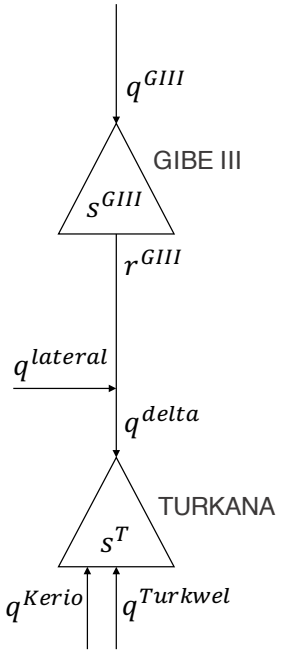

Figure A.4: *Topological Scheme of the OTB: comprising Gibe III reservoir, lake Turkana, the Omo river stretch connecting the two water bodies, and the external inflows.*

Appendix B

Supporting Information for: Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations

B.1 Lake Como integrated model and optimal control problem formulation

The Lake Como integrated model is composed of three main components as illustrated in Figure B.1 (further details can be found in (Giuliani et al., 2016d)):

- Catchment model - a conceptual hydrologic model Hydrologiska Byrans Vattenbalansavdelning (HBV), which relies on four storage units, one for snow and the other three for different soil layers.
- Lake Como operational model - the lake dynamics is described by a mass-

balance equation assuming a modelling and decision-making time step of 24 hours, where the lake releases are determined by the lake operating policy. According to the daily time step, the Adda River can be described by a plug-flow model to simulate the routing of the lake releases from the lake outlet to the intake of the irrigation canals. This diversion of the water from the Adda River into the irrigation canal is regulated by the water rights of the agricultural districts.

- Agricultural districts model - the dynamic processes internal to the irrigation districts are described by three distinct modules devoted to specific tasks: (i) a distributed-parameter water balance module (Facchi et al., 2004) simulating water sources, conveyance, distribution, and soil-crop water balance over a regular mesh of cells with a side length of 250 m. Each individual cell identifies a soil volume which is subdivided into two layers, where the upper one (evaporative layer) represents the upper 15 cm of the soil and the bottom one (transpirative layer) represents the root zone and has a time-varying depth, with the water percolating out of the bottom layer that constitutes the recharge to the groundwater system; (ii) a heat unit module (Neitsch et al., 2011) simulating the sequence of growth stages (e.g., root length, basal coefficient, leaf area index), including the sowing and harvesting dates, as a function of the temperature in terms of cumulated heat units; (iii) a crop yield module that first estimates the maximum yield achievable in optimal conditions and, then, reduces it to take into account the stresses due to insufficient water supply from rainfall and irrigation happened during the agricultural season, where the yield response to water stresses is estimated according to the empirical function proposed in the AquaCrop model (Steduto et al., 2009) and based on the approach proposed by FAO (Doorenbos and Kassam, 1979). Finally, the yield of the cultivated crops is used to estimate the farmers net profit, which also depends on the crop price and cost along with the subsidies derived from the EU's Common Agricultural Policy (Gandolfi et al., 2014).

In particular, the mass balance equation describing the regulated Lake Como dynamics in the operational model is formulated as follows:

$$s_{t+1} = s_t + n_{t+1} - r_{t+1} \tag{B.1}$$

where s_t is the lake storage [m^3], while n_{t+1} and r_{t+1} are the net inflow (i.e., inflow minus evaporation losses) and the outflow volumes in the time interval $[t, t+1)$, respectively. The release is defined as $r_{t+1} = f(s_t, u_t, n_{t+1})$, where $f(\cdot)$ describes the nonlinear, stochastic relation between the release decision determined by the operating policy (i.e., $u_t = p(\cdot)$), and the actual release r_{t+1} (Piccardi and Soncini-Sessa, 1991).

The lake operations is driven by multiple objectives, primarily flood control along the lake shorelines, water supply to four irrigation districts, and environmental protection of ecosystems in the Adda River downstream from the lake. This latter is accounted for by setting a minimum environmental flow constraint on the lake releases q^{MEF} = 5 m^3/s. On the basis of previous works (e.g., Castelletti et al., 2010a; Giuliani and Castelletti, 2016; Giuliani et al., 2016a; Denaro et al., 2017b), we formulated the flood control objective and the water supply objective as follows:

- *Flood control*: the average annual number of flooding days in the simulation horizon H, defined as days when the lake level h_t is higher than the flooding threshold $\bar{h}$=1.24 m:

$$J^F = \frac{1}{H/365} \sum_{t=1}^{H} \Lambda(h_t) \quad \text{where}$$
$$\Lambda(h_t) = \begin{cases} 1 & \text{if } h_t > \bar{h} \\ 0 & \text{otherwise} \end{cases} \tag{B.2}$$

- *Water supply deficit*: the daily average quadratic water deficit between the lake release r_{t+1} and the total water demand w_t of the downstream system, subject to the minimum environmental flow constraint q^{MEF}. This quadratic formulation aims to penalize severe deficits in a single time step, while allowing for more frequent, small shortages (Hashimoto et al., 1982):

$$J^D = \frac{1}{H} \sum_{t=1}^{H} \beta_t \times \max\left(w_t - \max(r_{t+1} - q^{MEF}, 0), 0\right)^2 \tag{B.3}$$

  where β_t is a time-varying parameter which depends on the phenological phases of the crops and is higher after germination to the beginning of phenological maturity.

According to the Evolutionary Multi-Objective Direct Policy Search method (Giuliani et al., 2016b, 2018), the operating policy is formulated as a network of Gaussian radial basis functions (RBFs) and the release decision is defined as

$$u_t = \sum_{i=1}^{N} w_i \varphi_i(\mathcal{I}_t) + \alpha \tag{B.4}$$

where N is the number of RBFs $\varphi(\cdot)$, w_i is the non-negative weight of the i-th RBF (i.e., $w_i \geqslant 0 \quad \forall i$), and α is a constant parameter. The single RBF is

defined as follows:

$$\varphi_i(\mathcal{I}_t) = \exp\left[-\sum_{j=1}^{M} \frac{((\mathcal{I}_t)_j - c_{j,i})^2}{b_{j,i}^2}\right] \tag{B.5}$$

where M is the number of policy inputs ($\mathcal{I}_t$) and $\mathbf{c}_i$, $\mathbf{b}_i$ are the M-dimensional center and radius vectors of the i-th RBF, respectively. The policy parameter vector $\boldsymbol{\theta}$ is therefore defined as $\boldsymbol{\theta} = \left[(c_1, \ldots, c_M)_1^N, (b_1, \ldots, b_M)_1^N, w_1^N, \alpha\right]$. The total number of policy parameters (i.e., decision variables) is therefore equal to $N(2M+1)+1 = 29$.

The optimal values of the policy parameters $\boldsymbol{\theta}^*$ are determined by solving the following optimization problem:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \mathbf{J}(\boldsymbol{\theta}) \tag{B.6}$$

where the decision variables are the policy parameters $\boldsymbol{\theta} \in \Theta$, the objective functions are the two operating objectives $\mathbf{J}$ defined in eqs. B.2-B.3 and the problem is constrained by the dynamics of the system (eq. B.1). Finding $\boldsymbol{\theta}^*$ is equivalent to finding the corresponding optimal RBF policy p_θ^* (as in eq. 3 of the main paper).

B.2 The Nino Index Phase Analysis framework

An overview of the Nino Index Phase Analysis (NIPA) framework originally proposed by Zimmerman et al. (2016) for predicting the spring precipitation over the Lower Colorado River Basin on the basis of the ENSO state is provided in Figure B.2.

B.3 Comparison of seasonal meteorological forecast models

The comparison of alternative empirical, data-driven forecast models for step 2 of the Climate State Intelligence framework is reported in Table B.1.

B.4 The Information Selection and Assessment framework

An overview of the Information Selection and Assessment (ISA) framework originally proposed by Giuliani et al. (2015) for selecting the most valuable information to improve water systems operations is provided in Figure B.3. The

Table B.1: *Coefficient of determination R2 between predicted and observed winter precipitation P_{JFM} for different seasonal meteorological forecast models in leave-one-out cross-validation.*

Model	Neurons	R2
Linear	-	0.3922
Artificial Neural Network	3	0.6043
	4	0.6332
	5	0.6334
Extreme Learning Machine	5	0.8143
	7	0.8174
	8	0.8179
	9	0.8093
	10	0.8374
	15	0.8274
	21	0.8005

first two steps of ISA for the Lake Como system, namely the quantification of the Expected Value of Perfect Information and the Information Selection in terms of best inflow forecast lead time, are described in Denaro et al. (2017b). Building on this work, we focus the Climate State Intelligence experiments on hydrologic forecasts cumulated over a lead-time of 51 days. Among the different metrics introduced in the ISA framework to quantify the value of the selected information (see Figure B.4), in this work we focus on the hypervolume Indicator that captures both the convergence and diversity of the baseline or informed operating policies with respect to the best known Pareto front (i.e., the set of Perfect Operating Policies).

B.5 Detection of teleconnections for multiple climate signals

The results of the preliminary teleconnection detection analysis for multiple climate signals (step 1 of the Climate State Intelligence framework) are reported in Table B.2. These results suggest that ENSO and NAO are the climate oscillations that provide most skilful precipitation forecasts in the Lake Como basin.

Table B.2: *Pearson correlation coefficients between predicted and observed winter precipitation* P_{JFM} *with associated level of confidence for each phase of five different climate signals (predictions depend on preseason SST anomalies).*

Climate Signal	Phase	Pearson correlation	Level of confidence
NAO	negative	0.50	89.00%
	positive	0.52	69.70%
ENSO	negative	0.41	32.00%
	positive	0.51	96.20%
PDO	negative	0.50	67.50%
	positive	-0.02	51.00%
AMO	negative	0.32	19.90%
	positive	0.28	35.00%
IOD	negative	0.51	32.30%
	positive	0.44	97.1%

B.6 Detection of ENSO and NAO teleconnections for all seasons

The results of step 1 of of the Climate State Intelligence framework for all 3-month periods are reported in Table B.3.

B.7 Validation of the NIPA detection of climate teleconnections

In order to get a validation of the correlations patterns presented in Figure 3, we performed an independent test using a different precipitation dataset and focusing on the relationship between JFM precipitation and ENSO, which corresponds to the results illustrated in panel (a) of the figure. The mean value and the standard deviation of the MEI index over the 1950-2009 time period are used to classify each year as positive, negative, or neutral. A year is considered "positive" if the November-December MEI index is greater than the mean MEI value + 0.5 standard deviation; "negative" if the November-December index is lower than the mean - 0.5 standard deviations; "neutral" otherwise. The precipitation dataset considered is the GPCC (Global Precipitation Climatology Centre) Full Data Monthly Product Version 2018 at 0.5 degree resolution (Schneider et al.), from which we constructed the mean JFM precipitation time series over the basin for the same period 1951-2009. The Lake Como basin is defined selecting the six grid points covering the portion of the Alpine region

Table B.3: *Pearson correlation coefficients between predicted and observed seasonal precipitation (all 3-month periods) for each phase of NAO and ENSO.*

Signal	**ENSO**		**NAO**	
Phase	positive	negative	positive	negative
JFM	-0.69	-0.6	-0.64	-0.64
FMA	0.72	0.79	0.55	-0.74
MAM	0.41	0.45	0.65	0.7
AMJ	0.67	-0.65	-0.69	0.82
MJJ	0.61	0.42	0.58	0.69
JJA	0.65	0.64	-0.61	0.59
JAS	0.75	0.66	-0.51	-0.56
ASO	0.76	-0.67	0.57	0.64
SON	-0.63	-0.58	-0.53	-0.6
OND	0.61	0.65	0.63	nan
NDJ	-0.62	0.65	0.61	0.61
DJF	0.68	0.58	0.51	0.72

the Como Lake belongs to (i.e., gridpoints in the box lat N 45.5 - 46.5; lon E 8.5 - 10.0). Global monthly SST anomalies are taken from the NOAA Extended Reconstructed SST (ERSST) V5 dataset.

Figure B.5 reports the correlation pattern between JFM precipitation anomaly over the Como Lake region and SST anomalies in the previous months, at different time lags, for the positive ENSO years. Filled contours mean that the statistical significance is >95%, non-filled contours mean significance >70%. In each panel a different three-monthly mean of SSTs is considered, starting from FMA of the previous year in the top-left panel. Moving from left to right and from top to bottom, the three-months period is shifted one month forward in time (i.e., FMA, MAM, AMJ, etc.), thus diminishing the time lag between SST and precipitation anomalies. The correlation between OND SSTs and JFM precipitations anomalies, which is thus equivalent to the top-right panel of Figure 3 in the paper, is shown in panel (i). These results show how the OND situation is just a "snapshot" of the evolution of persistent anomalies of the climate state over the months which are correlated with the precipitation in the Lake Como basin. These correlations are high and significant over a large portion of the Central Pacific particularly in spring and summer, both in the tropics and at midlatitudes, and in both hemispheres. During the fall, correlations in the tropical Pacific are less evident but are still present, while a signal appears in the Eastern North Atlantic. Correlations over the Central Tropical Pacific and in the Eastern North Atlantic persist during winter and appear as a signal of an

existing teleconnection between the Pacific and the Atlantic sector driven by the atmosphere, which influences the European hydroclimatic conditions ultimately impacting on the Lake Como basin.

B.8 Cross-validation of seasonal precipitation forecast

The seasonal precipitation forecast model constructed in the second step of the Climate State Intelligence framework is cross-validated using the leave-one-out scheme. The model predictions obtained over the full dataset (in calibration only) are illustrated in Figure 4 of the paper and attain a Pearson correlation coefficient between observed and predicted values equal to 0.91. The scatter plot with the cross-validated results is shown in Figure B.6; in this case, the Person correlation coefficient between observed and predicted values decreases to 0.81.

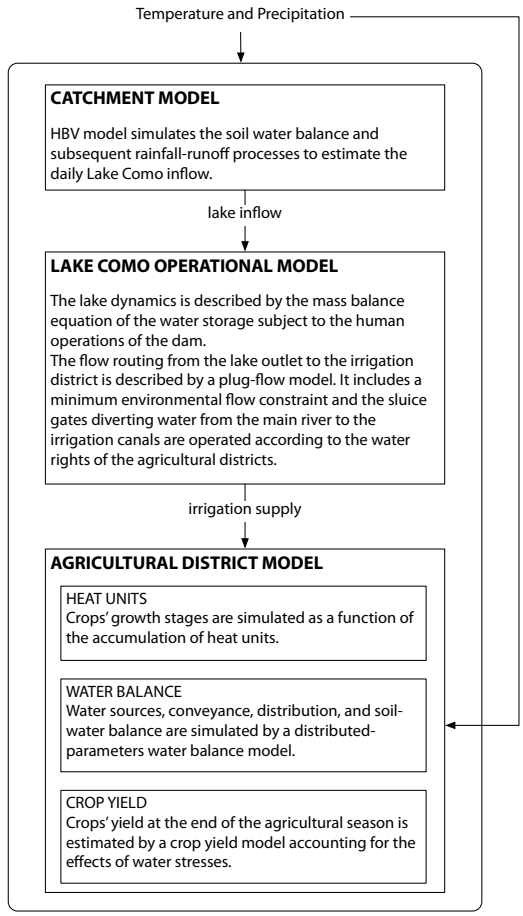

Figure B.1: *Schematic representation of the integrated simulation model of the Lake Como basin.*

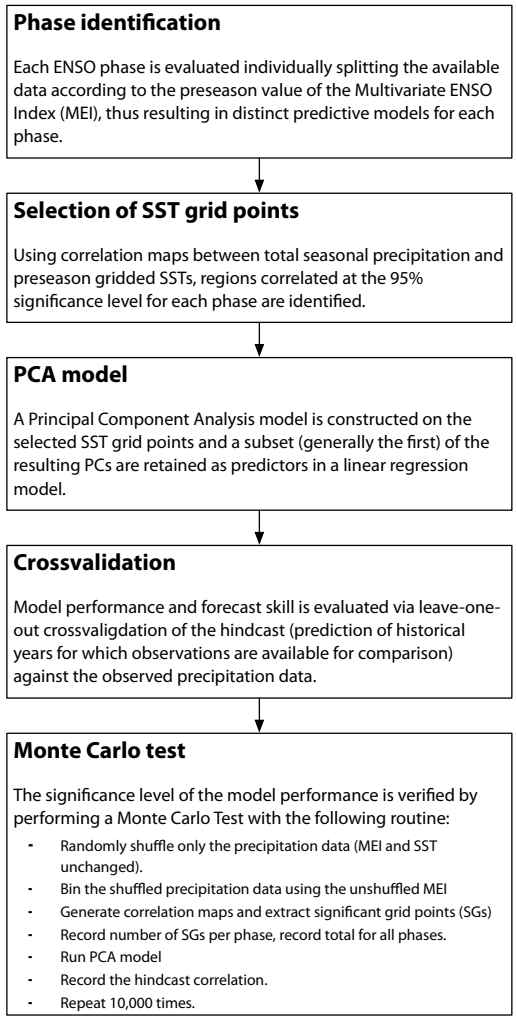

Figure B.2: *Illustration of the Nino Index Phase Analysis framework (Zimmerman et al., 2016).*

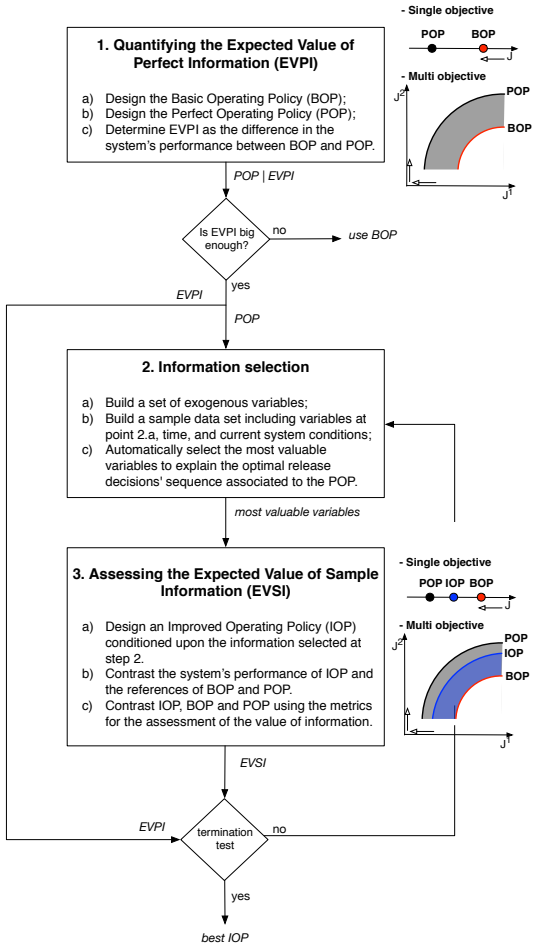

Figure B.3: *Illustration of the Information Selection and Assessment framework (Giuliani et al., 2015).*

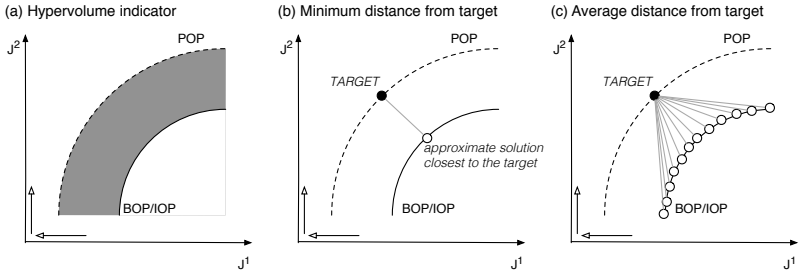

Figure B.4: *Illustration of the ISA metrics for assessing the value of information in multiobjective problems (Giuliani et al., 2015).*

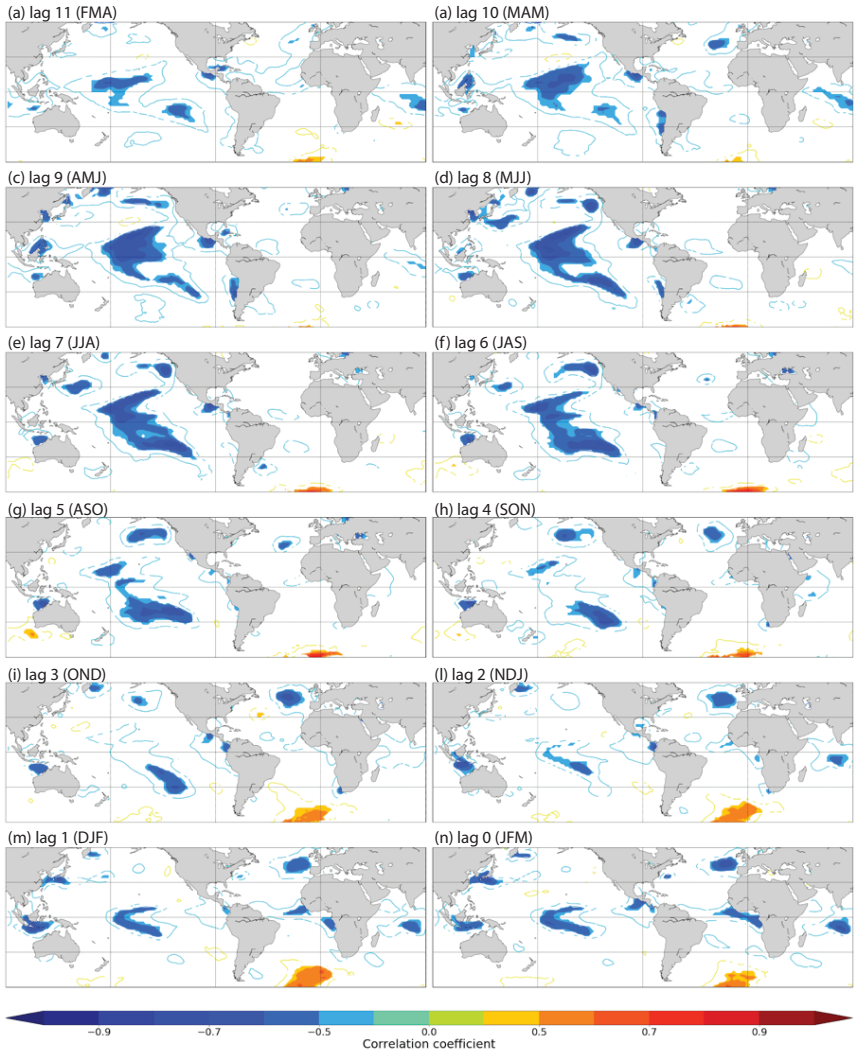

Figure B.5: *Correlation maps between JFM precipitation anomaly over the Como Lake region and SST anomalies for different lag-times during the ENSO positive phase, with the correlation between OND SSTs and JFM precipitations anomalies, which is thus equivalent to the top-right panel of Figure 3 in the paper, shown in panel (i).*

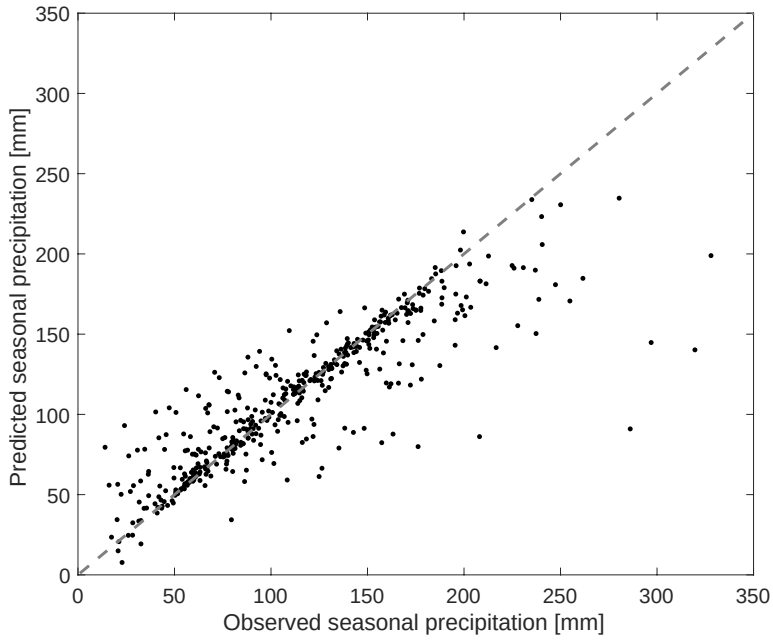

Figure B.6: *Scatterplot between observed and predicted seasonal precipitation in leave-one-out crossvalidation.*

Appendix C

Supporting Information for: Neuro-Evolutionary Direct Policy Search for Multi-Objective Optimal Control

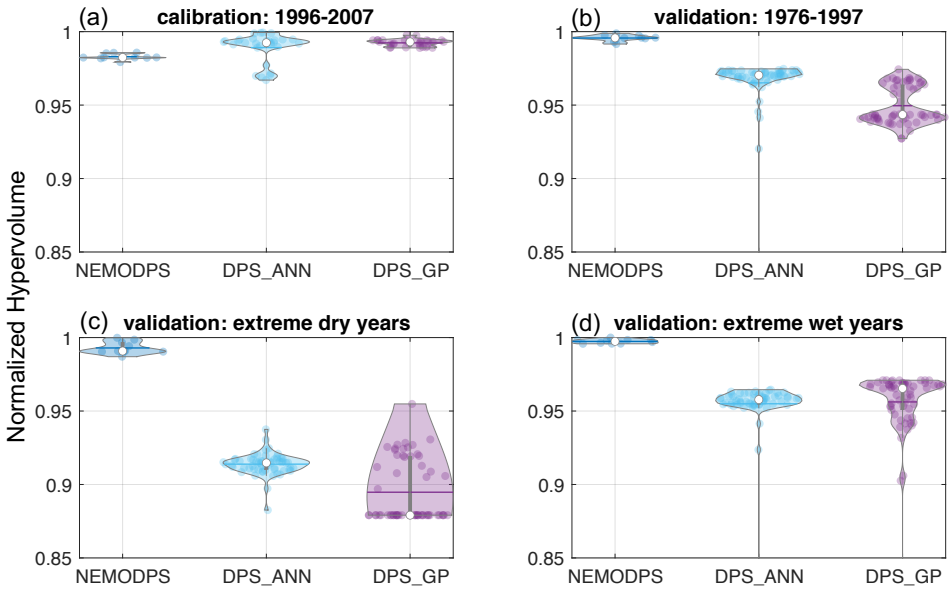

Figure C.1: *Violin plot showing the distribution of hypervolumes for different seeds in NEMODPS and DPS methods under calibration dataset (panel a) and validation datasets (panels b-d). Within each violin-shaped distribution, the white dot indicates the median, and the solid horizontal line the mean of the distribution. NEAT was not included in this analysis as single objective algorithms produce a single solution for each seed, preventing the single-seed hypervolume computation. NEMODPS hypervolume distributions are shown for 10 independent algorithmic runs, while DPS distribution comprise 10 runs for each predefined architecture (from 1 to 6 nodes) evaluated independently for the two activation functions considered (Sigmoidal for DPS-ANN and Gaussian for DPS-GP). NEMODPS distributions result much less variable than DPS across all the considered datasets, and especially for the three validation datasets, indicating higher consistency of NEMODPS solutions across independent random seeds.*

Appendix D

Supporting Information for: Dynamic retrieval of informative inputs for multi-sector reservoir policy design with diverse spatio-temporal objective scales

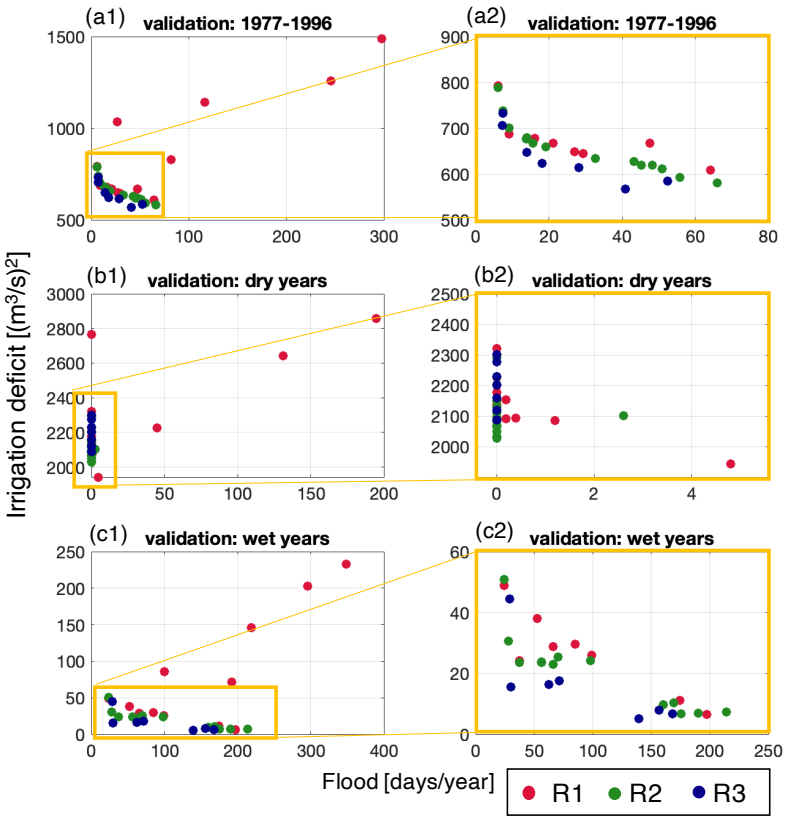

Figure D.1: *Validation of optimal policies for the three rounds of AFS-NEMODPS for a 20-year evaluation horizon (panel a1 and a2), and two 5-year evaluation horizons composed of extreme dry (panels b1 and b2) and wet years (panels c1 and c2). Panels a1, b1, and c1 show the entire objective space obtained for the policy re-evaluation: while most solutions are located in a limited area of the objective space (yellow boxes), few solutions produced in round R1 appear very distant, and located in largely inefficient areas of the objective space. Panel a2, b2, and c2 expand the yellow box to better compare the competitive solutions produced by the three rounds. The most informed round R3 consistently outperforms the other two in the 1977-1996 and wet-years datasets. In the dry years dataset, some R2 solutions achieve slightly lower irrigation deficit compared to R3, but with a fairly negligible difference. This analysis shows that a minimally informed policy, i.e., the one produced in R1, can incur in severe performance degradation when tested on new hydrological conditions that differ from the dataset used for its calibration. The addition of information in round R2 and R3 demonstrate the potential to greatly enhance the robustness of control policies across highly diverse hydrological conditions.*

Bibliography

Abdolmaleki, A., Lau, N., Reis, L.P., Peters, J., Neumann, G., 2016. Contextual policy search for linear and nonlinear generalization of a humanoid walking controller. Journal of Intelligent & Robotic Systems 83, 393–408.

Abolpour, B., Javan, M., Karamouz, M., 2007. Water allocation improvement in river basin using adaptive neural fuzzy reinforcement learning approach. Applied soft computing 7, 265–285.

AghaKouchak, A., 2015a. A multivariate approach for persistence-based drought prediction: Application to the 2010–2011 east africa drought. Journal of Hydrology 526, 127–135.

AghaKouchak, A., 2015b. Recognize anthropogenic drought. Nature 524, 409.

Akrour, R., Schoenauer, M., Sebag, M., 2012. April: Active preference learning-based reinforcement learning, in: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer. pp. 116–131.

Alcamo, J., Flörke, M., Märker, M., 2007. Future long-term changes in global water resources driven by socio-economic and climatic changes. Hydrological Sciences Journal 52, 247–275.

Alvernaz, S., Togelius, J., 2017. Autoencoder-augmented neuroevolution for visual doom playing, in: 2017 IEEE Conference on Computational Intelligence and Games (CIG), IEEE. pp. 1–8.

Andreu, J., Capilla, J., Sanchís, E., 1996. Aquatool, a generalized decision-support system for water-resources planning and operational management. Journal of hydrology 177, 269–291.

Andreu, J., Ferrer-Polo, J., Pérez, M., Solera, A., 2009. Decision support system for drought planning and management in the jucar river basin, spain, in: 18th World IMACS/MODSIM Congress, Cairns, Australia.

Anghileri, D., Pianosi, F., Soncini-Sessa, R., 2011. A framework for the quantitative assessment of climate change impacts on water-related activities at the basin scale. Hydrology and Earth System Sciences 15, 2025.

Anghileri, D., Voisin, N., Castelletti, A., Pianosi, F., Nijssen, B., Lettenmaier, D.P., 2016. Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments. Water Resources Research 52, 4209–4225.

Arnal, L., Cloke, H., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., Krzeminski, B., Pappenberger, F., 2018. Skilful seasonal forecasts of streamflow over Europe? Hydrology and Earth System Sciences 22, 2057–2072.

Asress, M.B., Simonovic, A., Komarov, D., Stupar, S., 2013. Wind energy resource development in ethiopia as an alternative energy future beyond the dominant hydropower. Renewable and Sustainable Energy Reviews 23, 366–378.

Assael, J.A.M., Wahlström, N., Schön, T.B., Deisenroth, M.P., 2015. Data-efficient learning of feedback policies from image pixels using deep dynamical models. arXiv preprint arXiv:1510.02173 .

Avery, S., 2013. What future for lake turkana. African Studies Centre, Oxford: University of Oxford .

Avery, S., 2017. Fears over ethiopian dams' costly impact on environment, people. URL: https://theconversation.com/fears-over-ethiopian-dams-costly-im-pact-on-environment-people-80757.

Avery, S., Eng, C., 2012. Lake turkana & the lower omo: hydrological impacts of major dam and irrigation developments. African Studies Centre, the University of Oxford .

Avery, S.T., Tebbs, E.J., 2018. Lake turkana, major omo river developments, associated hydrological cycle change and consequent lake physical and ecological change. Journal of Great Lakes research 44, 1164–1182.

Baglietto, M., Cervellera, C., Sanguineti, M., Zoppoli, R., 2010. Management of water resource systems in the presence of uncertainties by nonlinear approximation techniques and deterministic sampling. Computational Optimization and Applications 47, 349–376.

Bartolini, E., Claps, P., D'odorico, P., 2009. Interannual variability of winter precipitation in the european alps: relations with the north atlantic oscillation. Hydrology and Earth System Sciences 13, 17–25.

Basirat, M., Roth, P.M., 2018. The quest for the golden activation function. arXiv preprint arXiv:1808.00783 .

BBC, 2019. The 'water war' brewing over the new river nile dam. URL: https://www.bbc.com/news/world-africa-43170408.

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

Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence 35, 1798–1828.

Beniston, M., 1997. Variations of snow depth and duration in the swiss alps over the last 50 years: links to changes in large-scale climatic forcings, in: Climatic change at high elevation sites, pp. 49–68.

Benson, R., 2016. Reviewing reservoir operations: can federal water projects adapt to change. Columbia Journal of Environmental Law 42.

Bertoni, F., Giuliani, M., Castelletti, A., 2017. Scenario-based fitted q-iteration for adaptive control of water reservoir systems under uncertainty. IFAC-PapersOnLine 50, 3183–3188.

Bertsekas, D.P., 2019. Reinforcement learning and optimal control. Athena Scientific .

Bianchini, M., Scarselli, F., 2014. On the complexity of neural network classifiers: A comparison between shallow and deep architectures. IEEE transactions on neural networks and learning systems 25, 1553–1565.

Block, P., 2011. Tailoring seasonal climate forecasts for hydropower operations. Hydrology and Earth System Sciences 10, 15–1355.

Block, P., Goddard, L., 2012. Statistical and dynamical climate predictions to guide water resources in ethiopia. Journal of Water Resources Planning and Management 138, 287–298.

Block, P., Rajagopalan, B., 2007. Interannual variability and ensemble forecast of upper blue nile basin kiremt season precipitation. Journal of Hydrometeorology 8, 327–343.

Block, S., Webb, P., 2001. The dynamics of livelihood diversification in post-famine ethiopia. Food policy 26, 333–350.

Böhmer, W., Springenberg, J.T., Boedecker, J., Riedmiller, M., Obermayer, K., 2015. Autonomous learning of state representations for control: An emerging field aims to autonomously learn state representations for reinforcement learning agents from their real-world sensor observations. KI-Künstliche Intelligenz 29, 353–362.

Bowden, G.J., Dandy, G.C., Maier, H.R., 2005. Input determination for neural network models in water resources applications. Part 1 - Background and methodology. Journal of Hydrology 301, 75–92. doi:doi: 10.1016/j.jhydrol.2004.06.021.

Brandimarte, L., Baldassarre, G.D., Bruni, G., D?Odorico, P., Montanari, A., 2011. Relation between the north-atlantic oscillation and hydroclimatic conditions in mediterranean areas. Water Resources Management 25, 1269–1279.

de Broissia, A.d.F., Sigaud, O., 2016. Actor-critic versus direct policy search: a comparison based on sample complexity. arXiv preprint arXiv:1606.09152 .

de Bruin, T., Kober, J., Tuyls, K., Babuška, R., 2018. Integrating state representation learning into deep reinforcement learning. IEEE Robotics and Automation Letters 3, 1394–1401.

Bu, L., Babu, R., De Schutter, B., et al., 2008. A comprehensive survey of multiagent reinforcement learning. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 38, 156–172.

Busa-Fekete, R., Szörényi, B., Weng, P., Cheng, W., Hüllermeier, E., 2014. Preference-based reinforcement learning: evolutionary direct policy search using a preference-based racing algorithm. Machine Learning 97, 327–351.

Busoniu, L., Ernst, D., De Schutter, B., Babuska, R., 2011. Cross-entropy optimization of control policies with adaptive basis functions. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 41, 196–209.

Byun, H.R., Wilhite, D.A., 1999. Objective quantification of drought severity and duration. Journal of Climate 12, 2747–2756.

Carkoglu, A., Eder, M., 2001. Water conflict: The euphrates-tigris basin. Turkey in World Politics: An Emerging Multiregional Power, edited by Kemal Kirisci and Barry Rubin. Boulder: Lynne Rienner .

Carmona, M., Máñez Costa, M., Andreu, J., Pulido-Velazquez, M., Haro-Monteagudo, D., Lopez-Nicolas, A., Cremades, R., 2017. Assessing the effectiveness of multi-sector partnerships to manage droughts: The case of the jucar river basin. Earth's Future .

Castelletti, A., Fedorov, R., Fraternali, P., Giuliani, M., 2016. Multimedia on the mountaintop: Using public snow images to improve water systems operation, in: Proceedings of the 24th ACM international conference on Multimedia, pp. 948–957.

Castelletti, A., Galelli, S., Restelli, M., Soncini-Sessa, R., 2010a. Tree-based reinforcement learning for optimal water reservoir operation. Water Resources Research 46.

Castelletti, A., Galelli, S., Restelli, M., Soncini-Sessa, R., 2010b. Tree-based reinforcement learning for optimal water reservoir operation. Water Resources Research 46.

Castelletti, A., Galelli, S., Restelli, M., Soncini-Sessa, R., 2011a. Tree-based variable selection for dimensionality reduction of large-scale control systems, in: 2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL), IEEE. pp. 62–69.

Castelletti, A., Pianosi, F., Restelli, M., 2011b. Multi-objective fitted q-iteration: Pareto frontier approximation in one single run, in: 2011 International Conference on Networking, Sensing and Control, IEEE. pp. 260–265.

Castelletti, A., Pianosi, F., Restelli, M., 2012. Tree-based fitted q-iteration for multi-objective markov decision problems, in: The 2012 International Joint Conference on Neural Networks (IJCNN), IEEE. pp. 1–8.

Castelletti, A., Pianosi, F., Restelli, M., 2013. A multiobjective reinforcement learning approach to water resources systems operation: Pareto frontier approximation in a single run. Water Resources Research 49, 3476–3486.

Castelletti, A., Pianosi, F., Soncini-Sessa, R., 2008a. Water reservoir control under economic, social and environmental constraints. Automatica 44, 1595–1607.

Castelletti, A., Pianosi, F., Soncini-Sessa, R., 2008b. Water reservoir control under economic, social and environmental constraints. Automatica 44, 1595–1607.

Celeste, A.B., Billib, M., 2009. Evaluation of stochastic reservoir operation optimization models. Advances in Water Resources 32, 1429–1443.

Chang, C.H., 2015. Deep and shallow architecture of multilayer neural networks. IEEE transactions on neural networks and learning systems 26, 2477–2486.

Changnon, S.A., 1987. Detecting drought conditions in illinois. Circular (Illinois State Water Survey); 169 .

CHD, 2007. Plan especial de actuación en situaciones de alerta y eventual sequía.

CHE, 2007. Plan especial de actuación en situaciones de alerta y eventual sequia en la cuenca hidrográfica del ebro. MARM. Zaragoza .

CHG, 2007. Plan especial de actuación en situaciones de alerta y eventual sequía de la cuenca hidrográfica del guadalquivir. CHG: Seville, Spain .

Chiew, F., Zhou, S., McMahon, T., 2003. Use of seasonal streamflow forecasts in water resources management. Journal of Hydrology 270, 135–144.

CHJ, 2007a. Anejo2 - Plan especial de alerta y eventual sequía en la confederación hidrográfica del Júcar. Confederación Hidrográfica del Júcar, Jucar River Basin Management Authority, Ministry of Agriculture, Food and Environment, Spanish Government, Valencia, Spain (in Spanish) .

CHJ, 2007b. Plan especial de alerta y eventual sequía en la confederación hidrográfica del Júcar. Confederación Hidrográfica del Júcar, Jucar River Basin Management Authority, Ministry of Agriculture, Food and Environment, Spanish Government, Valencia, Spain (in Spanish) .

Ciarapica, L., Todini, E., 2002. Topkapi: A model for the representation of the rainfall-runoff process at different scales. Hydrological Processes 16, 207–229.

Clapham, C., 2018. The ethiopian developmental state. Third World Quarterly 39, 1151–1165.

Cloke, H., Pappenberger, F., 2009. Ensemble flood forecasting: a review. Journal of Hydrology 375, 613–626.

Coello, C.A.C., Lamont, G.B., Van Veldhuizen, D.A., et al., 2007. Evolutionary algorithms for solving multi-objective problems. volume 5. Springer.

Costanzo, G.T., Iacovella, S., Ruelens, F., Leurs, T., Claessens, B.J., 2016. Experimental analysis of data-driven control for a building heating system. Sustainable Energy, Grids and Networks 6, 81–90.

Cunningham, P., 2008. Dimension reduction, in: Machine learning techniques for multimedia. Springer, pp. 91–112.

Curran, W., Brys, T., Aha, D., Taylor, M., Smart, W.D., 2016. Dimensionality reduced reinforcement learning for assistive robots, in: 2016 AAAI Fall Symposium Series.

Dai, A., 2011. Drought under global warming: a review, in: Interdisciplinary Reviews: Climate Change, pp. 45–65.

Degefu, M.A., Rowell, D.P., Bewket, W., 2017. Teleconnections between ethiopian rainfall variability and global ssts: observations and methods for model evaluation. Meteorology and Atmospheric Physics 129, 173–186.

Deisenroth, M.P., Neumann, G., Peters, J., et al., 2013. A survey on policy search for robotics. Foundations and Trends® in Robotics 2, 1–142.

Denaro, S., Anghileri, D., Giuliani, M., Castelletti, A., 2017a. Informing the operations of water reservoirs over multiple temporal scales by direct use of hydro-meteorological data. Advances in water resources 103, 51–63.

Denaro, S., Anghileri, D., Giuliani, M., Castelletti, A., 2017b. Informing the operations of water reservoirs over multiple temporal scales by direct use of hydro-meteorological data. Advances in Water Resources 10, 51–63.

Dominey, P.F., 1995. Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning. Biological cybernetics 73, 265–274.

Doorenbos, J., Kassam, A., 1979. Yield response to water. Irrigation and drainage paper , 257.

Dracup, J.A., Lee, K.S., Paulson, E.G., 1980. On the definition of droughts. Water Resources Research 16, 297–302. doi:doi: 10.1029/WR016i002p00297.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., et al., 2012. Sentinel-2: Esa's optical high-resolution mission for gmes operational services. Remote sensing of Environment 120, 25–36.

Durand, Y., Giraud, G., Laternser, M., Etchevers, P., Mérindol, L., Lesaffre, B., 2009. Reanalysis of 47 years of climate in the french alps (1958–2005): climatology and trends for snow cover. Journal of applied meteorology and climatology 48, 2487–2512.

Dutta, S., Patchaikani, P.K., Behera, L., 2015. Near-optimal controller for nonlinear continuous-time systems with unknown dynamics using policy iteration. IEEE transactions on neural networks and learning systems 27, 1537–1549.

Efe, M.Ö., 2008. Novel neuronal activation functions for feedforward neural networks. Neural processing letters 28, 63–79.

Efthymiadis, D., Jones, P., Briffa, K., Böhm, R., Maugeri, M., 2007. Influence of large-scale atmospheric circulation on climate variability in the greater alpine region of europe. Journal of Geophysical Research: Atmospheres 112.

El-Fakdi, A., Carreras, M., Ridao, P., 2006. Towards direct policy search reinforcement learning for robot control, in: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE. pp. 3178–3183.

Estrela, T., Vargas, E., 2012. Drought management plans in the european union. the case of spain. Water resources management 26, 1537–1553.

Ethiopian Electrical Power Company (EEPCO), 2009. Gibe iii hydroelectric project - environmental and social impact assessment.

EU, 2007. Water Scarcity and Droughts, Second Interim Report. Technical Report.

Faber, B., Stedinger, J., 2001. Reservoir optimization using sampling sdp with ensemble streamflow prediction (esp) forecasts. Journal of Hydrology 10, 22–1694.

Facchi, A., Ortuani, B., Maggi, D., Gandolfi, C., 2004. Coupled svat–groundwater model for water resources simulation in irrigated alluvial plains. Environmental Modeling & Software 19, 1053–1063.

Falkenmark, M., Lundqvist, J., Widstrand, C., 1989. Macro-scale water scarcity requires micro-scale approaches, in: Natural resources forum, Wiley Online Library. pp. 258–267.

Fitiwi Tekle, M., 2016. The Role of Gibe III Dam in Achieving Effective Cooperation between Ethiopia and Kenya. Ph.D. thesis. Addis Ababa University.

Floreano, D., Dürr, P., Mattiussi, C., 2008. Neuroevolution: from architectures to learning. Evolutionary Intelligence 1, 47–62.

Folland, C., Knight, J., Linderholm, H., Fereday, D., Ineson, S., Hurrell, J., 2009. The summer north atlantic oscillation: past, present, and future. Journal of Climate 22, 1082–1103.

Formentin, S., Karimi, A., Savaresi, S.M., 2013. Optimal input design for direct data-driven tuning of model-reference controllers. Automatica 49, 1874–1882.

Francis, R., Estlin, T., Doran, G., Johnstone, S., Gaines, D., Verma, V., Burl, M., Frydenvang, J., Montaño, S., Wiens, R., et al., 2017. Aegis autonomous targeting for chemcam on mars science laboratory: Deployment and results of initial science team use. Science Robotics 2, eaan4582.

Galelli, S., Castelletti, A., 2013. Tree-based iterative input variable selection for hydrological modeling. Water Resources Research 49, 4295–4310.

Galelli, S., Humphrey, G.B., Maier, H.R., Castelletti, A., Dandy, G.C., Gibbs, M.S., 2014. An evaluation framework for input variable selection algorithms for environmental data-driven models. Environmental Modelling and Software 62, 33–51. URL: http://dx.doi.org/10.1016/j.envsoft.2014.08.015, doi:doi: 10.1016/j.envsoft.2014.08.015.

Gandolfi, C., Sali, G., Facchi, A., Tediosi, A., Bulgheroni, C., Rienzner, M., Weber, E., 2014. Integrated modelling for agricultural policies and water resources planning coordination. biosystems engineering 128, 100–112.

Gao, B.C., 1996. Ndwi-a normalized difference water index for remote sensing of vegetation liquid water from space. Remote sensing of environment 58, 257–266.

Garrote, L., Martin-Carrasco, F., Flores-Montoya, F., Iglesias, A., 2007. Linking drought indicators to policy actions in the tagus basin drought management plan. Water resources management 21, 873–882.

Gashler, M., Giraud-Carrier, C., Martinez, T., 2008. Decision tree ensemble: Small heterogeneous is better than large homogeneous, in: 2008 Seventh International Conference on Machine Learning and Applications, IEEE. pp. 900–905.

Gauci, J., Stanley, K., 2007. Generating large-scale neural networks through discovering geometric regularities, in: Proceedings of the 9th annual conference on Genetic and evolutionary computation, ACM. pp. 997–1004.

Gaudel, R., Sebag, M., 2010. Feature Selection as a One-Player Game, in: International Conference on Machine Learning, Haifa, Israel. pp. 359–366. URL: https://hal.inria.fr/inria-00484049.

Gebre, S., 2019. Ethiopia insists dam will benefit region as egypt seeks mediator. URL: https://www.bloomberg.com/news/articles/2019-10-22/egypt-ethiopia-leaders-to-meet-in-coming-days-over-gerd-dam.

Gelati, E., Madsen, H., Rosbjerg, D., 2011. Stochastic reservoir optimization using el ni no information: case study of daule peripa, ecuador. Hydrology Research 42, 413–431.

Georgakakos, K., Graham, N., 2008. Potential benefits of seasonal inflow prediction uncertainty for reservoir release decisions. Journal of Applied Meteorology and Climatology 47, 1297–1321.

Georgakakos, K., Seo, D., Gupta, H., Schaake, J., Butts, M., 2004. Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. Journal of Hydrology 298, 222–241.

Giorgi, F., Lionello, P., 2008. Climate change projections for the mediterranean region. Global and planetary change 63, 90–104.

Giudici, F., Castelletti, A., Garofalo, E., Giuliani, M., Maier, H.R., 2019. Dynamic, multi-objective optimal design and operation of water-energy systems for small, off-grid islands. Applied Energy 250, 605–616.

Giuliani, M., Castelletti, A., 2016. Is robustness really robust? how different definitions of robustness impact decision-making under climate change. Climatic Change 135, 409–424.

Giuliani, M., Castelletti, A., 2019. Data-driven control of water reservoirs using el niño southern oscillation indexes, in: 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), IEEE. pp. 1–5.

Giuliani, M., Castelletti, A., Fedorov, R., Fraternali, P., 2016a. Using crowdsourced web content for informing water systems operations in snow-dominated catchments. Hydrology and Earth System Sciences 10, 20–5049.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., Reed, P.M., 2016b. Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. Journal of Water Resources Planning and Management 142, 04015050.

Giuliani, M., Galelli, S., Soncini-Sessa, R., 2014a. A dimensionality reduction approach for many-objective markov decision processes: Application to a water reservoir operation problem. Environmental Modelling & Software 57, 101–114.

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

Giuliani, M., Li, Y., Castelletti, A., , C., 2016c. A coupled human-natural systems analysis of irrigated agriculture under changing climate. Water Resources Research 52, 6928–6947.

Giuliani, M., Li, Y., Castelletti, A., Gandolfi, C., 2016d. A coupled human-natural systems analysis of irrigated agriculture under changing climate. Water Resources Research 52, 6928–6947.

Giuliani, M., Pianosi, F., Castelletti, A., 2015. Making the most of data: an information selection and assessment framework to improve water systems operations. Water Resources Research 51, 9073–9093.

Giuliani, M., Quinn, J.D., Herman, J.D., Castelletti, A., Reed, P.M., 2018. Scalable multiobjective control for large-scale water resources systems under uncertainty. IEEE Transactions on Control Systems Technology 26, 1492–1499.

Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., Block, P., 2019. Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. Water Resources Research 55, 9133–9147.

Gómez, C.M.G., Blanco, C.D.P., 2012. Do drought management plans reduce drought risk? a risk assessment model for a mediterranean river basin. Ecological Economics 76, 42–48.

Gong, M., Liu, J., Li, H., Cai, Q., Su, L., 2015. A multiobjective sparse feature learning model for deep neural networks. IEEE transactions on neural networks and learning systems 26, 3263–3277.

Goroshin, R., Mathieu, M.F., LeCun, Y., 2015. Learning to linearize under uncertainty, in: Advances in Neural Information Processing Systems, pp. 1234–1242.

Grantz, K., Rajagopalan, B., Zagona, E., Clark, M., 2007. Water management applications of climate-based hydrologic forecasts: case study of the truckee-carson river basin. Journal of Water Resources Planning and Management 133, 339–350.

Grimm, A., Tedeschi, R., 2009. Enso and extreme rainfall events in south america. Journal of Climate 22, 1589–1609.

Guariso, G., Orlovski, S., Rinaldi, S., Soncini-Sessa, R., 1984. An application of the risk-averse approach to the management of lake como. Journal of Applied Systems Analysis 5, 54–64.

Guariso, G., Rinaldi, S., Soncini-Sessa, R., 1986. The management of lake como: A multiobjective analysis. Water Resources Research 22, 109–120.

Gustard, A., Demuth, S., et al., 2009. Manual on low-flow estimation and prediction.

Guyon, I., 2003. An Introduction to Variable and Feature Selection 3, 1157–1182.

Hachiya, H., Sugiyama, M., 2010. Feature selection for reinforcement learning: Evaluating implicit state-reward dependency via conditional mutual information, in: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer. pp. 474–489.

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

Hadka, D., Reed, P., 2013. Borg: An auto-adaptive many-objective evolutionary computing framework. Evolutionary computation 21, 231–259.

Hagg, A., Mensing, M., Asteroth, A., 2017. Evolving parsimonious networks by mixing activation functions, in: Proceedings of the Genetic and Evolutionary Computation Conference, ACM. pp. 425–432.

Hamlet, A., Lettenmaier, D., 1999. Columbia river streamflow forecasting based on enso and pdo climate signals. Journal of water resources planning and management 125, 333–341.

Hao, Z., AghaKouchak, A., 2013. Multivariate standardized drought index: a parametric multi-index model. Advances in Water Resources 57, 12–18.

Haro, D., Solera, A., Paredes, J., Andreu, J., 2014a. Methodology for drought risk assessment in within-year regulated reservoir systems. application to the orbigo river system (spain). Water resources management 28, 3801.

Haro, D., Solera, A., Pedro-Monzonís, M., Andreu, J., 2014b. Optimal management of the jucar river and turia river basins under uncertain drought conditions. Procedia Engineering 89, 1260–1267.

Haro-Monteagudo, D., Solera, A., Andreu, J., 2017. Drought early warning based on optimal risk forecasts in regulated river systems: Application to the jucar river basin (spain). Journal of Hydrology 544, 36–45.

Hashimoto, T., Stedinger, J., Loucks, D., 1982. Reliability, resilience, and vulnerability criteria for water resource system performance evaluation. Water Resources Research 18, 14–20.

Hausknecht, M., Lehman, J., Miikkulainen, R., Stone, P., 2014. A neuroevolution approach to general atari game playing. IEEE Transactions on Computational Intelligence and AI in Games 6, 355–366.

Heidrich-Meisner, V., Igel, C., 2008. Evolution strategies for direct policy search, in: International Conference on Parallel Problem Solving from Nature, Springer. pp. 428–437.

Heim Jr, R.R., 2002. A review of twentieth-century drought indices used in the united states. Bulletin of the American Meteorological Society 83, 1149.

Hejazi, M.I., Cai, X., Ruddell, B.L., 2008a. The role of hydrologic information in reservoir operation–learning from historical releases. Advances in water resources 31, 1636–1650.

Hejazi, M.I., Cai, X., Ruddell, B.L., 2008b. The role of hydrologic information in reservoir operation learning from historical releases. Advances in Water Resources 10, 1636–1650.

Hornik, K., 1989. Multilayered feedforward networks are universal approximators. Neural Networks 2, 359–366.

Huang, G.B., Zhou, H., Ding, X., Zhang, R., 2012. Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 42, 513–529.

Huang, G.B., Zhu, Q.Y., Siew, C.K., 2006a. Extreme learning machine: theory and applications. Neurocomputing 70, 489–501.

Huang, G.B., Zhu, Q.Y., Siew, C.K., 2006b. Extreme learning machine: theory and applications. Neurocomputing 70, 489–501.

Huang, J., Higuchi, K., Shabbar, A., 1998. The relationship between the north atlantic oscillation and el ni no-southern oscillation. Geophysical Research Letters 25, 2707–2710.

Human Rights Watch, 2017. Ethiopia: Dams, plantations a threat to kenyans. URL: https://www.hrw.org/news/2017/02/14/ethiopia-dams-plantations-threat-kenyans.

Huntsberger, T., Aghazarian, H., Tunstel, E., 2005. Onboard adaptive learning for planetary surface rover control in rough terrain, in: Proceedings of the 2005 IEEE International Conference on Robotics and Automation, IEEE. pp. 4156–4163.

Hurrell, J., Loon, H.V., 1997. Decadal variations in climate associated with the north atlantic oscillation, in: Climatic change at high elevation sites, pp. 69–94.

International Energy Agency, 2014. Sustainable Energy for All 2013-2014: Global Tracking Framework Report. The World Bank.

Isotta, F.A., Frei, C., Weilguni, V., Tadić, M.P., Lassegues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S.M., et al., 2014. The climate of daily precipitation in the alps: development and analysis of a high-resolution grid dataset from pan-alpine rain-gauge data. International Journal of Climatology 34, 1657–1675.

IUCN, 2018. Lake turkana listed as in danger due to impacts from dam, as advised by iucn. URL: https://www.iucn.org/news/iucn-42whc/201806/lake-turkana-listed-danger-due-impacts-dam-advised-iucn.

James, D., Tucker, P., 2004. A comparative analysis of simplification and complexification in the evolution of neural network topologies, in: Proc. of Genetic and Evolutionary Computation Conference.

Jang, S.H., Yoon, J.W., Cho, S.B., 2009. Optimal strategy selection of non-player character on real time strategy game using a speciated evolutionary algorithm, in: Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on, IEEE. pp. 75–79.

Jiang, Y., Fan, J., Chai, T., Lewis, F.L., Li, J., 2017. Tracking control for linear discrete-time networked control systems with unknown dynamics and dropout. IEEE transactions on neural networks and learning systems 29, 4607–4620.

Jiau, M.K., Huang, S.C., 2018. Self-organizing neuroevolution for solving carpool service problem with dynamic capacity to alternate matches. IEEE transactions on neural networks and learning systems 30, 1048–1060.

Johnson, S., Stockdale, T., Ferranti, L., Balmaseda, M., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S., Mogensen, K., Zuo, H., Monge-Sanz, B., 2019. Seas5: the new ecmwf seasonal forecast system. Geoscientific Model Development 12, 1087–1117.

Joliffe, I., 2002. Principal Component Analysis. Springer, New York, N.Y.

Jolliffe, I., 2002. Principal component analysis. Springer.

Jozaghi, A., Alizadeh, B., Hatami, M., Flood, I., Khorrami, M., Khodaei, N., Ghasemi Tousi, E., 2018. A comparative study of the ahp and topsis techniques for dam site selection using gis: A case study of sistan and baluchestan province, iran. Geosciences 8, 494.

Kahya, E., Dracup, J., 1993. Us streamflow patterns in relation to the el ni no/southern oscillation. Water resources research 29, 2491–2503.

Kamruzzaman, J., Aziz, S.M., 2002. A note on activation function in multilayer feedforward learning, in: Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on, IEEE. pp. 519–523.

Karakaya, G., Galelli, S., Ahipasaoglu, S.D., Taormina, R., 2015. Identifying (Quasi) Equally Informative Subsets in Feature Selection Problems for Classification: A Max-Relevance Min-Redundancy Approach. IEEE Transactions on Cybernetics PP, 1. URL: http://ieeexplore.ieee.org/ielx7/6221036/6352949/07150365.pdf?tp={\{}{&}{\}}arnumber=7150365{\{}{&}{\}}isnumber=6352949http://ieeexplore.ieee.org/xpls/abs{\{}{_}{\}}all.jsp?arnumber=7150365{\{}{&}{\}}tag=1, doi:doi:10.1109/TCYB.2015.2444435.

Karakovskiy, S., Togelius, J., 2012. The mario ai benchmark and competitions. IEEE Transactions on Computational Intelligence and AI in Games 4, 55–67.

Keller, P.W., Mannor, S., Precup, D., 2006. Automatic basis function construction for approximate dynamic programming and reinforcement learning, in: Proceedings of the 23rd international conference on Machine learning, pp. 449–456.

Kelman, J., Stedinger, J.R., Cooper, L.A., Hsu, E., Yuan, S.Q., 1990. Sampling stochastic dynamic programming applied to reservoir operation. Water Resources Research 26, 447–454.

Keyantash, J., Dracup, J.A., 2002. The quantification of drought: an evaluation of drought indices. Bulletin of the American Meteorological Society 83, 1167–1180.

Keyantash, J.A., Dracup, J.A., 2004. An aggregate drought index: Assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. Water Resources Research 40.

Kim, Y., Palmer, R., 1997a. Value of seasonal flow forecasts in bayesian stochastic programming. Journal of Water Resources Planning and Management 123, 327–335.

Kim, Y.O., Palmer, R.N., 1997b. Value of seasonal flow forecasts in bayesian stochastic programming. Journal of Water Resources Planning and Management 123, 327–335.

King, A., Block, P., 2014. An assessment of reservoir filling policies for the grand ethiopian renaissance dam. Journal of Water and Climate Change 5, 233–243.

Kingston, D., Lawler, D., (2006a), G.M., . Linkages between atmospheric circulation, climate and streamflow in the northern north atlantic: research prospects. Progress in Physical Geography 30, 143–174.

Kingston, D., McGregor, G., Hannah, D., Lawler, D., 2006. River flow teleconnections across the northern north atlantic region. Geophysical Research Letters 33.

Knight, W., 2017. 10 breakthrough technologies: Reinforcement learning .

Kohl, N., Miikkulainen, R., 2008. Evolving neural networks for fractured domains, in: Proceedings of the 10th annual conference on Genetic and evolutionary computation, ACM. pp. 1405–1412.

Kohl, N., Miikkulainen, R., 2012. An integrated neuroevolutionary approach to reactive control and high-level strategy. IEEE Transactions on Evolutionary Computation 16, 472–488.

Kollat, J.B., Reed, P.M., 2005. The value of online adaptive search: a performance comparison of nsgaii, ε-nsgaii and εmoea, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer. pp. 386–398.

Kollat, J.B., Reed, P.M., 2006. Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. Advances in Water Resources 29, 792–807.

Kolter, J.Z., Ng, A.Y., 2009. Regularization and feature selection in least-squares temporal difference learning, in: Proceedings of the 26th annual international conference on machine learning, pp. 521–528.

Kondolf, G.M., Schmitt, R.J., Carling, P., Darby, S., Arias, M., Bizzi, S., Castelletti, A., Cochrane, T.A., Gibson, S., Kummu, M., et al., 2018. Changing sediment budget of the mekong: Cumulative threats and management strategies for a large river basin. Science of the total environment 625, 114–134.

Krčah, P., 2012. Effects of speciation on evolution of neural networks in highly dynamic environments, in: International Conference on Learning and Intelligent Optimization, Springer. pp. 425–430.

Kroon, M., Whiteson, S., 2009. Automatic feature selection for model-based reinforcement learning in factored mdps, in: 2009 International Conference on Machine Learning and Applications, IEEE. pp. 324–330.

Kucukgocmen, A., 2013. Turkey starts filling huge tigris river dam, activists say. URL: https://www.reuters.com/article/us-turkey-dam/turkey-starts-filling-huge-tigris-river-dam-activists-say-idUSKCN1US194.

Kummu, M., Ward, P.J., de Moel, H., Varis, O., 2010. Is physical water scarcity a new phenomenon? global assessment of water shortage over the last two millennia. Environmental Research Letters 5, 034006.

Laaha, G., Gauster, T., Tallaksen, L.M., Vidal, J.P., Stahl, K., Prudhomme, C., Heudorfer, B., Vlnas, R., Ionita, M., Van Lanen, H.A., et al., 2016. The european 2015 drought from a hydrological perspective. Hydrology and Earth System Sciences Discussions .

Labadie, J.W., 2004. Optimal operation of multireservoir systems: state-of-the-art review. Journal of water resources planning and management 130, 93–111.

Lanckriet, S., Frankl, A., Adgo, E., Termonia, P., Nyssen, J., 2015. Droughts related to quasi-global oscillations: a diagnostic teleconnection analysis in north ethiopia. International Journal of Climatology 35, 1534–1542.

Laudani, A., Lozito, G.M., Fulginei, F.R., Salvini, A., 2015. On training efficiency and computational costs of a feed forward neural network: a review. Computational intelligence and neuroscience 2015, 83.

Lee, D., Ward, P., Block, P., 2018. Attribution of large-scale climate patterns to seasonal peak-flow and prospects for prediction globally. Water Resources Research 54, 916–938.

Lehman, J., Stanley, K.O., 2010. Revising the evolutionary computation abstraction: minimal criteria novelty search, in: Proceedings of the 12th annual conference on Genetic and evolutionary computation, ACM. pp. 103–110.

Lesort, T., Díaz-Rodríguez, N., Goudou, J.F., Filliat, D., 2018. State representation learning for control: An overview. Neural Networks 108, 379–392.

Li, Y., Giuliani, M., Castelletti, A., 2017. A coupled human–natural system to assess the operational value of weather and climate services for agriculture. Hydrology and Earth System Sciences 21, 4693–4709.

Libisch-Lehner, C., Nguyen, H., Taormina, R., Nachtnebel, H., Galelli, S., 2019. On the value of enso state for urban water supply system operators: Opportunities, trade-offs, and challenges. Water Resources Research 55, 2856–2875.

Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D., 2015. Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971 .

Lindsey, R., 2016. Global impacts of el niño and la niña. URL: https://www.climate.gov/news-features/featured-images/global-impacts-el-ni%C3%B1o-and-la-ni%C3%B1.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S., 1997. Development and test of the distributed hbv-96 hydrological model. Journal of hydrology 201, 272–288.

Liu, C., Xu, X., Hu, D., 2014. Multiobjective reinforcement learning: A comprehensive overview. IEEE Transactions on Systems, Man, and Cybernetics: Systems 45, 385–398.

Liu, D.R., Li, H.L., Wang, D., 2015. Feature selection and feature learning for high-dimensional batch reinforcement learning: A survey. International Journal of Automation and Computing 12, 229–242.

Liu, J., Gong, M., Miao, Q., Wang, X., Li, H., 2017. Structure learning for deep neural networks based on multiobjective optimization. IEEE transactions on neural networks and learning systems 29, 2450–2463.

López-Moreno, J., Vicente-Serrano, S., Morán-Tejeda, E., Lorenzo-Lacruz, J., Zabalza, J., Kenawy, A.E., Beniston, M., 2011. Influence of winter north atlantic oscillation index (nao) on climate and snow accumulation in the mediterranean mountains, in: Hydrological, Socioeconomic and Ecological Impacts of the North Atlantic Oscillation in the Mediterranean Region, pp. 73–89.

Lorenzo-Lacruz, J., Vicente-Serrano, S.M., L??pez-Moreno, J.I., Beguer??a, S., Garc??a-Ruiz, J.M., Cuadrat, J.M., 2010. The impact of droughts and water management on various hydrological systems in the headwaters of the Tagus River (central Spain). Journal of Hydrology 386, 13–26. URL: http://dx.doi.org/10.1016/j.jhydrol.2010.01.001, doi:doi: 10.1016/j.jhydrol.2010.01.001.

Loscalzo, S., Wright, R., Acunto, K., Yu, L., 2012. Sample aware embedded feature selection for reinforcement learning, in: Proceedings of the 14th annual conference on Genetic and evolutionary computation, pp. 887–894.

Loscalzo, S., Wright, R., Yu, L., 2015. Predictive feature selection for genetic policy search. Autonomous Agents and Multi-Agent Systems 29, 754–786.

Lu, M., Lall, U., Robertson, A.W., Cook, E., 2017. Optimizing multiple reliable forward contracts for reservoir allocation using multitime scale streamflow forecasts. Water Resources Research 53, 2035–2050.

Macian-Sorribes, H., Pulido-Velazquez, M., 2017. Integrating historical operating decisions and expert criteria into a dss for the management of a multireservoir system. Journal of Water Resources Planning and Management 143, 04016069.

MacKay, D.J., 2003. Information theory, inference and learning algorithms. Cambridge university press.

Maidment, R.I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., Allan, R.P., Stein, T., Nkonde, E., Senkunda, S., et al., 2017. A new, long-term daily satellite-based rainfall dataset for operational monitoring in africa. Scientific data 4, 170063.

Maillard, O.A., Ryabko, D., Munos, R., 2011. Selecting the state-representation in reinforcement learning, in: Advances in Neural Information Processing Systems, pp. 2627–2635.

Mandel, T., Liu, Y.E., Brunskill, E., Popović, Z., 2016. Offline evaluation of online reinforcement learning algorithms, in: Thirtieth AAAI Conference on Artificial Intelligence.

Mandel, T., Liu, Y.E., Levine, S., Brunskill, E., Popovic, Z., 2014. Offline policy evaluation across representations with applications to educational games., in: AAMAS, pp. 1077–1084.

Marcos-Garcia, P., Lopez-Nicolas, A., Pulido-Velazquez, M., 2017. Combined use of relative drought indices to analyze climate change impact on meteorological and hydrological droughts in a mediterranean basin. Journal of Hydrology 554, 292–305.

Mariotti, A., Zeng, N., Lau, K.M., 2002. Euro-mediterranean rainfall and enso?a seasonally varying relationship. Geophysical research letters 29, 59–1.

Matyasovszky, I., 2003. The relationship between nao and temperature in hungary and its nonlinear connection with enso. Theoretical and Applied Climatology 74, 69–75.

Maurer, E.P., Lettenmaier, D.P., 2004. Potential effects of long-lead hydrologic predictability on missouri river main-stem reservoirs. Journal of Climate 17, 174–186.

McKee, T.B., Doesken, N.J., Kleist, J., et al., 1993. The relationship of drought frequency and duration to time scales, in: Proceedings of the 8th Conference on Applied Climatology, American Meteorological Society Boston, MA. pp. 179–183.

McPhaden, M., Zebiak, S., Glantz, M., 2006. Enso as an integrating concept in earth science. science 314, 1740–1745.

Metzen, J.H., Edgington, M., Kassahun, Y., Kirchner, F., 2008. Analysis of an evolutionary reinforcement learning method in a multiagent domain, in: Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1, International Foundation for Autonomous Agents and Multiagent Systems. pp. 291–298.

Ministerio del Medio Ambiente, 2000. Plan hidrológico nacional.

Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. Journal of Hydrology 391, 202–216. doi:doi: 10.1016/j.jhydrol.2010.07.012.

Mishra, A.K., Singh, V.P., 2011. Drought modeling - A review. Journal of Hydrology 403, 157–175. URL: http://dx.doi.org/10.1016/j.jhydrol.2011.03.049, doi:doi: 10.1016/j.jhydrol.2011.03.049, arXiv:arXiv:1011.1669v3.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al., 2015. Human-level control through deep reinforcement learning. Nature 518, 529–533.

Mondal, M.A.H., Bryan, E., Ringler, C., Rosegrant, M., 2017. Ethiopian power sector development: Renewable based universal electricity access and export strategies. Renewable and Sustainable Energy Reviews 75, 11–20.

Moriguchi, H., Honiden, S., 2010. Sustaining behavioral diversity in neat, in: Proceedings of the 12th annual conference on Genetic and evolutionary computation, ACM. pp. 611–618.

Morimoto, J., Hyon, S.H., Atkeson, C.G., Cheng, G., 2008. Low-dimensional feature extraction for humanoid locomotion using kernel dimension reduction, in: 2008 IEEE International Conference on Robotics and Automation, IEEE. pp. 2711–2716.

Munk, J., Kober, J., Babuška, R., 2016. Learning state representation for deep actor-critic control, in: 2016 IEEE 55th Conference on Decision and Control (CDC), IEEE. pp. 4667–4673.

Narasimhan, B., Srinivasan, R., 2005. Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. Agricultural and Forest Meteorology 133, 69–88. doi:doi: 10.1016/j.agrformet.2005.07.012.

Nayak, M., Herman, J., Steinschneider, S., 2018. Balancing flood risk and water supply in california: Policy search integrating short-term forecast ensembles with conjunctive use. Water Resources Research 54, 7557–7576.

Neitsch, S., Arnold, J., Kiniry, J., Williams, J., 2011. Soil and water assessment tool theoretical documentation version 2009. Tech. Rep. 406 Grassland, Soil and Water Research Laboratory–Agricultural Research Service Blackland Research Center–Texas AgriLife Research, College Station, Texas .

Nichele, S., Ose, M.B., Risi, S., Tufte, G., 2018. Ca-neat: evolved compositional pattern producing networks for cellular automata morphogenesis and replication. IEEE Transactions on Cognitive and Developmental Systems 10, 687–700.

Nouri, A., Littman, M.L., 2010. Dimension reduction and its application to model-based exploration in continuous spaces. Machine Learning 81, 85–98.

Nowak, K., Prairie, J., Rajagopalan, B., Lall, U., 2010. A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow. Water Resources Research 10, 1029.

Oh, J., Singh, S., Lee, H., 2017. Value prediction network, in: Advances in Neural Information Processing Systems, pp. 6118–6128.

Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. science 313, 1068–1072.

Padowski, J., Gorelick, S., Thompson, B., Rozelle, S., Fendorf, S., 2015. Assessment of human–natural system characteristics influencing global freshwater supply vulnerability. Environmental Research Letters 10.

Palmer, T., Hagedorn, R., 2006. Predictability of weather and climate. Cambridge University Press.

Palmer, W.C., 1965. Meteorological drought. volume 30. US Department of Commerce, Weather Bureau Washington, DC.

Parr, R., Painter-Wakefield, C., Li, L., Littman, M., 2007. Analyzing feature generation for value-function approximation, in: Proceedings of the 24th international conference on Machine learning, pp. 737–744.

Paschalis, A., Fatichi, S., Molnar, P., Rimkus, S., Burlando, P., 2014. On the effects of small scale space–time variability of rainfall on basin flood response. Journal of Hydrology 514, 313–327.

Paul, C.J., Weinthal, E., 2019. The development of ethiopia's climate resilient green economy 2011–2014: implications for rural adaptation. Climate and Development 11, 193–202.

Pedro-Monzonís, M., Ferrer, J., Solera, A., Estrela, T., Paredes-Arquiola, J., 2014. Water accounts and water stress indexes in the european context of water planning: the jucar river basin. Procedia Engineering 89, 1470–1477.

Pedro-Monzonìs, M., Solera, A., Ferrer, J., Estrela, T., Paredes-Arquiola, J., 2015. A review of water scarcity and drought indexes in water resources planning and management. Journal of Hydrology 527, 482–493. doi:doi: 10.1016/j.jhydrol.2015.05.003.

Piccardi, C., Soncini-Sessa, R., 1991. Stochastic dynamic programming for reservoir optimal control: dense discretization and inflow correlation assumption made possible by parallel computing. Water Resources Research 27, 729–741.

Poveda, G., Alvarez, D., Rueda, O., 2011. Hydro-climatic variability over the andes of colombia associated with enso: a review of climatic processes and their impact on one of the earth?s most important biodiversity hotspots. Climate Dynamics 36, 2233–2249.

Powell, W.B., 2007. Approximate Dynamic Programming: Solving the curses of dimensionality. volume 703. John Wiley & Sons.

Quinn, J.D., Reed, P.M., Giuliani, M., Castelletti, A., 2017. Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems. Water Resources Research 53, 7208–7233.

Quinn, J.D., Reed, P.M., Giuliani, M., Castelletti, A., 2019. What is controlling our control rules? opening the black box of multireservoir operating policies using time-varying sensitivity analysis. Water Resources Research 55, 5962–5984.

Quinn, J.D., Reed, P.M., Giuliani, M., Castelletti, A., Oyler, J.W., Nicholas, R.E., 2018. Exploring how changing monsoonal dynamics and human pressures challenge multireservoir management for flood protection, hydropower production, and agricultural water supply. Water Resources Research 54, 4638–4662.

Rajagopalan, B., Lall, U., Tarboton, D., Bowles, D., 1997. Multivariate nonparametric resampling scheme for generation of daily weather variables. Stochastic Hydrology and Hydraulics 11, 65–93.

Rajeswaran, A., Ghotra, S., Ravindran, B., Levine, S., 2016. Epopt: Learning robust neural network policies using model ensembles. arXiv preprint arXiv:1610.01283 .

Raskin, P., Gleick, P., Kirshen, P., Pontius, G., Strzepek, K., 1997. Water futures: Assessment of long-range patterns and problems. Comprehensive assessment of the freshwater resources of the world. SEI.

Raso, L., Schwanenberg, D., van de Giesen, N., van Overloop, P.J., 2014. Short-term optimal operation of water systems using ensemble forecasts. Advances in water resources 71, 200–208.

Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B., 2013. Evolutionary multiobjective optimization in water resources: The past, present, and future. Advances in water resources 51, 438–456.

Rempis, C., Pasemann, F., 2012. An interactively constrained neuro-evolution approach for behavior control of complex robots, in: Variants of Evolutionary Algorithms for Real-World Applications. Springer, pp. 305–341.

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.K., et al., 2011. Merra: Nasa's modern-era retrospective analysis for research and applications. Journal of climate 24, 3624–3648.

Rijsberman, F.R., 2006. Water scarcity: fact or fiction? Agricultural water management 80, 5–22.

Rippl, W., 1883. The capacity of storage-reservoirs for water-slpply.(including plate)., in: Minutes of the Proceedings of the Institution of Civil Engineers, Thomas Telford-ICE Virtual Library. pp. 270–278.

Risi, S., Hughes, C.E., Stanley, K.O., 2010. Evolving plastic neural networks with novelty search. Adaptive Behavior 18, 470–491.

Risi, S., Togelius, J., 2015. Neuroevolution in games: State of the art and open challenges. IEEE Transactions on Computational Intelligence and AI in Games 9, 25–41.

Rodell, M., Famiglietti, J., Wiese, D., Reager, J., Beaudoing, H., Landerer, F., Lo, M., 2018. Emerging trends in global freshwater availability. Nature 1.

Rolnick, D., Donti, P., Kaack, L., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A., Maharaj, T., Sherwin, E., Mukkavilli, S.K., Kording, K., Gomes, C., Ng, A., Hassabis, D., Platt, J., Creutzig, F., Chayes, J., Bengio, Y., 2019. Tackling Climate Change with Machine Learning.

Roth, K., 2019. Kenya - events of 2018. URL: `https://www.hrw.org/world-report/2019/country-chapters/kenya`.

Rouse Jr, J., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the great plains with erts. NASA special publication 351, 309.

Rusu, A.A., Rabinowitz, N.C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., Hadsell, R., 2016. Progressive neural networks. arXiv preprint arXiv:1606.04671 .

Saavedra Valeriano, O.C., Koike, T., Yang, K., Graf, T., Li, X., Wang, L., Han, X., 2010. Decision support for dam release during floods using a distributed biosphere hydrological model driven by quantitative precipitation forecasts. Water Resources Research 46.

Salazar, J.Z., Reed, P.M., Quinn, J.D., Giuliani, M., Castelletti, A., 2017. Balancing exploration, uncertainty and computational demands in many objective reservoir optimization. Advances in Water Resources 109, 196–210.

Sallab, A.E., Abdou, M., Perot, E., Yogamani, S., 2017. Deep reinforcement learning framework for autonomous driving. Electronic Imaging 2017, 70–76.

Sarachik, E., Cane, M., 2010. The El Niño-southern oscillation phenomenon. Cambridge University Press.

Scherrer, S., Appenzeller, C., Laternser, M., 2004. Trends in swiss alpine snow days: The role of local-and large-scale climate variability. Geophysical Research Letters 31.

Schmitt, R., Bizzi, S., Castelletti, A., Opperman, J., Kondolf, G., 2019. Planning dam portfolios for low sediment trapping shows limits for sustainable hydropower in the mekong. Science Advances 5, eaaw2175.

Schmitt, R.J., Bizzi, S., Castelletti, A., Kondolf, G., 2018. Improved trade-offs of hydropower and sand connectivity by strategic dam planning in the mekong. Nature Sustainability 1, 96–104.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., . Gpcc full data monthly product version 2018 at 0.5Â°: Monthly land-surface precipitation from rain-gauges built on gts-based and historical data (global precipitation climatology centre, 2018).

Schoenauer, M. (Ed.), 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II.

Schrum, J., Miikkulainen, R., 2008. Constructing complex npc behavior via multi-objective neuroevolution. AIIDE 8, 108–113.

Schwatke, C., Dettmering, D., Bosch, W., Seitz, F., 2015. Dahiti - an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. Hydrology and Earth System Sciences 19, 4345–4364.

Scott, D.W., 2012. Multivariate density estimation and visualization, in: Handbook of computational statistics. Springer, pp. 549–569.

Shafer, B., Dezman, L., 1982. Development of a surface water supply index (swsi) to assess the severity of drought conditions in snowpack runoff areas, in: Proceedings of the western snow conference, Colorado State University Fort Collins, CO. pp. 164–175.

Shalev-Shwartz, S., Shammah, S., Shashua, A., 2016. Safe, multi-agent, reinforcement learning for autonomous driving. arXiv preprint arXiv:1610.03295 .

Shaman, J., 2014. The seasonal effects of enso on european precipitation: Observational analysis. Journal of Climate 27, 6423–6438.

Sharma, A., 2000a. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1-a strategy for system predictor identification. Journal of Hydrology 239, 232–239.

Sharma, A., 2000b. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 3 – a nonparametric probabilistic forecast model. Journal of Hydrology 239, 249–258.

Sharma, A., Luk, K., Cordery, I., Lall, U., 2000. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 2 – predictor identification of quarterly rainfall using ocean-atmosphere information. Journal of Hydrology 239, 240–248.

Sharma, A., Mehrotra, R., 2014. An information theoretic alternative to model a natural system using observational information alone. Water Resources Research 50, 650–660.

Shelhamer, E., Mahmoudieh, P., Argus, M., Darrell, T., 2016. Loss is its own reward: Self-supervision for reinforcement learning. arXiv preprint arXiv:1612.07307 .

Shortreed, S.M., Laber, E., Lizotte, D.J., Stroup, T.S., Pineau, J., Murphy, S.A., 2011. Informing sequential clinical decision-making through reinforcement learning: an empirical study. Machine learning 84, 109–136.

Si, W., Li, J., Ding, P., Rao, R., 2017. A multi-objective deep reinforcement learning approach for stock index future's intraday trading, in: 2017 10th International symposium on computational intelligence and design (ISCID), IEEE. pp. 431–436.

Sigaud, O., Stulp, F., 2019. Policy search in continuous action domains: an overview. Neural Networks .

Simonovic, S.P., 1992. Reservoir systems analysis: closing gap between theory and practice. Journal of Water Resources Planning and Management 118, 262–280.

Singh, B., Kushwaha, N., Vyas, O.P., et al., 2014. A feature subset selection technique for high dimensional data using symmetric uncertainty. Journal of Data Analysis and Information Processing 2, 95.

Smarra, F., Jain, A., de Rubeis, T., Ambrosini, D., D'Innocenzo, A., Mangharam, R., 2018. Data-driven model predictive control using random forests for building energy optimization and climate control. Applied energy 226, 1252–1272.

Soncini-Sessa, R., Weber, E., Castelletti, A., 2007. Integrated and participatory water resources management-theory. volume 1. Elsevier.

Souza Filho, F.A., Lall, U., 2003. Seasonal to interannual ensemble streamflow forecasts for ceara, brazil: Applications of a multivariate, semiparametric algorithm. Water Resources Research 39.

Spinage, C.A., 2012. African ecology: benchmarks and historical perspectives. Springer Science & Business Media.

Spinoni, J., Naumann, G., Carrao, H., Barbosa, P., Vogt, J., 2014. World drought frequency, duration, and severity for 1951–2010. International Journal of Climatology 34, 2792–2804.

Spinoni, J., Naumann, G., Vogt, J., Barbosa, P., 2016. Meteorological droughts in europe .

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 15, 1929–1958.

Stahl, K., Kohn, I., Blauhut, V., Urquijo, J., De Stefano, L., Acácio, V., Dias, S., Stagge, J.H., Tallaksen, L.M., Kampragou, E., et al., 2016. Impacts of european drought events: insights from an international database of text-based reports. Natural Hazards and Earth System Sciences 16, 801–819.

Stanley, K.O., Bryant, B.D., Miikkulainen, R., 2005. Real-time neuroevolution in the nero video game. IEEE transactions on evolutionary computation 9, 653–668.

Stanley, K.O., Miikkulainen, R., 2002. Efficient reinforcement learning through evolving neural network topologies, in: Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation, Morgan Kaufmann Publishers Inc.. pp. 569–577.

Stanley, K.O., Miikkulainen, R., 2003. A taxonomy for artificial embryogeny. Artif. Life 9, 93–130.

Staudinger, M., Stahl, K., Seibert, J., 2014. A drought index accounting for snow. Journal of Hydrology 6, 2108–2123. doi:doi: 10.1002/2012WR013085.Received.

Stedinger, J.R., Sule, B.F., Loucks, D.P., 1984. Stochastic dynamic programming models for reservoir operation optimization. Water resources research 20, 1499–1505.

Steduto, P., Hsiao, T., Raes, D., Fereres, E., 2009. Aquacrop –the fao crop model to simulate yield response to water: I. concepts and underlying principles. Agronomy Journal 101, 426–437.

Steirou, E., Gerlitz, L., Apel, H., Merz, B., 2017. Links between large-scale circulation patterns and streamflow in central europe: A review. Journal of hydrology 549, 484–500.

Studley, M., Bull, L., 2007. Using the xcs classifier system for multi-objective reinforcement learning problems. Artificial Life 13, 69–86.

Sturtevant, N.R., White, A.M., 2006. Feature construction for reinforcement learning in hearts, in: International Conference on Computers and Games, Springer. pp. 122–134.

Sullivan, C.A., Meigh, J.R., Giacomello, A.M., 2003. The water poverty index: development and application at the community scale, in: Natural Resources Forum, Wiley Online Library. pp. 189–199.

Sutton, R.S., Barto, A.G., et al., 1998. Introduction to reinforcement learning. volume 2. MIT press Cambridge.

Szita, I., 2012. Reinforcement learning in games, in: Reinforcement learning. Springer, pp. 539–577.

Tallaksen, L.M., Van Lanen, H.A., 2004. Hydrological drought: processes and estimation methods for streamflow and groundwater. volume 48. Elsevier.

Tan, M., Deklerck, R., Cornelis, J., Jansen, B., 2013. Phased searching with neat in a time-scaled framework: experiments on a computer-aided detection system for lung nodules. Artificial intelligence in medicine 59, 157–167.

Tan, M., Deklerck, R., Jansen, B., Cornelis, J., 2012. Analysis of a feature-deselective neuroevolution classifier (fd-neat) in a computer-aided lung nodule detection system for ct images, in: Proceedings of the 14th annual conference companion on Genetic and evolutionary computation, pp. 539–546.

Tangkaratt, V., Morimoto, J., Sugiyama, M., 2016. Model-based reinforcement learning with dimension reduction. Neural Networks 84, 1–16.

Taormina, R., Galelli, S., Karakaya, G., Ahipasaoglu, S., 2016. An information theoretic approach to select alternate subsets of predictors for data-driven hydrological models. Journal of Hydrology 542, 18–34. URL: http://linkinghub.elsevier.com/retrieve/pii/S0022169416304759, doi:doi: 10.1016/j.jhydrol.2016.07.045.

Tessama, Z., Davis, M., Tella, P.V., Lambe, F., 2013. Mainstreaming sustainable energy access into national development planning: The case of Ethiopia. Stockholm Environment Institute.

The Economist, 2016a. Ethiopia opens africa's tallest and most controversial dam.

The Economist, 2016b. Sharing the nile. URL: https://www.economist.com/middle-east-and-africa/2016/01/16/sharing-the-nile.

Thornthwaite, C.W., et al., 1948. An approach toward a rational classification of climate. Geographical review 38, 55–94.

Todini, E., 2014. The role of predictive uncertainty in the operational management of reservoirs. Proceedings of the ICWRS2014 Evolving Water Resources Systems: Understanding, Predicting and Managing Water-Society Interactions, Bologna, Italy , 4–6.

Trenberth, K., 2011. Changes in precipitation with climate change. Climate Research 47, 123–138.

Tsitsiklis, J.N., Van Roy, B., 1996. Feature-based methods for large scale dynamic programming. Machine Learning 22, 59–94.

Turco, M., Ceglar, A., Prodhomme, C., Soret, A., Toreti, A., Francisco, J., 2017. Summer drought predictability over europe: empirical versus dynamical forecasts. Environmental Research Letters 12.

Turner, S., Bennett, J., Robertson, D., Galelli, S., 2017. Complex relationship between seasonal streamflow forecast skill and value in reservoir operations. Hydrology and Earth System Sciences 21, 4841–4859.

Turner, S., Galelli, S., 2016. Regime-shifting streamflow processes: Implications for water supply reservoir operations. Water Resources Research 52, 3984–4002.

UNESCO, 2018. Lake turkana national parks (kenya) inscribed on list of world heritage in danger. URL: http://whc.unesco.org/en/news/1842.

Vamplew, P., Dazeley, R., Berry, A., Issabekov, R., Dekker, E., 2011. Empirical evaluation methods for multiobjective reinforcement learning algorithms. Machine learning 84, 51–80.

Vamplew, P., Yearwood, J., Dazeley, R., Berry, A., 2008. On the limitations of scalarisation for multi-objective reinforcement learning of pareto fronts, in: Australasian Joint Conference on Artificial Intelligence, Springer. pp. 372–378.

Van Hoof, H., Chen, N., Karl, M., van der Smagt, P., Peters, J., 2016. Stable reinforcement learning with autoencoders for tactile and visual data, in: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE. pp. 3928–3934.

Van Loon, A., Van Lanen, H., 2012. A process-based typology of hydrological drought. Hydrology and Earth System Sciences 16, 1915.

Van Loon, A.F., Van Lanen, H.A.J., 2013. Making the distinction between water scarcity and drought using an observation-modeling framework. Water Resources Research 49, 1483–1502. doi:doi: 10.1002/wrcr.20147.

Vargas, D.V., Murata, J., 2017. Spectrum-diverse neuroevolution with unified neural models. IEEE transactions on neural networks and learning systems 28, 1759–1773.

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. Journal of Climate 23, 1696–1718.

Vicente-Serrano, S.M., López-Moreno, J.I., 2005. Hydrological response to different time scales of climatological drought: an evaluation of the standardized precipitation index in a mountainous mediterranean basin. Hydrology and Earth System Sciences Discussions 9, 523–533.

Vigorito, C.M., Barto, A.G., 2009. Incremental structure learning in factored mdps with continuous states and actions. University of Massachusetts Amherst-Department of Computer Science, Tech. Rep .

Wahlström, N., Schön, T.B., Deisenroth, M.P., 2015. From pixels to torques: Policy learning with deep dynamical models. arXiv preprint arXiv:1502.02251 .

Wanders, N., Van Lanen, H.A., van Loon, A.F., 2010. Indicators for drought characterization on a global scale. Technical Report. Wageningen Universiteit.

Wang, T., Gao, H., Qiu, J., 2015. A combined adaptive neural network and nonlinear model predictive control for multirate networked industrial process control. IEEE Transactions on Neural Networks and Learning Systems 27, 416–425.

Ward, P.J., Beets, W., Bouwer, L.M., Aerts, J.C., Renssen, H., 2010. Sensitivity of river discharge to ENSO. Geophysical Research Letters 37.

Ward, P.J., Jongman, B., Kummu, M., Dettinger, M.D., Weiland, F.C.S., Winsemius, H.C., 2014a. Strong influence of el niño southern oscillation on flood risk around the world. Proceedings of the National Academy of Sciences 111, 15659–15664.

Ward, P.J., Jongman, B., Kummu, M., Dettinger, M.D., Weiland, F.C.S., Winsemius, H.C., 2014b. Strong influence of el niño southern oscillation on flood risk around the world. Proceedings of the National Academy of Sciences 111, 15,659–15,664.

Watter, M., Springenberg, J., Boedecker, J., Riedmiller, M., 2015. Embed to control: A locally linear latent dynamics model for control from raw images, in: Advances in neural information processing systems, pp. 2746–2754.

Welch, R., Limonadi, D., Manning, R., 2013. Systems engineering the curiosity rover: A retrospective, in: 2013 8th International Conference on System of Systems Engineering, IEEE. pp. 70–75.

Wheeler, K.G., Basheer, M., Mekonnen, Z.T., Eltoum, S.O., Mersha, A., Abdo, G.M., Zagona, E.A., Hall, J.W., Dadson, S.J., 2016. Cooperative filling approaches for the grand ethiopian renaissance dam. Water International 41, 611–634.

Whiteson, S., Stone, P., Stanley, K.O., Miikkulainen, R., Kohl, N., 2005. Automatic feature selection in neuroevolution, in: Proceedings of the 7th annual conference on Genetic and evolutionary computation, pp. 1225–1232.

Witten, I.H., Frank, E., 2005. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.

Wolter, K., Timlin, M., 1998. Measuring the strength of enso events: How does 1997/98 rank? Weather 53, 315–324.

Wolter, K., Timlin, M.S., 2011. El niño/southern oscillation behaviour since 1871 as diagnosed in an extended multivariate enso index (mei. ext). International Journal of Climatology 31, 1074–1087.

World Bank, 2019. Sustainable energy for all (se4all)(database). URL: https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?end=2017&locations=ET&name_desc=false&start=2000&view=chart.

Wright, R., Loscalzo, S., Yu, L., 2012. Embedded incremental feature selection for reinforcement learning. Technical Report. AIR FORCE RESEARCH LAB ROME NY INFORMATION DIRECTORATE.

Wurbs, R.A., 1993. Reservoir-system simulation and optimization models. Journal of water resources planning and management 119, 455–472.

Xu, H., Jagannathan, S., 2014. Neural network-based finite horizon stochastic optimal control design for nonlinear networked control systems. IEEE transactions on neural networks and learning systems 26, 472–485.

Yakowitz, S., 1982. Dynamic programming applications in water resources. Water resources research 18, 673–696.

Yang, H., Reichert, P., Abbaspour, K.C., Zehnder, A.J., 2003. A water resources threshold and its implications for food security.

Yang, Y., Pedersen, J.O., 1997. A comparative study on feature selection in text categorization 97, 412–420.

Yeh, W.W.G., 1985. Reservoir management and operations models: A state-of-the-art review. Water resources research 21, 1797–1818.

Yihdego, Z., Rieu-Clarke, A., Cascão, A.E., 2017. The Grand Ethiopian Renaissance Dam and the Nile Basin: implications for transboundary water cooperation. Routledge.

Yuan, X., Wood, E.F., Ma, Z., 2015. A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development, in: Interdisciplinary Reviews: Water, pp. 523–536.

Zanchettin, D., Franks, S., Traverso, P., Tomasino, M., 2008. On enso impacts on european wintertime rainfalls and their modulation by the nao and the pacific multi-decadal variability described through the pdo index. International Journal of Climatology: A Journal of the Royal Meteorological Society 28, 995–1006.

Zaniolo, M., Giuliani, M., Castelletti, A., 2019. Data-driven modeling and control of droughts. IFAC-PapersOnLine 52, 54–60.

Zaniolo, M., Giuliani, M., Castelletti, A., 2020. Neuro-evolutionary direct policy search for multi-objective optimal control. IEEE transactions on neural networks and learning systems (under review) .

Zaniolo, M., Giuliani, M., Castelletti, A., Pulido-VelÃ zquez, M., 2018a. Raw and processed hydrometeorological variables of Jucar river basin for feature selection. URL: https://doi.org/10.5281/zenodo.1185084, doi:doi: 10.5281/zenodo.1185084.

Zaniolo, M., Giuliani, M., Castelletti, A.F., Pulido-Velazquez, M., 2018b. Automatic design of basin-specific drought indexes for highly regulated water systems. Hydrology and Earth System Sciences 22, 2409–2424.

Zarfl, C., Berlekamp, J., He, F., Jähnig, S.C., Darwall, W., Tockner, K., 2019. Future large hydropower dams impact global freshwater megafauna. Scientific Reports 9, 1–10.

Zarfl, C., Lumsdon, A.E., Berlekamp, J., Tydecks, L., Tockner, K., 2015. A global boom in hydropower dam construction. Aquatic Sciences 77, 161–170.

Zargar, A., Sadiq, R., Naser, B., Khan, F.I., 2011. A review of drought indices. Environmental Reviews 19, 333–349.

Zatarain, J.S., Reed, P.M., Herman, J.D., Giuliani, M., Castelletti, A., 2016. A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. Advances in water resources 92, 172–185.

Zhang, T., 2009. Adaptive forward-backward greedy algorithm for sparse learning with linear models, in: Advances in Neural Information Processing Systems, pp. 1921–1928.

Zhang, Y., Erkyihum, S.T., Block, P., 2016. Filling the gerd: evaluating hydroclimatic variability and impoundment strategies for blue nile riparian countries. Water international 41, 593–610.

Zhao, T., Cai, X., Yang, D., 2011. Effect of streamflow forecast uncertainty on real-time reservoir operation. Advances in water resources 34, 495–504.

Zhao, T., Zhao, J., Lund, J.R., Yang, D., 2014. Optimal hedging rules for reservoir flood operation from forecast uncertainties. Journal of Water Resources Planning and Management 140, 04014041.

Zhu, H., Jin, Y., 2019. Multi-objective evolutionary federated learning. IEEE transactions on neural networks and learning systems .

Ziervogel, G., Johnston, P., Matthew, M., Mukheibir, P., 2010. Using climate information for supporting climate change adaptation in water resource management in south africa. Climatic Change 103, 537–554.

Zilberstein, S., Washington, R., Bernstein, D.S., Mouaddib, A.I., 2002. Decision-theoretic control of planetary rovers, in: Advances in Plan-Based Control of Robotic Agents. Springer, pp. 270–289.

Zimmerman, B.G., Vimont, D.J., Block, P.J., 2016. Utilizing the state of enso as a means for season-ahead predictor selection. Water resources research 52, 3761–3774.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Da Fonseca, V.G., 2003. Performance assessment of multiobjective optimizers: An analysis and review. IEEE Transactions on evolutionary computation 7, 117–132.

Zoppoli, R., Sanguineti, M., Parisini, T., 2002. Approximating networks and extended ritz method for the solution of functional optimization problems. J. Optim. Theory Appl. 112, 403–439.