IFAC Journal of Systems and Control 19 (2022) 100185

Contents lists available at ScienceDirect

IFAC Journal of Systems and Control

journal homepage: www.elsevier.com/locate/ifacsc

Learning-based hierarchical control of water reservoir systems[☆]

Pauline Kergus[a,*,1], Simone Formentin[b], Matteo Giuliani[b], Andrea Castelletti[b]

[a] Department of Automatic Control, Lund University, Sweden
[b] Dipartimento di Elettronica, Informazione e Biongegneria, Politecnico di Milano, Italy

ARTICLE INFO

Article history:
Received 5 July 2021
Received in revised form 17 October 2021
Accepted 6 January 2022
Available online 14 January 2022

Keywords:
Water reservoir
Control application
Data-driven control
Predictive control
Constrained control

ABSTRACT

The optimal control of a water reservoir system represents a challenging problem, due to uncertain hydrologic inputs and the need to adapt to changing environment and varying control objectives. In this work, we propose a real-time learning-based control strategy based on a hierarchical predictive control architecture. Two control loops are implemented: the inner loop is aimed to make the overall dynamics similar to an assigned linear model through data-driven control design, then the outer economic model-predictive controller compensates for model mismatches, enforces suitable constraints, and boosts the tracking performance. The effectiveness of the proposed approach is illustrated on an accurate simulator of the Hoa Binh reservoir in Vietnam. Results show that the proposed approach outperforms stochastic dynamic programming.

1. Introduction

On a global scale, population and economic growth result in an increasing energy demand. At the same time, climate change and growing populations pressure freshwater resources (McDonald, et al., 2011). Hydropower dams could constitute a response to these big energy challenges since they provide a wide range of benefits such as reduced fossil fuel consumption, irrigation and urban water supply. Yet, building new dams implies substantial financial and environmental costs and, as explained in Ansar, Flyvbjerg, Budzier, and Lunn (2014), many large storage projects worldwide are failing to produce the level of benefits that would economically justify their development. Therefore, operating existing infrastructures more efficiently, rather than planning new ones, represents an opportunity for these facilities to fulfill their potential, which is a critical challenge. New operating policies should be able to adapt release decisions to uncertain hydrologic conditions and to evolving objectives such as growing water demands (Soncini-Sessa, Weber, & Castelletti, 2007).

This problem has received much attention from different research fields since the 70s but, as explained in Castelletti, Pianosi, and Soncini-Sessa (2008), it remains challenging for different reasons: water reservoirs are highly non-linear and multiple and conflicting interests are at stakes, usually formulated as non-linear and strongly asymmetric objective functions. In addition, the system is affected by strong uncertainties, such as the inflow of water, which cannot be neglected. In the literature, dynamic programming (DP) and its stochastic extension (SDP) are among the most widely used methods for designing optimal operating policies for water reservoirs, see Castelletti et al. (2008). In practice, the use of SDP is limited by the curse of dimensionality (Bellman, 1957) and the curse of modeling (Powell, 2007). For these reasons, new control strategies are needed. Such methods should:

(i) account for the partially unmodeled and complex behavior of water reservoirs. Indeed, the hydrological conditions affect the system in a nonlinear nonparametric way;

(ii) adapt to evolving hydrological conditions. Traditional control policies, such as the ones based on DP, rely on past observations and need to be fully re-designed to account for new disturbance profiles;

(iii) balance conflicting objectives, e.g., maximization of hydropower production and minimization of flood damages.

To overcome these challenges, we propose to control water reservoirs using a hierarchical learning-based approach as in Piga, Formentin, and Bemporad (2017), Polverini, Formentin, Merzagora, and Rocco (2019) and Piga, Forgione, Formentin, and Bemporad (2019). It combines data-driven control to tackle **(i)**, and an online model-based controller in order to handle **(ii)** and **(iii)**. The outline of the proposed strategy is as follows. First, a

☆ This project was partially supported by the Italian Ministry of University and Research under the PRIN'17 project "Data-driven learning of constrained control systems", contract no. 2017J89ARP, and by Regione Lombardia, Italy under the project ADDApt.

* Corresponding author.
E-mail addresses: pauline.kergus@control.lth.se (P. Kergus), simone.formentin@polimi.it (S. Formentin), matteo.giuliani@polimi.it (M. Giuliani), andrea.castelletti@polimi.it (A. Castelletti).

1 This author is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No 834142 (ScalableControl).

https://doi.org/10.1016/j.ifacsc.2022.100185

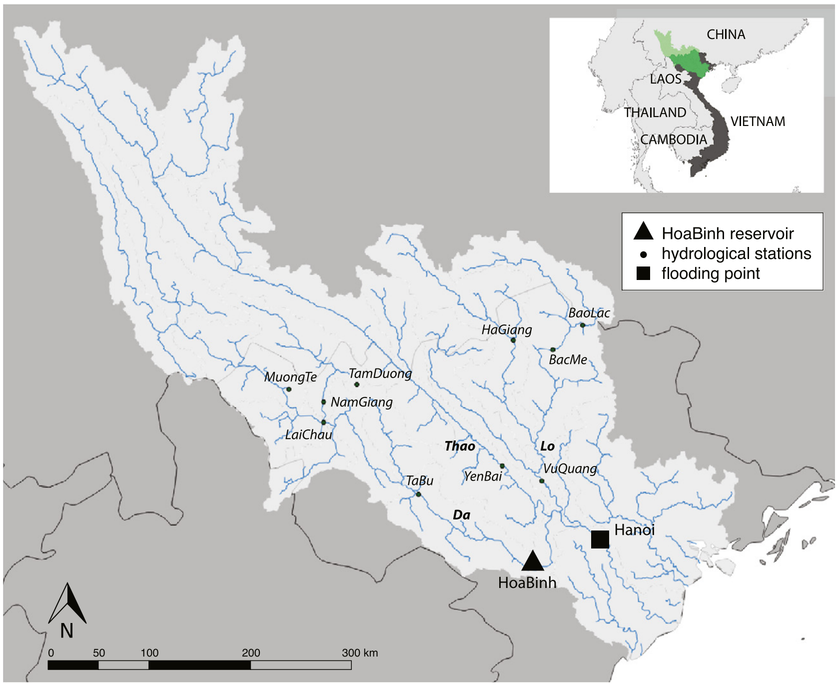


Fig. 1. Map of the Red River Basin.

parametric controller is designed to match some desired closed-loop behavior using the Virtual Reference Feedback Tuning (VRFT) approach (Campi, Lecchini, & Savaresi, 2002), which is a one-shot data-driven control technique based on open-loop time-domain measurements. A Model Predictive Controller (MPC) is then used as a reference governor, enabling to enforce constraints on the water release and to optimize performances online. These two techniques, VRFT and MPC, are indeed complementary as pointed out in Boeira, Bordignon, Eckhard, and Campestrini (2018), and the proposed strategy allows to combine their benefits. The resulting hierarchical approach shows a number of advantages over traditional SDP. First, its online feature allows one to adapt to evolving hydrological conditions thanks to the incorporation of short-term disturbance forecasts. In addition, it does not require a full mathematical description of the reservoir or the disturbances profile.

The main contribution of this paper lies in the application of this hierarchical control strategy to a hydropower system. This also leads to methodologic novelties: the overall control objective is formulated in an economic MPC fashion (see Faulwasser, Grüne, Müller, et al. (2018)) and time-varying constraints are considered in the design of the outer control. In addition, the proposed method presents a number of tuning knobs, whose practical selection is also discussed throughout the paper. This work is illustrated on the Hoa Binh water reservoir system in Vietnam, for which a simulation model is available along with historical data. The objective is to control the operations of this water reservoir to maximize hydroelectricty production and minimize flooding while taking into account the constraints of the system and the hydrologic inputs. The whole paper demonstrates the effectiveness of the proposed approach on this case study, and compares its performance with classical SDP as used in Castelletti et al. (2008).

The remainder of the paper is organized as follows. First, in Section 2, the Hoa Binh case study is presented, its model is detailed and the control objectives are introduced. The classical SDP approach and its challenges are recalled in order to motivate the present work. The available data, used in this work, is also listed and separated into a training and validation data-set. Section 3 describes the proposed hierarchical approach, starting with the data-driven design of the inner-loop based on VRFT (Campi et al., 2002), and then introducing the outer economic MPC loop, which plays the role of a reference governor. Simulation results are then reported in Section 4 showing that, in comparison with SDP, the proposed approach allows to obtain better trade-offs. These good results are demonstrated on both the training and validation time-periods. Concluding remarks, along with outlooks for future research, are presented in the last section.

2. Preliminaries

2.1. The Hoa Binh case study

The Hoa Binh reservoir is one of the largest reservoir in Vietnam, characterized by a surface area of about 198 km^2 and an active storage capacity of about 6 billion m^3. The reservoir is located along the Da River (see Fig. 1), which is the main tributary of the Red River, with this latter being the second largest river basin of Vietnam accounting for a total area of about 169,000 km^2. The dam is connected to a power plant equipped with eight turbines, for a total design capacity of 1,920 MW, which guarantees a large share of the national electricity production. Moreover, the dam operation contributes to the control of downstream floods, particularly in the highly densely populated capital city Hanoi.

The Hoa Binh system is here modeled as a discrete-time, periodic, non-linear, stochastic Markov Decision Process (MDP) by combining conceptual and data-driven models. This model, initially presented in Castelletti, Pianosi, Quach and Soncini-Sessa (2012), is recalled in this paragraph. The Hoa Binh dynamics is represented by the mass balance equation of the water volume s_t stored in the reservoir, affected by a stochastic disturbance q^D_{t+1} (namely, the inflow to the reservoir in the time interval $[t, t+1)$) and the water release r_{t+1}:

$$s_{t+1} = s_t + q^D_{t+1} - r_{t+1} \tag{1}$$

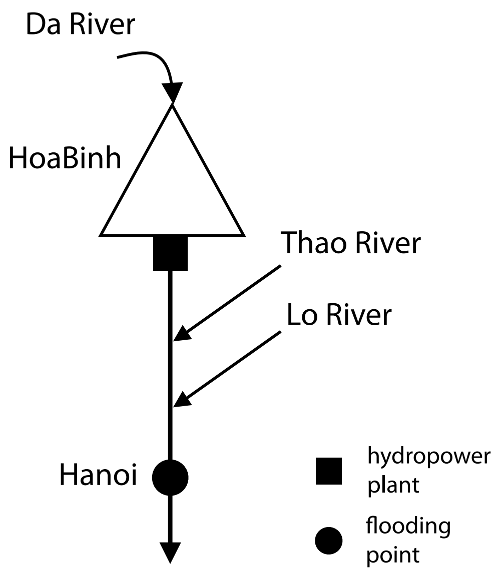


Fig. 2. Schematic representation of the model.

The nonlinear dynamics of the reservoir are due to the release function, which determines the actual release r_{t+1} as

$$r_{t+1} = f(s_t, u_t, q^D_{t+1}) \tag{2}$$

where u_t is the water release decision (i.e., the control signal). The minimum and maximum releases that can be produced in the time interval $[t, t+1)$ starting from s_t with inflows q^D_{t+1} (by keeping all the dam's gates completely closed and completely open, respectively) are embedded in the model through the release function f, thus guaranteeing the feasibility of the designed solutions (Soncini-Sessa et al., 2007).

Remark 1. In the adopted notation, the time subscript of a variable indicates the instant when its value is deterministically known. The reservoir storage s_t is observed at time t, whereas the inflow and release have subscript $t+1$ as they depend on the realization of the stochastic process in the time interval $[t, t+1)$.

The routing of the water released is represented schematically in Fig. 2. The disturbances affecting the system are the flows q^D, q^T and q^L, from the Da, Thao and Lo rivers respectively. The Thao and Lo rivers join the Da river 50 km and 60 km from the Hoa Binh dam respectively and affect the water level in Hoa Binh. The Hoa Binh dam itself is situated on the Da river, which is responsible for the reservoir inflow. In Pianosi, Castelletti, and Lovera (2012), a feedforward neural network has been trained on the basis of historical data. It is used in this work to simulate the water level in the city of Hanoi as a function of the Hoa Binh release r_{t+1} along with the natural discharges of the Thao (q^T_{t+1}) and Lo (q^L_{t+1}) rivers.

Considering the model described in this paragraph, the overall objective when designing a control strategy is to anticipate the disturbance flows to avoid floodings in Hanoi by reducing the water release from the reservoir, while releasing as much water as possible in order to produce hydropower in the Hoa Binh power dam. These objectives are obviously conflicting and they will be formalized in the next subsection, before introducing the traditional policy design based on dynamic programming.

2.2. Control objectives

In this work, two conflicting interests are at stakes when controlling the Hoa Binh operation: hydropower production should be maximized and floodings in Hanoi should be minimized. They are described by the following objective formulations (Giuliani, Castelletti, Pianosi, Mason, & Reed, 2016):

- **Hydropower production**: the energy production (kWh) for the period $[t, t+1)$ is defined by

$$J^H_{t+1} = \eta g \gamma_w \bar{h}_t q^{Turb}_{t+1} \cdot 10^{-6} \tag{3}$$

where $g = 9.81$ (m/s^2) is the gravitational acceleration, $\gamma_w = 1000$ (kg/m^3) is the water density. The turbined flow q^{Turb}_{t+1} is equal to the water release r_{t+1} as long as it lies between a feasible range for the turbine: $q^{Turb} \in [38, 2360]$m^3/s. The hydraulic head $\bar{h}_t$ in the reservoir is equal to the reservoir level h_{up} minus the tailwater level h_{down}. The reservoir level h_{up} is a function of the water storage s and depends on the reservoir geometry. This function is known only for a finite set of storage values. The tailwater level h_{down} is given by a function depending on the water release r. Finally, η is the turbine efficiency (which depends on the hydraulic head $\bar{h}_t$);
- **Flood control**: the average excess level J^F (cm^2/day) in Hanoi for the period $[t, t+1)$ is defined by

$$J^F_{t+1} = max(h^{Hanoi}_{t+1} - \bar{h}, 0)^2 \tag{4}$$

where $\bar{h}$ = 950 cm is the flooding threshold and h^{Hanoi}_{t+1} is the water level in Hanoi estimated by the routing model.

2.3. Prior control approach: dynamic programming

Traditionally, Pareto optimal control policies of the Hoa Binh reservoir are obtained offline for a considered time-period $t \in [0, T]$ via Dynamic Programming (DP) (Bellman, 1957). The multi-objective problem is reformulated as a single-objective problem, defined as follows:

$$\min_p -\alpha \frac{1}{T} \sum_{t=0}^{T-1} J^H_{t+1} + (1-\alpha) \frac{1}{T} \sum_{t=0}^{T-1} J^F_{t+1}, \tag{5}$$

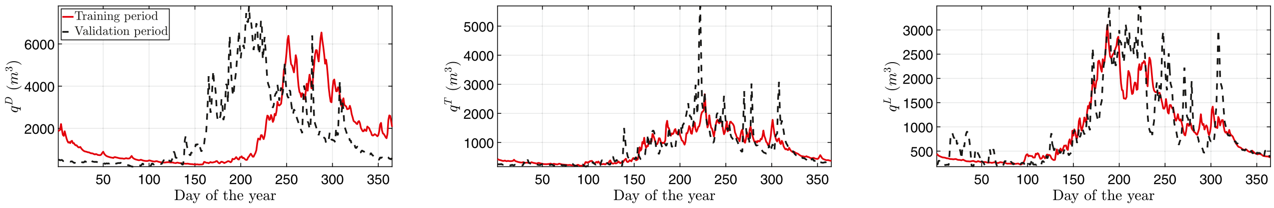
subject to the dynamics of Hoa Binh reservoir, see Eq. (1)–(2). The ponderation α balances the two competing objectives (3) and (4), which are averaged over the whole considered period. The Pareto optimal policies are obtained by solving Problem (5) for different values of α ranging between 0 and 1.

The disturbances $[q^D_{t+1}, q^T_{t+1}, q^L_{t+1}]$ can be considered to be deterministic over the considered period, in which case the resolution of the problem is referred to as Deterministic Dynamic Programming (DDP). The corresponding performance should be regarded to as ideal, as the optimization is performed *a-posteriori* with full knowledge available. As a consequence, the resulting policies cannot be applied to other time-periods for which the disturbances would not be exactly the same.

In Giuliani et al. (2016), the disturbances are considered to be stochastic, which constitutes a more realistic approach. In this case, the single-objective problem are solved through Stochastic Dynamic Programming (SDP). The Hoa Binh inflow q^D_{t+1} (from the Da river) is then described as a log-normal distribution $\mathcal{L}_t$:

$$q^D_{t+1} \sim \mathcal{L}_t, \tag{6}$$

while the two tributaries q^T_{t+1} and q^L_{t+1} (Thao and Lo rivers) are produced by a simple linear model that describes the spatial



(a) Inflow q^D to the Hoa Binh reservoir. (b) Inflow q^T from the Thao river. (c) Inflow q^L from the Lo river.

Fig. 3. Yearly disturbance pattern for the training and validation periods: by evaluating the patterns, one may check whether the proposed control strategy can be generalized to different hydroclimatic conditions.

correlation across the rivers, and the model residuals are then represented using 2 normal distributions $\mathcal{N}^T$ and $\mathcal{N}^L$:

$$\begin{aligned} q^T_{t+1} &= \alpha^T q^D_{t+1} + \varepsilon^T_{t+1}, \quad \varepsilon^T_{t+1} \sim \mathcal{N}^T, \\ q^L_{t+1} &= \alpha^L q^D_{t+1} + \varepsilon^L_{t+1}, \quad \varepsilon^L_{t+1} \sim \mathcal{N}^L. \end{aligned} \tag{7}$$

The stochastic disturbance models of the inflows, namely the three probability distributions $\mathcal{L}_t$, $\mathcal{N}^T$ and $\mathcal{N}^L$, as well as the coefficients α^T and α^L, were calibrated on the historical observations from the time-period 1962–1969. As long as the disturbances follow the considered stochastic model, the SDP policies can be employed. However, when the inflows profiles drift away from (6)–(7), as it has been often the case over the last decade because of climate changes, the SDP policies are no longer optimal. It is then necessary to collect a sufficient amount of data in order to update the stochastic disturbance models before applying SDP once again. To that extent, SDP does not constitute a flexible enough solution to face evolving hydrological conditions.

Furthermore, if in principle, SDP can solve problem (5) under relatively mild assumptions (Castelletti, Pianosi & Soncini-Sessa, 2012), the use of SDP is limited in practice. Indeed, the application of SDP in large-scale control schemes is constrained by the well-known *curse of dimensionality* (Bellman, 1957). Besides, SDP is also constrained by the *curse of modeling* (Tsitsiklis & Van Roy, 1996), as any input of the control policy must be explicitly modeled, and the *curse of multiple objectives* (Powell, 2007) as the generation of the full set of Pareto optimal solutions factorially scales with the growth in the number of the objectives. These three curses, along with the challenges related to the adaptation to variable hydrologic regimes, motivate the search for scalable and more flexible solutions.

2.4. Available data

Before moving on to the proposed control strategy, it is worth recalling that the proposed design relies on different types of data, listed hereafter:

- **Historical data** from the Hoa Binh reservoir: the daily water inflows from the Da, Thao and Lo rivers are known from 1959 to 2008. In this paper, two time periods are considered. The first one is 1962–1969, as selected in Giuliani et al. (2016) as it comprises normal, wet, and dry years. It is used as a training data-set, while the second period is 2007–2008 and is used as a verification set for the proposed control strategy. As it is the most recent data available, it allows to test the different approaches under different hydroclimatic conditions, see Fig. 3. The data is sampled daily;
- The system is partially unmodeled: the **release function** f and the water level h_{up} in the reservoir are not parametrized. Their value is known only on a grid of values of reservoir storage and inflows from the Da river;
- The **optimal policy** obtained through DDP with $\alpha = 0.05$ for the first time period (1962–1969) is denoted $\{u^\star, s^\star\}$ and embeds the historical data from the same period. It allows to compute the ideal mean annual behavior $\{u^c, s^c\}$, which is used in the proposed approach to learn a low-level controller through data-driven design.

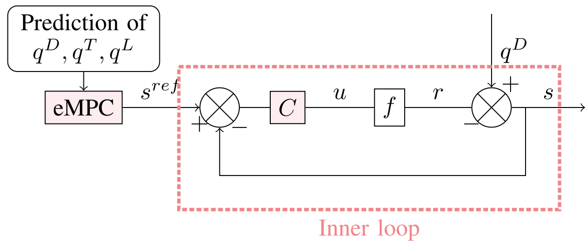


Fig. 4. The proposed control architecture: the inner-loop is designed to simplify the model dynamics for the outer reference governor, while the latter consists of an economic model predictive controller generating s^{ref}.

3. Hierarchical control system design

This paper aims at proposing an online data-driven strategy based on economic Model Predictive Control (eMPC), to operate the water reservoir system described in the previous section. The motivation behind this choice is that eMPC would enable to handle the uncertainty related to the evolution of the hydrological inputs. Indeed, eMPC also allows one to include short-term predictions of the disturbances, meaning that changes in the disturbances profiles can be taken into account. The selection of the best prediction strategy to obtain such forecasts is outside the scopes of this paper. Traditional model-based eMPC would allow us to enforce path constraints like in the classical SDP approach, ensuring that the obtained policy is feasible. However, the complexity and non-linearity of water reservoir systems constitutes a major obstacle to the implementation of such strategy as the online resolution of the associated optimization problem would represent a significant computational burden.

To that extent, the hierarchical data-driven approach introduced in Piga et al. (2017) allows us to overcome this issue while still offering the benefits of an online strategy. Indeed, the first step consists in reducing the complexity of the system dynamics by closing an inner loop through a model-free feedback controller: the design of such a low-level controller C allows to assimilate the inner-loop to a simpler linear model, which can be used as a prediction model in an outer-loop eMPC. This whole hierarchical approach, represented in Fig. 4, is detailed in this section.

3.1. Data-driven inner-loop design

To start with, a low-level controller C is designed to reduce the complexity of the system's behavior and to transform it into a known linear inner-loop model. The system is described by the balance equation (1) and by the nonlinear release function

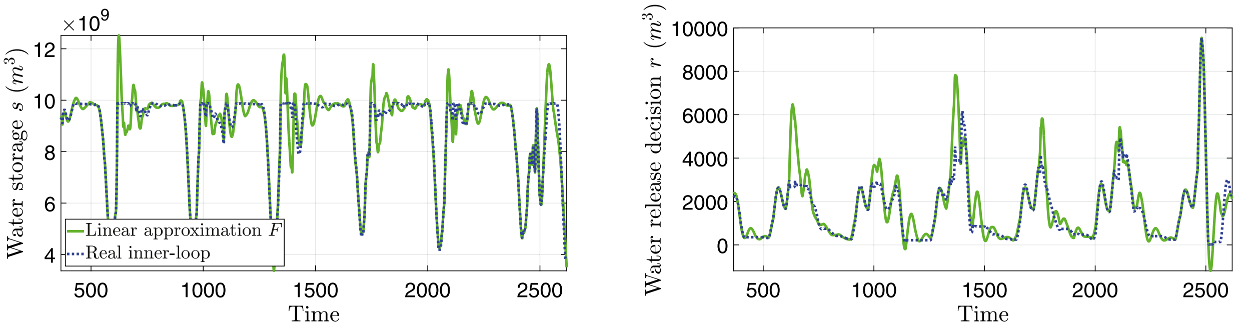


Fig. 5. Performance of the inner-loop and linear approximation F of the inner-loop, see Eq. (13). The water release decision u (left) and water storage s (right) are obtained by simulating the true inner-loop and its linear approximation over the period 1962–1969. The signal s^{ref} is taken as a desirable water storage signal, obtained by solving the optimal control problem through DDP over the same period.

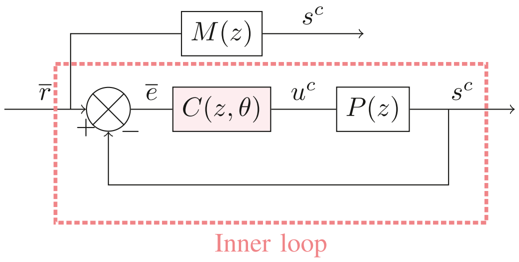


Fig. 6. Inner-loop design using VRFT: M is the reference model, P is the system to be controlled and C represents the controller to be tuned on the basis of the signals u^c and s^c. The signals $\bar{r}$ and $\bar{e}$ are the virtual reference and virtual error respectively.

f. Since f does not have a parametrized expression, it motivates the use of a data-driven technique, such as the VRFT (Campi et al., 2002), to design the inner-loop controller.

Based on the available time-domain data $\{u^c, s^c\}$, a desired closed-loop behavior M and a controller structure $C(z,\theta)$, the objective of the VRFT is to find the controller parameters $\theta^\star$ such that the resulting closed-loop is as close as possible to the reference model M, see Fig. 6. The key idea is the computation of the virtual reference signal

$$\bar{r}_t = M^{-1}(z)s^c_t, \tag{8}$$

as the reference that would feed the loop if the complementary sensitivity function was exactly M. Like that, the ideal controller (i.e., the one achieving M in closed-loop) can be computed as the system producing u^c when fed by the virtual error $\bar{e} = \bar{r} - s^c$. Since the virtual error and the virtual reference are computed off-line based on the available dataset, the controller can be retrieved by solving the following *one-shot* optimization problem:

$$\theta^\star = \arg\min_{\theta} \frac{1}{N}\sum_{t=1}^{N}(u^c_t - C(z,\theta)\bar{e}_t)^2. \tag{9}$$

In practice, a linear parameterization of the controller $C(\theta)$ (like a PID) leads to a simple quadratic problem that can be solved via least squares formulas.

In this work, the cyclostationary signals u^c (release decision) and s^c (storage), which represent the annual mean behavior obtained from the DDP ideal policy for 1962–1969, are used as the dataset for the identification of the controller. The desired closed-loop behavior M is defined as

$$M(z) = \frac{0.2z^{-1}}{1-0.8z^{-1}}. \tag{10}$$

Unlike the classical VRFT framework, here the reference model does not need to represent the desired closed-loop performance: this will be handled by the outer-loop eMPC controller. The reference model only needs to be practically achievable, which is most likely the case when specifying a low-performance closed-loop behavior M.

In the present case, the inner-loop controller has a PID structure:

$$C(z,\theta) = \theta_1 + \frac{\theta_2}{1-z^{-1}} + \theta_3(1-z^{-1}) \tag{11}$$

and the coefficients obtained through the VRFT procedure are

$$\theta^\star = 10^{-6}\,[-0.4439 \; -0.1063 \; -0.1898]^T. \tag{12}$$

The purpose of the inner-loop controller C is to assimilate the inner-loop to a simple linear model, which will be used later for prediction in the outer-loop to simplify the eMPC problem. To that extent, the nonlinear term f of the system (Fig. 4) is neglected, i.e. $u \approx r$, which leads to the linear model F approximating the inner-loop behavior:

$$F: \begin{cases} s(z){=}z^{-1}s(z) + T_s(q^D(z) - z^{-1}u(z)) \\ u(z){=}C(z)(s^{ref}(z) - s(z)) \end{cases}. \tag{13}$$

Fig. 5 compares the water-storage signals that are obtained using the inner-loop linear approximation F or by simulating the real inner-loop visible on Fig. 4, both fed with the reference signal $s^{ref} = s^\star$. The signal $s^\star$ is found through DDP when solving the optimal control problem for the period 1962–1969. The results indicate that the linear model F can reasonably be used to predict the inner-loop behavior. Even though the prediction model is not perfect because of the nonlinear part f of the true inner-loop. Since f corresponds to the physical limitations of the dam in terms of water release, this aspect can be handled in the outer-loop through constraints enforcement.

Finally, this simulation highlights the fact that the inner-loop alone is not sufficient for this application, for two reasons. First, the choice of a practically achievable and low performance reference model M as in (10) implies that the resulting inner-loop may be too slow to produce a sufficient level of benefit in terms of hydropower production and floodings' damages. Besides that, the reference s^{ref} to be tracked still needs to be defined. The optimal control strategy cannot be used for this purpose, as it requires to optimize over the whole considered period. The best option is therefore to perform an online optimization of the reference signal on a limited prediction horizon. Based on these considerations and on the simple description of the inner-loop obtained through this data-driven control design, an outer-loop model predictive controller is then designed as described in the next paragraph (see Fig. 6).

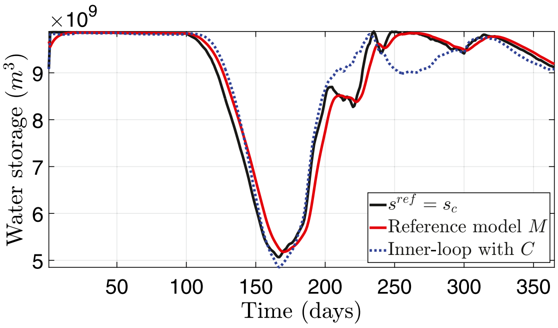


Fig. 7. Comparison of the response of the reference model M and the true inner-loop obtained with C, when the ideal water storage s^c is used as a reference.

Remark 2 (*Key Aspects for a Good Inner-Loop Design*).
The hierarchical strategy used in this work is relevant for the control of water reservoirs in general. However, in this section, the inner-loop design is specific to the Hoa Binh case study. To be applied to other cases, the following points should be considered:

- The **training data** should be rich enough to capture the dynamics of the considered system.
- The **choice of the reference model** is not as critical as in the original VRFT algorithm since the overall performance can obtain a boost – if possible – from the outer-loop. In this work, M only needs to be practically achievable in order to provide predictions of the output. To that extent, the first-order reference model has been chosen with a static gain equal to 1 (the inner loop is expected to compensate for static errors). Its time constant is selected to be of 5 days, which is more or less the open-loop one. In the end, it is always helpful to compare the response of the reference model M and that of the inner-loop like in Fig. 7, to see how well the inner-loop controller C matches the reference model. Note that it is still possible to iterate over several reference models in order to obtain a good inner-loop design if necessary.
- The **controller structure** can be changed. However, since the objective is not to have an ideal matching but a practical matching with the reference model, a simple controller structure, such as PID, is recommended (if it is not suitable, the reference model can be changed, by asking for less demanding performance as suggested in the previous point and in Piga et al. (2017)).

In the end, a good inner-loop design should provide a linear predictive model to be used in the outer-loop. This model does not need to be perfect as long as it is a good linear approximation of the main dynamics of the system, as it is the case here according to Fig. 5. Notice that the observed fluctuations are "in-line" with the active (controllable) storage of the Hoa Binh water reservoir system, which varies between a minimum at 3.8 billion m3 and a maximum of 10.89 million m3.

3.2. Model-based outer-loop design using economic MPC

While the inner-loop allows to have a simplified description of the problem, the outer-loop has a double objective: performance optimization and constraints enforcement. To that extent, an economic model predictive controller is implemented, solving the following problem at each time-step t:

$$\min_{s^{ref}} \quad -\alpha \frac{1}{N_p} \sum_{k=1}^{N_p} J^H_{t+k} + (1-\alpha) \frac{1}{N_p} \sum_{k=1}^{N_p} J^F_{t+k} \tag{14}$$

$$\text{s.t.} \quad \forall k = 1 \dots N_p :$$

$$\begin{cases} \quad x(t+k+1) \quad = A_F x(t+k) + B_F \begin{pmatrix} s^{ref}(t+k) \\ q^D(t+k) \end{pmatrix} \\ \begin{pmatrix} s(t+k+1) \\ u(t+k+1) \end{pmatrix} = C_F x(t+k) + D_F \begin{pmatrix} s^{ref}(t+k) \\ q^D(t+k) \end{pmatrix} \end{cases}$$

$$r^{min}(s_{t+k}, q^D_{t+k+1}) \leq u(t+k) \leq r^{max}(s_{t+k}, q^D_{t+k+1})$$

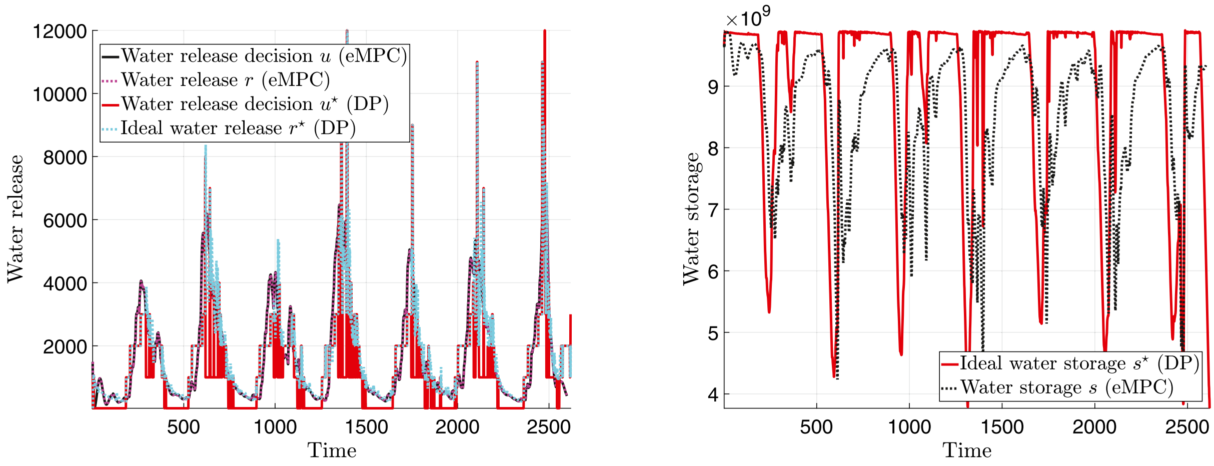
$$s^{min} \leq s(t+k) \leq s^{max}$$

First, the outer-loop controller acts as a reference governor for the inner-loop, determining the signal s^{ref} such that good performance is obtained over the prediction horizon N_p, both in terms of energy production and excess level in Hanoi. The major difference with respect to the DP approach is that the performance is here evaluated online on a shorter time window, starting at the current time step t and covering the prediction horizon N_p, instead of computing the optimal solution offline using the whole period information (in the present case the 1962–1969 period as in Giuliani et al. (2016)). As in the optimal control approach in (5), a ponderation $\alpha \in [0, 1]$ is used to balance the two conflicting objectives.

To evaluate the performances over the prediction horizon, the linear description F of the inner-loop, given in a state-space form (A_F, B_F, C_F, D_F), is used as prediction model (suitably scaled to avoid numerical errors). Its input are the reference water storage s^{ref}, which is the optimization variable, and the flow q^D from the Da river. The outputs of the inner-loop are the water storage s and the water release decision u.

Secondly, the outer-loop controller should enforce constraints on the release decision u so that the assumption $u \approx r$ made earlier holds. In addition, enforcing the constraints allow to avoid emptying the reservoir, which would happen otherwise since the flood damage objective is mostly equal to zero over time. To that extent, hard constraints are imposed on both outputs. The constraints are constant when it comes to the water storage s, with $s_{min} = 3.8Gm^3$ and $s_{max} = 9.9Gm^3$. The constraints on u are nonlinear and allow to take into account the nonlinearity f of the system: r_{min} and r_{max} represent the minimal and maximal water release, respectively. Like f, they depend on the current water storage s_t and the incoming flow q^D. These functions r_{min} and r_{max} are not parametrized: their value is known only for a given set of values of storage and flow. They are computed online according to the 2 nearest neighbors of the considered storage value and inflow from the Da river.

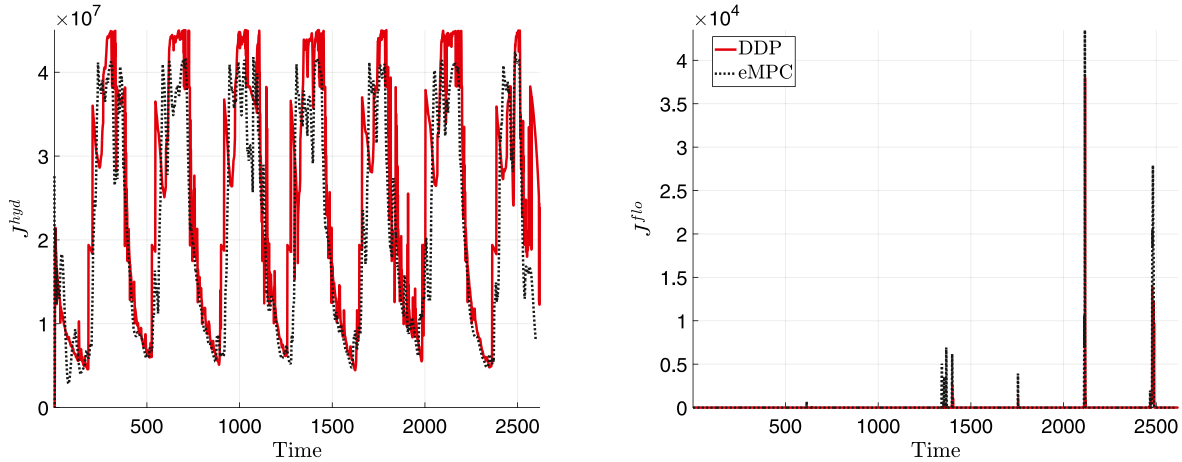
Remark 3 (*Recursive Feasibility of the eMPC*).
The constraints on the release embed the physical limitations of the dam. By definition of r_{min} and r_{max}, for any storage and inflow pair, the set of feasible water release u is always non-empty. In extreme cases, when the reservoir has a high (respectively low) level of water, this set is limited to one value corresponding to fully open gates (respectively fully closed). Storage constraints enforce that the water level in the reservoir stays between the desired bounds h_r^{min} and h_r^{max} despite the reservoir inflow. These bounds respectively correspond to the minimal and maximal storage s^{min} and s^{max}. In practice, the eMPC is recursively feasible for $s_0 \in [s^{min}, s^{max}]$ and for $0 \leq q^D \leq 25000$ m^3/day.



(a) Water release decision and actual water release.

(b) Water storage in the Hoa Binh reservoir.

Fig. 8. Trajectory of the controlled system when combining the inner-loop controller with the outer-loop economic MPC. The controlled system is simulated over the training period 1962–1969 and compared with the DDP approach.



(a) Daily average energy production.

(b) Daily average excess of water in Hanoi.

Fig. 9. Performance of the controlled system in terms of hydropower production and flood damages over the training period 1962–1969, compared with the DDP approach.

4. Simulation results

This section aims at evaluating the performance of the proposed strategy on the Hoa Binh case study, both on the training period (1962–1969) and the validation period (2007–2008). While the inner-loop design is performed using the VRFT toolbox (Carè, Torricelli, Campi, & Savaresi, 2019) in MATLAB®, the outer-loop optimization problem (14) is built with Yalmip (Löfberg, 2004).

4.1. Training 1962–1969

The behavior of the controlled system, including both the inner-loop and the outer loop controllers, is simulated over the period 1962–1969, using the available data regarding the incoming flows q^D, q^T and q^L. The performance of the controlled system can be easily appreciated when looking at the Pareto front given in Fig. 10, which represents the compromise between the average daily hydropower production J^{hyd} (on the y-axis) and the average daily excess of water in Hanoi J^{flo} (on the x-axis):

$$J^{hyd} = \frac{1}{T}\sum_{t=0}^{T-1} J^{H}_{t+1}, \qquad J^{flo} = \frac{1}{T}\sum_{t=0}^{T-1} J^{H}_{t+1}. \tag{15}$$

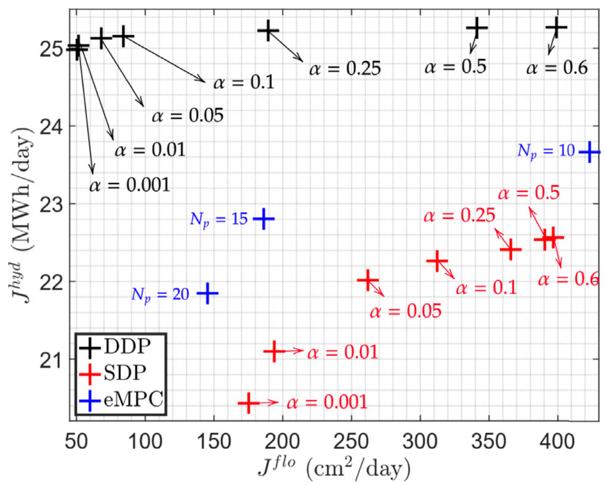
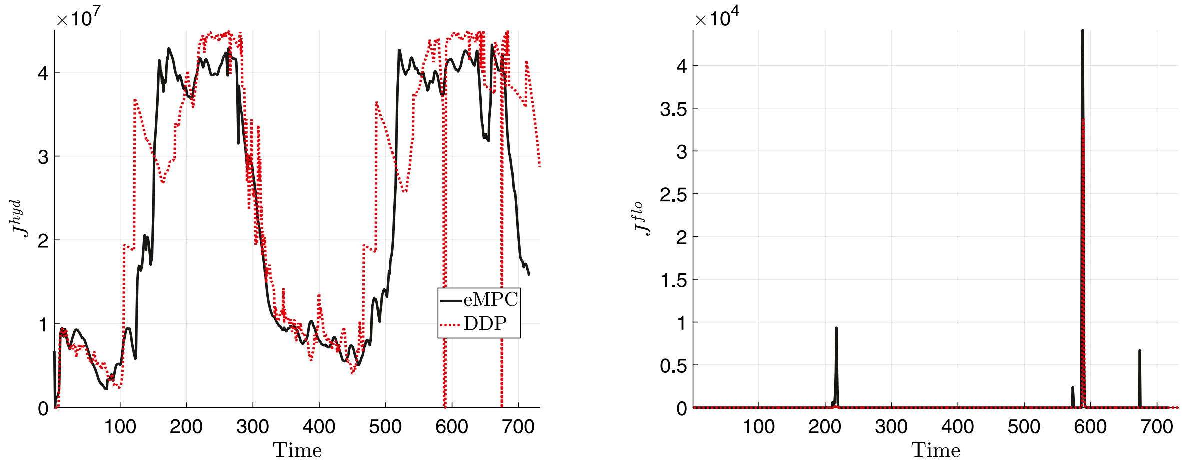


Fig. 10. Pareto front representing the performance trade-off between hydropower production and excess water level in Hanoi (1962–1969) for different policies: DDP, SDP and the proposed strategy.



(a) Daily average energy production.

(b) Daily average excess of water in Hanoi.

Fig. 11. Performance of the controlled system in terms of hydropower production and flood damages over the validation period 2007–2008, compared with the DDP approach.

Table 1
Performances of different control strategies during the validation period 2007–2008, compared with the results obtained on the training period 1962–1969 ($N_p = 15$ and $\alpha = 0.05$).

	1962–1969		2007–2008	
	J^{hyd}	J^{flo}	J^{hyd}	J^{flo}
VRFT + eMPC	**22.8049**	**186.1333**	**23.3126**	**212.9908**
SDP	22.015	261.872	23.149	274.186

Good performances on Fig. 10 therefore correspond to the top-left corner. The performance compromise is represented for the optimal control approaches DDP and SDP. The DDP approach can be conceived as an upper bound for the achievable performance. Unlike the proposed approach, the optimization is done *offline* for the whole considered period. According to Fig. 10, the ponderation $\alpha = 0.05$ (see (5)) is chosen as the one giving the best compromise between the conflicting objectives for both the DDP and SDP approaches as bigger values does improve significantly hydropower production while the daily excess level of water in Hanoi keeps on increasing.

The resulting water release and water storage are visible in Fig. 8. Panel (a) highlights that the constraints on the water release decision are satisfied so that $u_t = r_{t+1}$. The corresponding performance can be evaluated through the hydropower production and the excess amount of water in Hanoi, visible in Fig. 9. The results of the optimal control approach, solved through DDP over the same period, are also represented in both the figures and can be seen as an "upper" reference. In comparison, the proposed control strategy exhibits a slower reaction to the monsoon, mainly when refilling the reservoir every year, and does not reach the highest storage level of the DDP policy for the rest of the time.

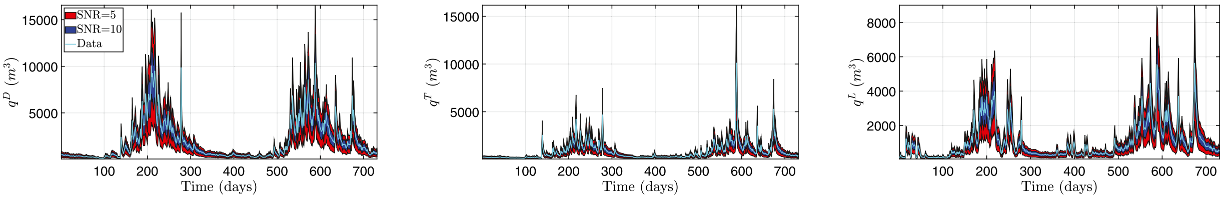
For the proposed approach, the performance compromise is computed over the training period 1962–1969 for three different values of the prediction horizon, $N_p = 10$, 15 and 20 days, while the ponderation in the cost function is selected as $\alpha = 0.05$ (as for the DDP and SDP approaches). Shortening the prediction horizon allows to increase the hydropower production, but it also increases the floodings in Hanoi. Indeed, a prediction horizon of $N_p = 10$ days is not sufficient to cover the time scale of the floodings, and makes it harder to anticipate this phenomenon by adjusting the release. On the other side, increasing the prediction horizon may lead to a significantly reduced hydropower production ($N_p = 20$ days) as the release is reduced to anticipate floodings occurring within a longer time-window. In addition, extending the prediction horizon increases the computational complexity. In the present case, $N_p = 15$ days constitutes a good performance trade-off and this value is chosen as the prediction horizon in the rest of this paper.

According to Fig. 10, the proposed approach performs better than SDP, which makes it a good candidate for its online implementation and use in different time periods.

Remark 4. The Pareto front of the DDP and SDP approaches correspond to different values of the ponderation α in (5) while, the proposed hierarchical approach, the tradeoff varies with the prediction horizons N_p. Indeed, in the proposed strategy, the ponderation α does not influence much the performance tradeoff for a given prediction horizon N_p. This is due to the fact that the excess level of water is mostly equal to zero by definition (4). In the future, it would be interesting to define another objective function for flood minimization purposes that would be more adequate on a short prediction horizon.

4.2. Validation 2007–2008

In order to demonstrate the ability of the proposed approach to handle different hydrologic conditions, a simulation is run for the validation period 2007–2008 with a prediction horizon of $N_p = 15$ days. The corresponding performance is visible in Fig. 11, along with the optimal one, computed by applying DDP for $\alpha = 0.05$ over the 2007–2008 period: as for the training period 1962–1969, the proposed approach allows us to obtain good performance in terms of hydropower production as compared to the optimal policy, but results in punctually higher levels of water in Hanoi. The proposed strategy is compared with the performance of the policy that was obtained offline through SDP for the training period 1962–1969. The average performances over both training and validation periods are given in Table 1 for the proposed approach (VRFT+eMPC), and for the SDP strategy with $\alpha = 0.05$ and trained on the period 1962–1969. First, the SDP policy is robust to the change of hydrological conditions between the training and validation period, as it takes uncertainty into account by considering the disturbances as stochastic processes. On the other side, the technique proposed in this paper handles the change of hydrologic conditions by performing online optimization over a limited prediction horizon. In the end, the



(a) Inflow q^D to the Hoa Binh reservoir. (b) Inflow q^T from the Thao river. (c) Inflow q^L from the Lo river.

Fig. 12. Considered noisy disturbances q^D, q^T and q^L over the validation periods: the noisy predictions are in the blue area for SNR=10 and in the red one for SNR=5, both area being centered on the real disturbance signal. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

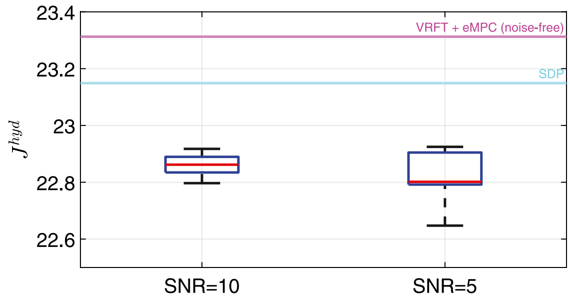
Table 2
Mean performance of the proposed strategy during the validation period 2007–2008 ($N_p = 15$ and $\alpha = 0.05$) with noisy disturbance predictions.

	J^{hyd}	J^{hyd}
VRFT + eMPC (noise-free)	23.3126	212.9908
VRFT + eMPC + SNR=10	22.8745	207.3910
VRFT + eMPC + SNR=5	22.8181	207.9877

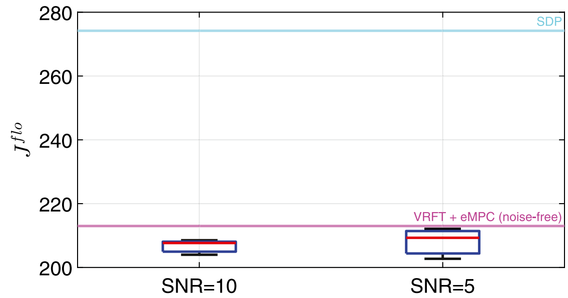
proposed online strategy performs the best over the validation period 2007–2008: it obtains both the highest daily average hydropower production and the lowest daily average excess level of water.

In practice, only predictions of the disturbances can enter the outer-loop economic predictive controller. Additional simulations have been run over the validation period 2007–2008, where random noise is added on the disturbances to mimick uncertain predictions over the prediction horizon. The noise is zero-mean and normally distributed, and two different SNRs (Signal to Noise Ratios) have been considered (SNR=10 and SNR=5). The corresponding noise levels are illustrated in Fig. 12. For each SNR, ten simulations are run. The mean performance is given in Table 2 and the results for each SNR are summed up in Fig. 13. It shows that noisy disturbance predictions tend to result in a slightly lower daily average hydropower, see Fig. 13(a), while the daily average excess level of water tends to slightly decrease, see Fig. 13(b). Even though SDP gives slightly better results in terms of hydropower production according to Table 1, the proposed approach outperforms SDP by maintaining a reasonable level of excess water in Hanoi despite the error in disturbance prediction.

Indeed, while these results are encouraging, the robustness of the eMPC approach to prediction errors needs further investigation. It is worth recalling that, by optimizing online on a shorter prediction horizon, the proposed approach makes it easier to update the prediction to climate change. Furthermore, notice that, while the distribution of the disturbances evolves, it is not necessary to determine a new disturbance distribution before rerunning SDP.



(a) Influence of noisy disturbance predictions on the average daily hydropower production J^{hyd}.



(b) Influence of noisy disturbance predictions on the average daily excess level of water J^{flo}.

Fig. 13. Performance obtained by the proposed strategy for different noise levels (SNR=5 and 10) applied to the disturbance predictions. Ten simulations have been performed for each SNR. The central red mark indicates the median, and the bottom and top edges of the blue box indicate the 25th and 75th percentiles, respectively. The black whiskers extend to the most extreme data points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusions

In this paper, a hierarchical data-driven control design strategy is proposed for water resources management, with a focus on the Hoa Binh reservoir case study, Vietnam. First, a linear controller is designed from data (with no use of the model of the system) to approximately assign a desired behavior to the inner-loop. The model of the inner loop is then used for the design of an outer economic model predictive control loop, which handles performance optimization and constraints enforcement. Compared to the resolution of the optimal control through traditional SDP tools, the proposed strategy leads to a better compromise between hydropower production and floodings in Hanoi for a prediction horizon of 15 days. The main strength of the proposed approach is its online nature and the corresponding ability to adapt to new environmental conditions.

Future research will focus on providing formal guarantees for the proposed schemes regarding stability, feasibility and robustness to disturbance uncertainties. Moreover, we will investigate the applicability of the approach to the other reservoirs in the Red River basin, to analyze their possible coordination.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ansar, A., Flyvbjerg, B., Budzier, A., & Lunn, D. (2014). Should we build more large dams? The actual costs of hydropower megaproject development. *Energy Policy*.

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

Boeira, E., Bordignon, V., Eckhard, D., & Campestrini, L. (2018). Comparing MIMO process control methods on a pilot plant. *Journal Of Control, Automation And Electrical Systems*.

Campi, M. C., Lecchini, A., & Savaresi, S. M. (2002). Virtual reference feedback tuning: a direct method for the design of feedback controllers. *Automatica*.

Carè, A., Torricelli, F., Campi, M. C., & Savaresi, S. M. (2019). A toolbox for virtual reference feedback tuning (VRFT). In: *2019 18th European Control Conference*.

Castelletti, A., Pianosi, F., Quach, X., & Soncini-Sessa, R. (2012). Assessing water reservoirs management and development in Northern Vietnam. *Hydrology And Earth System Sciences*.

Castelletti, A., Pianosi, F., & Soncini-Sessa, R. (2008). Water reservoir control under economic, social and environmental constraints. *Automatica*.

Castelletti, A., Pianosi, F., & Soncini-Sessa, R. (2012). Stochastic and robust control of water resource systems: Concepts, methods and applications. In *System identification, environmental modelling, and control system design*. Springer.

Faulwasser, T., Grüne, L., Müller, M. A., et al. (2018). Economic nonlinear model predictive control. *Foundations And Trends® In Systems And Control*.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., & Reed, P. M. (2016). Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. *Journal Of Water Resources Planning And Management*.

Löfberg, J. (2004). YALMIP : A Toolbox for Modeling and Optimization in MATLAB. *In Proceedings Of The CACSD Conference*, Taipei, Taiwan.

McDonald, R. I., Green, P., Balk, D., Fekete, B. M., Revenga, C., Todd, M., et al. (2011). Urban growth, climate change, and freshwater availability. *Proceedings Of The National Academy Of Sciences*.

Pianosi, F., Castelletti, A., & Lovera, M. (2012). Identification of a flow-routing model for the red river network. *IFAC Proceedings Volumes*.

Piga, D., Forgione, M., Formentin, S., & Bemporad, A. (2019). Performance-oriented model learning for data-driven MPC design. *IEEE Control Systems Letters*.

Piga, D., Formentin, S., & Bemporad, A. (2017). Direct data-driven control of constrained systems. *IEEE Transactions On Control Systems Technology*.

Polverini, M., Formentin, S., Merzagora, L., & Rocco, P. (2019). Mixed data-driven and model-based robot implicit force control: A hierarchical approach. *IEEE Transactions On Control Systems Technology*.

Powell, W. B. (2007). *Approximate dynamic programming: Solving the curses of dimensionality*. NJ: Wiley.

Soncini-Sessa, R., Weber, E., & Castelletti, A. (2007). *Integrated and participatory water resources management-theory*. Elsevier.

Tsitsiklis, J. N., & Van Roy, B. (1996). Feature-based methods for large scale dynamic programming. *Machine Learning*.