Hard-coupling water and power system models increases the complementarity of renewable energy sources

Rachel Koh \textsuperscript{a,}\textsuperscript{*}, Jordan Kern \textsuperscript{b}, Stefano Galelli \textsuperscript{a}

\textsuperscript{a} Pillar of Engineering Systems and Design, Singapore University of Technology and Design, Singapore 487372, Singapore
\textsuperscript{b} Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27607, United States of America

A R T I C L E  I N F O

Keywords:
Water-energy nexus
Renewable energy
Hard-coupling
Hydropower
Power systems

A B S T R A C T

The soft (one-way) coupling of water and power system models is the dominant approach for studying the impact of water availability on grid performance. Yet, such approach does not explicitly capture key dynamic interdependencies between the state of the grid and the operational decisions made at the water system level. Here, we address this gap and introduce a novel numerical modelling framework that hard-couples a multi-reservoir system model and a power system model. The framework captures two-way feedback mechanisms and thereby enables operational decisions to be made contingent upon the states of both the water and energy system. We evaluate the framework on a real-world case study based on the Cambodian grid. In light of the country’s plan to further decarbonize its grid, we tested the framework on three grid configurations—the as-is grid, and the grid with two different levels of installed solar capacity. Simulation experiments were run with and without feedback, while uncertainty in external forcings was explored through 1,000 stochastic time series of streamflow, solar production, and load. As demonstrated in our results, hard-coupling the water and energy systems reduces operating costs and CO\textsubscript{2} emissions while increasing the integration of renewables. Under favourable conditions (large reservoir inflow and low electricity demand), the system experienced a 44\% saving in annual operating costs and 53\% reduction of CO\textsubscript{2} emissions. A spatio-temporal analysis on the reservoir operations and transmission line usage reveals that the timing of the monsoon and interconnections between individual grid components also play significant roles in influencing the system’s responses to the hard coupling. Overall, simulation frameworks like this provide a modelling framework for testing management and planning solutions aimed to improve the performance of water-energy systems.

1. Introduction

Recognition of the interdependence between water and energy systems has motivated the research community to better understand the complex interactions within the water-energy nexus. Researchers often depend on numerical simulation models to represent these systems. Such models are particularly useful in evaluating potential system responses to important changes (e.g., integration of variable renewable energy into the grid) at both planning and operational stages. At the planning stage, longer-term simulations—usually employing monthly to annual time-steps over multi-decade horizons—facilitate the design of physical expansions of the grid or policies relevant to grid infrastructure \cite{1,2,3,4,5}. In contrast, operational modelling of power grids involves simulation of system dynamics at the hourly to daily time-step (e.g., \cite{6,7}). In this study, we focus on short-term operations of water-energy systems.

At the heart of the water-energy nexus, hydropower reservoirs play a paramount role. They fulfil multiple management objectives, including generating electricity for the power system, while being subject to varying hydro-meteorological conditions. The representation of hydropower reservoirs in numerical simulation models is therefore critical for a variety of applications. These applications include understanding the impact of droughts on power system performance \cite{7,8} or supporting the integration of renewables \cite{9}. Because reservoirs are a critical element of both water and energy systems, their study in a water-energy context typically builds on coupled models. This coupling process can be largely divided into two approaches: soft (one-way) and hard (two-way) \cite{10}. Soft coupling approaches generally model the water and power systems separately, with information being fed unilaterally from one model to the other. For example, hydropower availability could be first simulated using a hydrological model \cite{8,11}—or another ad hoc approach \cite{7,12,13}—and then passed as a time series input to the power system model simulating the dispatch of electricity. While useful, softly coupled models can

* Corresponding author.

E-mail address: rachel.koh@mymail.sutd.edu.sg (R. Koh).

https://doi.org/10.1016/j.apenergy.2022.119386

Received 10 February 2022; Received in revised form 12 May 2022; Accepted 29 May 2022

Available online 9 June 2022

0306-2619/© 2022 Elsevier Ltd. All rights reserved.
fail to capture complexities and nuances of water-energy systems. In particular, operating the water system components based only upon the hydrology and reservoir operating rules may erroneously assume that reservoir operators do not factor-in the state of the power system when making operational decisions. Yet, because of its limited complexity, the soft-coupling approach is, by far, the dominant approach used in research. Indeed, there are only a handful of studies that attempt to fully integrate two models—thereby capturing the feedbacks between one system to the other. For example, Ibanez et al. [14] hard-coupled RiverWare (a river routing model) with PLEXOS (a grid operations model) by iteratively optimizing the power system production costs while fulfilling the constraints of the water system. The application on the Western US showed favourable results in terms of production costs and curtailment of renewables. Recently, Stevanato et al. [15] demonstrated the importance of considering hydrological constraints on hydropower by integrating a reservoir model into Calliope, an energy modelling framework applied on the Zambezi River Basin in South Africa. Studying the possibility of storing hydropower when wind is available, Gebretsadik et al. [16] integrated an ad hoc reservoir system model with a single-node power system model and showed reduced reliance on thermal resources as a result.

Despite these recent studies, the full range of benefits from capturing these feedbacks is still poorly understood. Previous studies have focused exclusively on the generation mix as a key performance metric, but there are other benefits that have not been quantified yet. For example, transmission congestion can strongly affect the feasibility and value of hydropower production [17,18]. Enabling reservoir operational decisions to be contingent upon the states of both water and energy systems could help address this problem. Another important potential benefit is increasing the penetration of renewables, particularly wind and solar: given the intermittent nature of solar and wind generation, operationally flexible hydropower can be a complementary grid asset [19]. Hydropower reservoirs allow for storing and releasing water (and, therefore, potential energy) to buffer the variability of solar and wind. In addition, the benefits of hard-coupling water and energy models have until now been evaluated under deterministic conditions. Yet, it is widely known that uncertainties in load and variable renewable energy production can profoundly affect power system operations. It is thus important to incorporate these uncertainties in the further development and implementation of hard-coupled models. Furthermore, most grid representations in previous studies tend to be overly simple, which also limits our understanding of the importance of capturing feedbacks in water-energy systems models. For example, by modelling the electricity demand and supply of an entire region as a single-node, the role of important grid elements such as substations and transmission lines is not considered in the operational framework. Finally, we need to understand how the operations of non-hydropower grid components is affected by the representation of feedback mechanisms.

In this study, we contribute a novel numerical modelling framework for hard coupling water system and power system models. As highlighted in [20], grids with different kinds of renewable resources present multiple challenges, including spatial and temporal variability. For instance, reservoir releases that are made concurrently with high solar or wind generation may cause temporary “oversupply” of electricity, leading to forced curtailment or grid congestion issues [18]. To account for the effects of such external conditions, we introduce a reservoir re-operation model in our framework, a key element which gathers information from both the water and power systems. Specifically, the timing and amount of hydropower dispatched into the grid is compared to what is available to update water release decisions from hydropower reservoirs. This in turn allows the models more flexibility in responding to changes in external conditions. As we shall see, another important feature of the re-operation model is its capability to handle both single and multi-reservoir systems, thereby making it suitable for complex reservoir networks. Using a detailed representation of the Cambodian water-energy system, we demonstrate that the flexibility built into hard-coupled models can be exploited to improve the system’s performance, measured in terms of operating costs, CO₂ emissions, and penetration of renewables. We also explore how uncertainty in external conditions (i.e., inflow, load, and solar power production) influences the benefits derived from this hard coupling. Finally, we zoom into finer spatio-temporal scales and investigate how specific grid behaviours are modified by the use of a re-operation model.
2. Methods

2.1. Overview

We now provide an overview of our approach for creating a hard integration of water and energy models. Our integrated model is written in Python and consists of: (1) a reservoir system model; (2) a power system model; and (3) a re-operation model (blue, yellow and green box, respectively, in Fig. 1). Note that in state-of-the-art multi-sector models, it is typical to find only two components, i.e., a reservoir model and a power system model with a one-way transfer of information (usually hydropower availability) from the former to the latter (e.g., [8,11,21]). The added re-operation component enables the hard integration of the two models, allowing for a two-way feedback between them. The re-operation component evaluates the outcomes of the decisions made by the two primary models and provides a feedback (see dashed line leading back) to the respective models. Details of the reservoir, power system, and re-operation models are provided in Sections 2.2, 2.3, and 2.4, respectively.

2.2. Reservoir system model

Hydropower reservoir operations depend on several factors, including the network topography, hydro-meteorological processes, and site-specific operating objectives. Many models have been developed to simulate the behaviour of such reservoir networks (e.g., [22–25]). Although they come with different assumptions and require different inputs, in general, these models accomplish several things: they (1) route water through the river network; (2) solve the mass balance with a one-way transfer of information (usually hydropower availability) from the former to the latter (e.g., [8,11,21]). The added re-operation component enables the hard integration of the two models, allowing for a two-way feedback between them. The re-operation component evaluates the outcomes of the decisions made by the two primary models and provides a feedback (see dashed line leading back) to the respective models. Details of the reservoir, power system, and re-operation models are provided in Sections 2.2, 2.3, and 2.4, respectively.

2.3. Power system model

The power system model is a steady-state system model that is written in Python and consists of: (1) the power system model (blue box); (2) solar, wind, load, and thermal plant properties (grey box); and (3) other renewables (grey box). The model is run in a steady-state mode, with the objective of minimizing the total fuel consumption. The model calculates the total power demand and the available power from each power source. The difference between the two is then used to calculate the fuel consumption. The model then iterates until the total fuel consumption is minimized.

2.4. Re-operation model

The re-operation model is written in Python and consists of: (1) the re-operation model; (2) a decision variable for each reservoir; and (3) the total available hydropower. The model is run in a steady-state mode, with the objective of maximizing the total hydropower available. The model calculates the total storage in each reservoir and the amount of hydropower available. The difference between the two is then used to calculate the total hydropower available. The model then iterates until the total hydropower available is maximized.

The decision variable for each reservoir within the network is \( R_d \). This variable affects the operations of the reservoir and also influences the power system operations, since it determines the amount of hydropower available to the grid on a daily basis. To calculate initial values of \( R_d \), we use rule curves, a common approach adopted by reservoir operators [8,26]. Simply put, a rule curve represents the target storage (or level) that a reservoir should maintain on each calendar day. Following Dang et al. [24], we define the rule curves as piece-wise linear functions based on four parameters (Fig. 2): the initial values of \( R_d \), \( Q_{MEF,d} \), \( M_{MEF,d} \), and \( R_{max} \). The minimum and maximum water levels that a reservoir should reach within a year are \( (H_1 \) and \( H_2 \) and the time at which the two levels should be reached are \( T_1 \) and \( T_2 \)). For a given water system, these parameters can be reported by local agencies, retrieved from water storage data, or even inferred from satellite data [27]. The daily release decisions \( R_d \) from each reservoir are made to bring the actual storage as close to the target storage as possible, while being subjected to an upper

![Flowchart of the modelling framework](image)

Fig. 1. Modelling framework, comprising a reservoir model, power system model, and re-operation model. The dashed lines represent the feedback from the re-operation model to the two primary models. Symbols are introduced in Sections 2.2, 2.3, and 2.4.

The storage dynamics are described by the following daily mass balance:

\[
S_d = S_{d-1} + Q_d - R_d - spill_d,
\]

\[0 \leq S_d \leq S_{cap},\]

\[Q_{MEF,d} \leq R_d \leq R_{max},\]

where \( S_d \) is the reservoir storage, \( Q_d \) the reservoir inflow, \( R_d \) the volume of water released through the turbines, \( spill_d \) the volume of water spilt from the reservoir on day \( d \), \( S_{cap} \) the capacity of the dam, \( Q_{MEF,d} \) the downstream environmental flow requirement (described later), and \( R_{max} \) the maximum volume of water that can be turbined, representing the designed discharge capacity of the dam.

If the reservoir is located downstream of one or multiple reservoirs, the mass balance is modified to consider the natural inflow \( Q_d \) as well as the turbined and spill water from the upper reservoir(s) \( (R_{upper}, \text{spill}_{upper}) \):

\[
S_d = S_{d-1} + Q_d - R_d - spill_d + R_{upper} + \text{spill}_{upper},
\]

\[0 \leq S_d \leq S_{cap},\]

\[Q_{MEF,d} \leq R_d \leq R_{max}.
\]

The decision variable for each reservoir within the network is \( R_d \). This variable affects the operations of the reservoir and also influences the power system operations, since it determines the amount of hydropower available to the grid on a daily basis. To calculate initial values of \( R_d \), we use rule curves, a common approach adopted by reservoir operators [8,26]. Simply put, a rule curve represents the target storage (or level) that a reservoir should maintain on each calendar day. Following Dang et al. [24], we define the rule curves as piece-wise linear functions based on four parameters (Fig. 2): the initial values of \( R_d \), \( Q_{MEF,d} \), \( M_{MEF,d} \), and \( R_{max} \). The minimum and maximum water levels that a reservoir should reach within a year are \( (H_1 \) and \( H_2 \) and the time at which the two levels should be reached are \( T_1 \) and \( T_2 \)). For a given water system, these parameters can be reported by local agencies, retrieved from water storage data, or even inferred from satellite data [27]. The daily release decisions \( R_d \) from each reservoir are made to bring the actual storage as close to the target storage as possible, while being subjected to an upper

![Flowchart of the modelling framework](image)

Fig. 1. Modelling framework, comprising a reservoir model, power system model, and re-operation model. The dashed lines represent the feedback from the re-operation model to the two primary models. Symbols are introduced in Sections 2.2, 2.3, and 2.4.

The storage dynamics are described by the following daily mass balance:

\[
S_d = S_{d-1} + Q_d - R_d - spill_d,
\]

\[0 \leq S_d \leq S_{cap},\]

\[Q_{MEF,d} \leq R_d \leq R_{max},\]

where \( S_d \) is the reservoir storage, \( Q_d \) the reservoir inflow, \( R_d \) the volume of water released through the turbines, \( spill_d \) the volume of water spilt from the reservoir on day \( d \), \( S_{cap} \) the capacity of the dam, \( Q_{MEF,d} \) the downstream environmental flow requirement (described later), and \( R_{max} \) the maximum volume of water that can be turbined, representing the designed discharge capacity of the dam.

If the reservoir is located downstream of one or multiple reservoirs, the mass balance is modified to consider the natural inflow \( Q_d \) as well as the turbined and spill water from the upper reservoir(s) \( (R_{upper}, \text{spill}_{upper}) \):

\[
S_d = S_{d-1} + Q_d - R_d - spill_d + R_{upper} + \text{spill}_{upper},
\]

\[0 \leq S_d \leq S_{cap},\]

\[Q_{MEF,d} \leq R_d \leq R_{max}.
\]

The decision variable for each reservoir within the network is \( R_d \). This variable affects the operations of the reservoir and also influences the power system operations, since it determines the amount of hydropower available to the grid on a daily basis. To calculate initial values of \( R_d \), we use rule curves, a common approach adopted by reservoir operators [8,26]. Simply put, a rule curve represents the target storage (or level) that a reservoir should maintain on each calendar day. Following Dang et al. [24], we define the rule curves as piece-wise linear functions based on four parameters (Fig. 2): the initial values of \( R_d \), \( Q_{MEF,d} \), \( M_{MEF,d} \), and \( R_{max} \). The minimum and maximum water levels that a reservoir should reach within a year are \( (H_1 \) and \( H_2 \) and the time at which the two levels should be reached are \( T_1 \) and \( T_2 \)). For a given water system, these parameters can be reported by local agencies, retrieved from water storage data, or even inferred from satellite data [27]. The daily release decisions \( R_d \) from each reservoir are made to bring the actual storage as close to the target storage as possible, while being subjected to an upper

![Flowchart of the modelling framework](image)

Fig. 1. Modelling framework, comprising a reservoir model, power system model, and re-operation model. The dashed lines represent the feedback from the re-operation model to the two primary models. Symbols are introduced in Sections 2.2, 2.3, and 2.4.

The storage dynamics are described by the following daily mass balance:

\[
S_d = S_{d-1} + Q_d - R_d - spill_d,
\]

\[0 \leq S_d \leq S_{cap},\]

\[Q_{MEF,d} \leq R_d \leq R_{max},\]

where \( S_d \) is the reservoir storage, \( Q_d \) the reservoir inflow, \( R_d \) the volume of water released through the turbines, \( spill_d \) the volume of water spilt from the reservoir on day \( d \), \( S_{cap} \) the capacity of the dam, \( Q_{MEF,d} \) the downstream environmental flow requirement (described later), and \( R_{max} \) the maximum volume of water that can be turbined, representing the designed discharge capacity of the dam.

If the reservoir is located downstream of one or multiple reservoirs, the mass balance is modified to consider the natural inflow \( Q_d \) as well as the turbined and spill water from the upper reservoir(s) \( (R_{upper}, \text{spill}_{upper}) \):

\[
S_d = S_{d-1} + Q_d - R_d - spill_d + R_{upper} + \text{spill}_{upper},
\]

\[0 \leq S_d \leq S_{cap},\]

\[Q_{MEF,d} \leq R_d \leq R_{max}.
\]
and lower bound ($R_{\text{max}}$ and $Q_{\text{MEF},d}$, respectively), as shown in Eq. (1) (see Section 1 in the SI for additional details about the rule curves). Following the method used in [28], the daily $Q_{\text{MEF},d}$ are determined according to Eq. (3), where the daily reservoir inflows $Q_d$ are classified as low, medium, and high compared to $Q_{\text{MEF},d}$, the mean annual flow for each calendar day:

$$
Q_{\text{MEF},d} = \begin{cases} 
0.6 \cdot Q_d & \text{if } Q_d \leq 0.4 \cdot Q_{\text{MAF},d} \text{ (low)} \\
0.3 \cdot Q_d & \text{if } Q_d > 0.8 \cdot Q_{\text{MAF},d} \text{ (high)} \\
0.45 \cdot Q_d & \text{otherwise (medium)}
\end{cases}
$$

Finally, the daily available hydropower is calculated as follows:

$$
H_{\text{P}_d} = \eta \cdot \rho \cdot g \cdot R_d \cdot (H_{d-1} + H_d)/2,
$$

where $H_{\text{P}_d}$ is the available hydropower (MW) on day $d$, $\eta$ the turbine efficiency, $\rho$ the water density (1000 kg/m$^3$), $g$ the gravitational acceleration (9.81 m/s$^2$), and $H_d$ the hydraulic head, taken as the average between days $d-1$ and $d$. To ensure that the storage dynamics do not deviate excessively from the rule curves, a verification of the reservoir operations model can be found in Figure S2 in the SI.

In softly-coupled models, the one-way flow of information involves the amount of available hydropower ($H_{\text{P}_d}$) being communicated from the reservoir model to the power system model (refer to Fig. 1). This is usually done at the beginning of each day to set the upper bound for hydropower allocation in the power system model. Yet, there is no guarantee that the full hydropower allocation is needed by the grid; it could instead contribute to temporary oversupply, leading to $H_{\text{P}_d}$ not being fully dispatched. The goal of our hard-coupled framework is to tighten the coordination of water-power systems, thereby reducing the curtailment of renewable energy caused by excess of supply. On days when the full hydropower allocation is not needed by the grid, there may be benefits in re-operating the reservoir network (releasing less water and therefore producing less hydropower). The actual operations of the network will thus only be decided after gathering a feedback from the power system model.

### 2.3. Power system model

Multiple models have been developed to simulate the dynamics of large-scale power systems [14,21,29]. Here, we use PowNet [12], a highly customizable production-cost model that has been applied to Cambodia [12] and the Greater Mekong [8] for studying the vulnerability of the grid to hydro-climatic variability. PowNet solves a mixed integer linear program with the objective of meeting dynamic electricity demand at all substations at minimum cost. The decisions made by PowNet include (1) which generating units to start-up and shut down (unit commitment) and (2) the amount of power supplied by each unit (economic dispatch). The objective function of the model thus comprises operating costs, start-up costs, fuel prices and variable costs for the domestic thermal units, as well as import costs for the import nodes. The operations are also subjected to multiple constraints, such as ramping limits of the thermal plants, energy balance at each node, and capacities of the transmission lines. Importantly, modelling the power flow through the grid can provide insights into potential transmission congestion problems—such as those that may arise when incorporating increasing intermittent supply from renewables [17,18]. In contrast to the reservoir model, PowNet works at the hourly time step and with a planning horizon of 24 h. For each simulated day, PowNet outputs include hourly time series of operating costs, CO$_2$ emissions, and generation mix.

The model inputs include the parameters of the transmission lines (e.g., susceptance, capacity), the hourly time series of electricity demand at each sub-station, the parameters of the dispatchable units (e.g., capacity, O&M costs), and the hourly time series of power availability for renewable resources. The daily generation mix of systems with a large percentage of renewables can fluctuate significantly due to the influence of seasonality and anomalies in hydro-meteorological processes. Unlike thermal plants—where operations are constrained by the plant specifications—the dispatch of electricity from renewable sources is limited by the availability of the resource. For non-dispatchables like solar and wind, the input to PowNet are the hourly time series of availability over 24 h, while for hydropower the input is the available hydropower ($H_{\text{P}_d}$) calculated by the reservoir system model (Section 2.2).

#### 2.4. Re-operation model

The reservoir and power system models have independent operating guidelines, so an additional model component is required to coordinate the two-way transfer of information between the two models [30]. The re-operation model (Fig. 1) orchestrates this exchange of information in our framework, including translating variables (i.e., available and dispatched hydropower, and reservoir release) between the two models by (dis)aggregating them to a finer or coarser resolution as necessary. The final goal of the re-operation model is to implement a 'globally optimal' reservoir operation strategy that accounts for the overall state of the water-energy system, using information from both reservoir and power models on a day-to-day basis. Importantly, the re-operation modelling framework can be applied to hydropower reservoirs in both individual and cascade configurations. In our explanation, we will first describe how the model works for single reservoirs, and then highlight the modifications required for cascade systems.

The re-operation model works in two steps:

**Step 1: Compare and re-operate.** On each day, the model compares the quantity of available hydropower $H_{\text{P}_d}$ simulated by the reservoir model with the quantity of hydropower dispatched by the power system model over 24 h for each reservoir $i$, $H_{\text{P}_d}$, where $t = 1, 2, \ldots, 24$ (from PowNet). This yields one of three outcomes, as detailed below. Depending on the outcome, the model decides whether or not to re-operate the reservoir and power system. At this point, it is important to note that the inflow to a reservoir with at least one other dam upstream will be altered if its upper reservoir has been re-operated to release less water. When evaluating such reservoirs, we have to update the mass balance (using Eq. (2)) based on the new inflow and derive the new hydropower availability $H_{\text{P}_d}$ (using Eq. (4)) before comparing it against the output of the power system model.
Outcome 1: $\sum H P_i^{out} = H P_j^d$. All daily, available hydropower in reservoir $i$ is fully dispatched. This is the simplest outcome. No re-operation is required in this case, and the final strategy for reservoir $i$ is to release $R_d^i$ according to the rule curve (Fig. 2).

Outcome 2: $\sum H P_i^{out} < H P_j^d$. The available hydropower in reservoir $i$ is not fully dispatched by the power system model. This triggers the re-operation of reservoir $i$, with the goal of generating only $\sum H P_i^{out}$. In turn, this translates into a smaller release and more water storage available for future use. The re-operation steps are summarized in Algorithm 1. The main equations involved are the mass balance equation (Eq. (1)) and hydropower equation (Eq. (4)), so as to ensure consistency between the reservoir and power system operations. In Algorithm 1, we begin by substituting $HP_i$ in Eq. (4) with $\sum H P_i^{out}$, as this is the resulting hydropower that we would like to generate. Since all parameters from the previous time step $(d-1)$ are known, the only unknowns in this equation are $R_d$ and $H_d$. To solve the equation, we first assume that the water level remains unchanged from the previous day, i.e. $H_d = H_{d-1}$. The next steps involve iteratively solving Eqs. (1)–(4) until the difference in the values of the release $r$ obtained in two consecutive iterations are within the defined convergence limit $\epsilon$, where $\epsilon$ is a small arbitrary constant. In sum, we proceed as follows:

(a) Calculate $r$ from Eq. (4).
(b) Compare the obtained $r$ against $Q_{MEF,F}$, $d$ (derived from Eq. (3)) and adjust it accordingly if it does not fulfill the minimum flow requirements.
(c) Calculate spill using Eq. (1).
(d) With the estimated $r$ and spill, find the reservoir storage $S_i$ using Eq. (1). If the reservoir has inflow from other reservoirs, Eq. (2) is used instead.
(e) Update the hydraulic head $H_d$ with the derived reservoir storage $S_i$.

**Step 2: Implement best release strategy.** The best release strategy based on the current system states is conveyed back to the reservoir model for implementation. This concludes the system operations for day $d$ and the simulation will advance to the next day.

At the end of the simulation, we expect to achieve a better performance of the coupled system through more operationally flexible and informed use of the reservoirs—leading, for example, to lower system costs or better penetration of renewables. However, one natural consequence of re-operating the reservoirs is some deviation from the target storage levels recommended by the rule curves. As we shall see, strategic re-operation of reservoirs often involves (temporarily) storing more and releasing less water. As a result, more spills may be observed during wet seasons if the demand for hydropower is less than what is available. This is equivalent in some respects to curtailed wind and solar dollars during periods of over-supply. To address such a matter, one could modify the re-operation model to explicitly account for other operating factors, such as hydroupeaking, by tracking for example the day-to-day release decisions. However, these modifications are left as future explorations, because our chief objective is to identify the largest operational space (and corresponding benefits) associated with a tighter integration of water and energy systems.

### 3. Design of experiments

#### 3.1. Case study

The modelling framework is tested on the Cambodian water-energy system, illustrated in Fig. 3. Specifically, our model implementation is based on the infrastructure built and operated in 2016, for which detailed data are available [12]. Domestically, the electricity demand is fulfilled by three coal-fired units (totaling 400 MW of installed capacity), 15 oil-fired units (totaling 282 MW), and six hydropower plants (totaling 1048 MW). Near the borders, Cambodia can import hydropower from Laos and thermal power from Vietnam (200 MW) and Thailand (120 MW). The domestic hydropower availability in Cambodia is largely influenced by the seasonality of the hydro-climate. Driven by the summer Monsoon, reservoirs tend to receive more inflow during wet seasons if the demand for hydropower is less than what is available. This makes the operation infeasible, and a feedback will be sent to PowNet to re-allocate its resources for day $d$ based on an updated hydropower availability of reservoir $i$ ($HP_i^d$). Once we obtain a new generation mix from PowNet, we re-evaluate the performance of all reservoirs again, i.e., repeat Step 1 for day $d$ again.

**Algorithm 1 Reservoir re-operation algorithm (for Outcome 2)**

**Input:**

$H P_i^d$, $i = 1,...,24$

$S_{i,d-1}$, $Q_d$

**Output:** $R_d^i$, $S_d^i$, spill$_{d}^i$

**Initialize:** $H_d = H_{d-1}$

**repeat**

$r_k = \sum_{i=1}^{24} H P_i^d/(g \times \rho \times g \times (H_{d-1} + H_j^d))/2$ \quad $\triangleright$ Eq. (4)

$r_k = \max(r_k, Q_{MEF,F})$ \quad $\triangleright$ Eq. (4)

$\text{spill}_i = S_{i,d-1} + Q_d - r_k - S_{i,sp}$ \quad $\triangleright$ Eq. (1)

$S_i = S_{i,d-1} + Q_d - r_k - \text{spill}_i$ \quad $\triangleright$ Eq. (1) or Eq. (2)

$H_d \leftarrow S_i$ \quad $\triangleright$ Update hydraulic head

$k = k + 1$

until $r_k - r_{k-1} \leq \epsilon$ or $k + 1 \geq n$ \quad $\triangleright$ Convergence

**return** $R_d^i = r_k$, $S_d^i = S_i$, $\text{spill}_{d}^i = \text{spill}_i$

When $r$ converges—or we have reached the pre-defined maximum number of iterations $n$—we complete the re-operation for reservoir $i$ on day $d$ and update the new state of the reservoir accordingly (release, storage, and spill on day $d$). The best strategy is thus for reservoir $i$ to release $R_d^i = r_k$.

Outcome 3: $\sum H P_i^{out} > H P_j^d$. The dispatched hydropower in reservoir $i$ exceeds the availability. Typical operations will not result in this outcome as the hydropower allocation in PowNet is bounded by $HP_i^d$. This outcome is thus unique to downstream reservoirs with inflow from other reservoir(s) when both of the following conditions occur: (1) at least one of the upstream reservoirs is re-operated (and thus releases less water), and (2) this reduction in inflow is significant enough to generate less hydropower than the power system model allocated from reservoir $i$. This makes the operation infeasible, and a feedback will be sent to PowNet to re-allocate its resources for day $d$ based on an updated hydropower availability of reservoir $i$ ($HP_i^d$). Once we obtain a new generation mix from PowNet, we re-evaluate the performance of all reservoirs again, i.e., repeat Step 1 for day $d$ again.
demand profiles for weekdays and weekends. To get the hourly load for each substation, we relied on a spatial and temporal disaggregation of the national data. In particular, we first interpolated between the monthly peak demand data to obtain the daily peak national demand, and then disaggregated it further using weekday-weekend and peak-off-peak demand profiles to account for the variation among days in a week and hours in a day. As for the spatial disaggregation, we dispersed the national demand to the substations based on their voltage levels to yield the hourly load at each substation. We note that the electricity demand time series so derived do not show day-to-day changes in demand due to weather; however, they implicitly capture the seasonal sensitivity to air temperatures—a driver of air conditioning need—which slightly increases in pre-monsoon months. At the substation with the highest demand, we added a hypothetical “slack” generator with arbitrarily assigned parameters and very high production cost to prevent electricity shortage [12]. Further details about the model setup and validation are reported in [12].

3.2. Stochastic inputs

To characterize the benefits of co-simulating the water and energy systems, we conduct a probabilistic assessment of the system performance by generating stochastic replicates of the input variables. Since the water-energy system is susceptible to both hydrological and meteorological uncertainty, the inputs required in our set-up include streamflow (for the reservoir model), and electricity demand and solar power production (for the power system model). In particular, each variable is synthetically expanded to produce 1000 single-year time series. A random combination of the expanded time series will allow us to have a better understanding of how the system performance evolves with different inputs as they form potential scenarios that drive the system’s responses. In the stochastic generation process, we assume that the three processes are independent—an assumption that was verified through the observational historical datasets.

3.2.1. Streamflow

It is especially important to explore the variability in streamflow as Cambodia is subjected to pronounced hydro-climatic variability due to the summer monsoon (Section 3.1). The stochastic model that we require should therefore be able to capture the monsoonal effects in the generated streamflow while ensuring that the spatial correlation of the inflow across the six reservoirs is well represented. The Kirsch–Nowak Streamflow Generator fulfils the required features, with a stochastic model that produces correlated synthetic daily streamflow time series at multiple sites [34]. Besides the generator engine, the model also consists of a monsoonal rescaling function to capture inter-annual variability in the monsoon [35]. This stochastic model has been used to generate synthetic streamflow in other regions of Southeast Asia, such as Vietnam [36,37].

Daily inflow data for the six reservoirs in our study were obtained for 37 years (1980–2016) from the Global Flood Awareness System (GloFAS) dataset [38]. The streamflow data of each reservoir were then synthetically expanded to yield the daily time series for 1000 years. The top panels of Fig. 4 compare the historical observed records to the generated synthetic time series at multiple sites [34]. The streamflow data shown (top left panel) are the aggregated total across the six reservoirs for each year. Note that the median of both records is similar in pattern and scale, while the envelope of the synthetic data extends beyond that of the historical record. This allows us to explore plausible inflow conditions outside of recent observed data.
3.2.2. Solar power

As we shall see in Section 3.3, our experimental setup includes a few scenarios where solar plants are integrated in the Cambodian grid. As solar power is non-dispatchable, the required input to PowNet are the hourly time series of solar availability. Solar power production is derived from solar irradiance using the GSEE package [39], which takes as input time series of solar irradiance, the coordinates of the solar plants, and the plant specifications (e.g., installed capacity, tilt angle). The hourly solar irradiance data for Cambodia were obtained for 39 years (1981–2019) from Renewables.ninja [40].

To generate the synthetic data, we model the deviations of observed irradiance from “clear sky” (cloudless) conditions, a method also used in [21,41]. We begin by identifying the 365-day “clear sky” profile and subtracting the observed irradiance for each calendar day to find the “losses” due to cloud cover:

\[ I_d^l = I_d^o - I_d^c, \]

where \( I_d^l \) represents the irradiance “losses” on day \( d \), \( I_d^o \) the “clear sky” irradiance on calendar day \( n \), and \( I_d^c \) the observed irradiance on day \( d \). The irradiance “losses” are then standardized to approximate a Gaussian distribution:

\[ \tilde{I}_d^l = \frac{I_d^l - \mu(I^n_d)}{\sigma(I^n_d)}, \]

where \( \tilde{I}_d^l \) is the standardized irradiance loss on day \( d \), and \( \mu(I^n_d) \) and \( \sigma(I^n_d) \) are the average and standard deviation of \( I_d^l \) on calendar day \( n \), respectively. We then model the resulting \( \tilde{I}_d^l \) as an AutoRegressive Moving Average (ARMA) process with order \((n_a, n_c)\):

\[ \tilde{I}_d^l = \sum_{i=1}^{n_a} a_i I_{d-i} + \sum_{i=1}^{n_c} c_i e_{d-i} + e_d, \]

where \( a_i \) and \( c_i \) are the autoregressive and moving average coefficients, and \( e_d \sim \mathcal{N}(0, \sigma^2) \) the error generated by a white Gaussian stochastic process. ARMA(2, 2) was selected to model the standardized irradiance losses as it resulted in the lowest Akaike Information Criterion (AIC) value (details of the ARMA model results can be found in Figure S3 in the SI). We then synthetically generate 1000 years of \( \tilde{I}_d^l \) by first generating random values of \( e_d \) and then propagating them through the ARMA model. The new time series of irradiance losses are obtained by reversing Eq. (6):

\[ I_d^{new} = \tilde{I}_d^l \times \sigma(I^n_d) + \mu(I^n_d), \]

and the time series of solar irradiance are obtained by subtracting the losses from the clear sky profile,

\[ I_d^{new} = I_d^o - I_d^{new}. \]

As shown in Fig. 4 (top central panel), the synthetic solar irradiance has an envelope of variability going above and below the observed ones, representing conditions where we get varying amounts of solar.

3.2.3. Load

As mentioned in Section 2.3, the objective of PowNet is to meet the electricity demand at all substations at minimum cost. The hourly scheduling of power plants (including hydroelectric dams) is thus constrained by available supply, but also strongly driven by demand. This makes it important to explore the coupled system’s responses under different load conditions. Since the only available time series of load data from technical reports in Cambodia are the monthly peak national demand data [32], these form the basis of our synthetic data generation. In particular, we employed a two-step approach to model the peak demand. First, we identify the variable related to temperature that makes it important to explore the coupled system’s responses under different load conditions. Since the only available time series of load data from technical reports in Cambodia are the monthly peak national demand data [32], these form the basis of our synthetic data generation. In particular, we employed a two-step approach to model the peak demand. First, we identify the variable related to temperature that has an envelope of variability going above and below the observed ones, representing conditions where we get varying amounts of solar.

We obtained the monthly peak demand data from [32] for six years (2011–2016), and daily minimum and maximum temperature data from the Climate Forecast System Reanalysis (CFSR) dataset [44] for 42 years (1975–2016). After performing a correlation analysis between the normalized peak demand and multiple variables related to temperature (see Figure S4 for additional details) for the common available period (2011–2016), we found that the peak demand in Cambodia has the strongest correlation \( r = 0.65 \) with the monthly mean of the minimum air temperature averaged across Cambodia (see Figure S5 in the SI). This yields the following linear model:

\[ d_{m}^{pred} = a \times \bar{T}_{m,\text{min}} + b, \]

where \( d_{m}^{pred} \) is the predicted peak demand in month \( m \), \( \bar{T}_{m,\text{min}} \) the mean of the daily minimum temperature in month \( m \), and \( a \) and \( b \) the...
coefficients of the linear regression model. The difference between the observed and fitted peak demand is \( D_m = d^{	ext{obs}}_m - \hat{D}^\text{syn}_m \), where \( D_m \) is the time series of monthly difference, and \( d^{	ext{obs}}_m \) the observed peak demand. We then model two variables stochastically: (i) the daily minimum air temperature averaged across Cambodia, time series of which are then translated to monthly mean values for use in the peak demand model, and (ii) the stochastic difference between the predicted and actual peak demand \( \hat{D}_m \). Similar to the process used for solar irradiance, we standardize the two variables using Eqs. (11) and (12) respectively:

\[
\hat{\text{\( t \)}}_{\text{\( m \),\( d \)}} = \frac{\text{\( t \)}}_{\text{\( m \),\( d \)}} - \mu(\text{\( t \)})_{\text{\( m \),\( d \)}} \quad (11)
\]

\[
\hat{D}_m = \frac{D_m - \mu(D_m)}{\sigma(D_m)} \quad (12)
\]

where \( \hat{\text{\( t \)}}_{\text{\( m \),\( d \)}} \) and \( \hat{D}_m \) are the standardized temperature on day \( d \) and residual between the predicted and actual peak demand in month \( m \), respectively. \( \text{\( t \)}_{\text{\( m \),\( d \)}} \) denotes the daily observed minimum temperature, and \( \mu(\text{\( t \)})_{\text{\( m \),\( d \)}} \) and \( \sigma(\text{\( t \)})_{\text{\( m \),\( d \)}} \) are its mean and standard deviation on calendar day \( d \) in month \( m \), while \( \mu(D_m) \) and \( \sigma(D_m) \) are the mean and standard deviation of \( D_m \) for each calendar month \( p \). In the next step, we model \( \hat{\text{\( t \)}}_{\text{\( m \),\( d \)}} \) and \( \hat{D}_m \) as separate ARMA processes, then synthetically generate 1000 years of time series for each of them (details of the ARMA model results can be found in Figure S6 in the SI). With the regression model found in Eq. (10), we then translate these 1000-year ensembles to synthetic time series of monthly peak demand \( d^\text{syn}_m \),

\[
d^\text{syn}_m = a \times \hat{\text{\( t \)}}_{\text{\( m \),\( d \)}} + c + D^\text{syn}_m \quad (13)
\]

where \( \hat{\text{\( t \)}}_{\text{\( m \),\( d \)}} \) is the mean of the synthetic minimum temperature in month \( m \), and \( D^\text{syn}_m \) the synthetic errors between the predicted and actual peak demand in month \( m \). The peak demand data are then transformed to hourly load at each substation following the method described in Section 3.1.

Fig. 4 (top right panel) shows the variation in the daily national load. The profile shows prominent rise and falls indicating the weekdays and weekends, and also has a synthetic envelope that allows us to explore the response of the system in conditions with high and low loads.

3.3. Experimental setup

Recall that the goal of our study is to understand the benefits of hard-coupling water and power system models. Our experimental design explores the benefits of this integration under: (1) stationary hydro-meteorological uncertainty (Section 3.2); and (2) current and future grid infrastructure scenarios, including future additions of renewable energy.

As Cambodia strives to both expand and decarbonize its grid, multiple solar farms are planned to be built for fulfilling the electricity demand up until 2030 [31]. We thus explore the addition of two solar nodes, one at Kampong Chhnang and another at Kampong Speu province, both near the capital Phnom Penh (Fig. 3). We model three different infrastructure configurations: (1) the 2016 as-is grid (400 MW coal, 282 MW oil, 1048 MW hydropower, and 320 MW imports), (2) the grid with 300 MW of solar capacity installed (two plants with 140 MW and 160 MW capacity), and (3) the grid with 600 MW of solar installed (two plants with 280 MW and 320 MW capacity). The location of the solar plants is based on the aforementioned reports, while the plant capacities are loosely based on the amount expected to be installed in the coming decade (so as to model the system under a broad range of conditions). These three grid configurations are referred to throughout the rest of this paper as Case 1, Case 2, and Case 3, respectively. Each grid configuration is simulated under two modelling implementations: (1) with hard-coupling of the water-energy models; and (2) without hard-coupling. The three grid configurations paired with hard-coupling of the water-energy models are referred to as Case 1H, Case 2H, and Case 3H, respectively. To determine the volume of simulations that should be conducted, we track the convergence between the 1st and 99th percentile values of the historical and synthetic inputs. As shown in Fig. 4 (bottom panel), we require at least 900, 300, and 200 simulations for streamflow, solar irradiance, and load, respectively, to achieve a stabilized distribution. In our study, we thus conducted our experiments with 1000 simulation years. Thus, our total number of simulated years is 6000 (1000 synthetic weather years x 3 grid configurations x 2 modelling implementations). In our study, we defined the convergence limit \( \epsilon = 10^{-4} \), and maximum number of iterations \( n = 10 \). Typically, the re-operation model takes up to three iterations to calculate the updated release of the reservoir. The runtime (minutes) of each year of simulation with respect to different external drivers is reported in Figure S7. Based on simulations ran on an Intel(R) Core(TM) i7-8700 CPU 3.2 GHz with 8 GB RAM running Windows 10, the runtime varies between 20 and 65 min for the scenarios without feedback (Cases 1, 2 and 3). The addition of the feedback mechanism causes a larger range in computation time (up to two times) across the scenarios. In general, cases with a larger reservoir inflow and smaller load are more computationally demanding. Depending on the experimental set-up and simulation–optimization environment, each year of simulation on the Cambodian grid can take up to two hours.

4. Results

First, we report results from our probabilistic assessment of the benefits provided by the hard-coupled framework (Section 4.1). Then, we identify the key factors that affect such benefits, so as to understand their relevance in a real-world context (Section 4.2). Finally, we explore the effects of the hard coupling mechanism on the re-operation of individual grid components (Section 4.3).

4.1. System response to co-simulation

To quantify the potential benefits of the co-simulation framework, we consider three metrics, namely annual operating costs, \( \text{CO}_2 \) emissions, and penetration of renewables in the generation mix. Fig. 5 (top panel) illustrates the effects of hard coupling for each grid configuration aggregated across the 1000 synthetic weather years. The first key result to notice is that the adoption of the hard-coupling nearly always increases the penetration of variable renewable energy (wind and solar) and decreases both annual operating costs and \( \text{CO}_2 \) emissions, indicating that the benefits of the co-simulation framework are robust with respect to stationary uncertainty in streamflow, solar irradiance, and load. Importantly, this finding is consistent across the three grid configurations (i.e., current grid, +300 MW, and +600 MW of solar capacity).

The second result worth discussing here is the magnitude of the benefits reaped by the hard-coupling. Starting with the existing grid (Case 1 and 1H), the cost savings from hard-coupling range between 2% and 40% per year, while the \( \text{CO}_2 \) reduction ranges between 4% and 50%, reflecting the benefits of the curbed reliance of fossil fuels and corresponding increased penetration of renewables (range between 1% and 20%). As hydropower is the only renewable resource in Case 1H, the results suggest that the system’s performance can be improved just by re-operating the reservoir, even before additional renewable energy is added. Shifting our focus to Scenarios 2 and 3, the addition of solar resources results in a slight increase in the penetration of renewables caused by hard-coupling, leading to further reductions in operating costs and \( \text{CO}_2 \) emissions. It is interesting to note that although the cumulative distribution functions of change in renewable integration do not seem to vary much across the three grid setups, the disparity of their effects on cost and \( \text{CO}_2 \) emissions are more pronounced. For example, there is a 50% chance of hard-coupling improving the operating costs by at least 15% in Case 1H, a chance that
increases to 63% and 74% in Case 2H and 3H, respectively. A similar observation applies to CO$_2$ emissions.

We carry out a similar analysis for the capacity factor of the four generation types (i.e., coal, oil, hydropower, and solar), whose cumulative distribution functions are shown in Fig. 5 (bottom panel). This analysis reveals that different generation types exhibit different responses to the feedback mechanism added via model hard-coupling. For instance, the capacity factor of coal plants always decreases, while that of oil plants has a ~90% probability of increasing. Hydropower usage always increases (despite the same amount of water being available with and without hard-coupling). As we shall see in Section 4.3, the varying responses of the dispatchable power plants (coal, oil, hydropower) to the hard-coupling arise from factors such as plant locations and the reservoirs they are interconnected with. For example, two out of the three coal plants are directly connected to Kamchay, the reservoir with the biggest storage capacity (see Fig. 3 and Table 1). When the large reservoir is re-operated, the re-distribution of water resources increases the plant’s hydropower production, reducing the capacity factor of the nearby coal plants by up to 25%. Conversely, the capacity factor of oil plants only varies between ~0.5% and 1%. This can be attributed to their location, as most of them are near Phnom Penh (the main load-centre) and do not have direct interconnections with any hydropower reservoir.

4.2. Factors affecting performance of the feedback mechanism

Fig. 5 shows that the use of a hardly-coupled model always out-performs the benchmark modelling approach. It also shows that the magnitude of the benefits varies widely across years. It is therefore essential to understand how the hydro-meteorological conditions control the effectiveness of such hard integration. This is done through a parallel coordinate plot (Fig. 6), where we visualize the annual values of streamflow, load, and the system’s response to the hard coupling obtained when moving from Case 1 to Case 1H. (similar analyses for Case 2/2H and 3/3H are reported in Figure S8 in the SI). The variables are shown across five parallel axes, so each line connecting the axes represents a single year of the 1000-year stochastic ensemble. Note that this analysis focuses on the ten years that result in the largest (99th percentile and above) and smallest (1st percentile and below) difference between the benchmark and hard-coupled formulations for operating costs and CO$_2$ emissions. Second, both total annual streamflow and load appear to control the extent of the benefits. Beginning with streamflow, it is intuitive that larger benefits will be reaped if the system has additional flexibility in storing water for more timely hydropower generation. If dams are better able to coordinate their release decisions with the state of the power grid, it follows that the grid should be able to rely less on fossil fuel power plants. This is evident in the “Annual available Q” axis, where the maximum benefit case corresponds to large values of Q in the top panel—while the highlighted lines are clustered near the bottom of the axis in the bottom panel. In other words, having a larger volume of water in the system allows greater flexibility in the re-operation (as shown in the third axis), thus hydropower generation can be better scheduled or deferred to allow for the use of solar (see Figure S8)—an opportunty that does not exist during dry years. Likewise, the frequency of reservoir re-operation also depends on load. We observe that years with high annual streamflow and low load (leading to more frequent periods of low net demand) are associated with the largest benefits from model hard-coupling, allowing annual cost savings of up to 44% and CO$_2$ reduction of up to 53%. Conversely, in years with low Q and high load, the benefits of greater flexibility in re-allocating water are diminished, resulting in an annual cost and CO$_2$ reduction of about 4%. Third, a prominent difference between the top and bottom panels stands in the dispersion of the values associated to the highlighted years—in the bottom panel the ten highlighted lines tend to converge around similar values. A possible explanation for this behaviour exists in the distribution of the inflow process: high-streamflow years are less concentrated around a
Fig. 6. Parallel coordinate plot illustrating the relation between input variables (streamflow ($Q$) and load) and the system’s response to the hard coupling (annual hydropower dispatched, costs, and CO$_2$ emissions) obtained when moving from Case 1 to Case 1H. Each line corresponds to one of the 1000 simulation scenarios. In the top panel we highlight the years with the largest benefits (above the 99th percentile) in terms of the operating costs and CO$_2$ emissions, while in the bottom panel we highlight the years with the smallest benefit (below the 1st percentile). The bold lines represent the year with the maximum and minimum benefits, respectively, in the top and bottom panels. The analysis for Case 2/2H and 3/3H is reported in Figure S8 in the SI.

Fig. 7. Probabilistic assessment of the annual operating costs as a function of annual values of streamflow ($Q$, horizontal axis) and load (vertical axis). Their bivariate distribution (estimated with a Gaussian kernel) is represented by the isolines inside the plot. Each symbol represents a year that led to the largest (99th percentile and above) and smallest (1st percentile and below) operating cost reductions. Squares, triangles, and crosses correspond to the three grid configurations.

Having identified annual streamflow and load as the key factors that influence the benefits of hard-coupling in a given year, we then evaluate this finding across the different grid set-ups. Fig. 7 illustrates the joint distribution of total annual streamflow and load (estimated with a Gaussian kernel), and highlights the 1st and 99th percentile years in terms of annual operating cost for all grid set-ups. A similar plot highlighting the performance in terms of CO$_2$ emissions can be found in the SI (Figure S9). A visual comparison between the two panels shows that the points tend to cluster in specific regions of the streamflow-load space—the ones with the least benefits mostly lie on
Fig. 8. Variability in input variables (streamflow ($Q$) and load), number of reservoir re-operations, and the resulting system’s performance (increase in the penetration of renewables and decrease in operating costs) across different months.

with the top left quadrant (small inflow, large load), while the ones with the largest benefits lie on the bottom right (large inflow, small load). This reinforces the finding that streamflow and load are the main factors that determine the effectiveness of the feedback mechanism in a given weather year.

4.3. Spatial and temporal effects of the feedback mechanism

With an understanding of how the external conditions influence the overall system’s performance, we now explore how the re-operation mechanism modifies the grid behaviour at a finer spatio-temporal scale. This is achieved through an assessment of the month-to-month variability in the re-operation frequency for each hydropower reservoir, illustrated in Fig. 8 (central panels). By looking at the temporal patterns revealed by the violin plots, it is evident that the re-operation mechanism is heavily influenced by the seasonality of the monsoon (see the monthly total streamflow illustrated in Fig. 8, upper panel). For all six reservoirs, the number of days in which the re-operation is adopted is fairly modest in the pre-monsoon season (January to April), gradually increasing during the summer monsoon (May to October) and then decreasing in November–December. This pattern is a direct consequence of the control that water availability exerts on the benefits of the re-operation mechanism (Section 4.2). The violin plots also suggest that the monthly number of re-operations is rather stable during the central part of the monsoon season (i.e., July to September), while it tends to vary at the beginning and end—a result explained by the inter-annual variability in the monsoon season onset, withdrawal, and length [45].

The temporal patterns of the re-operation mechanism are reflected in the monthly operating costs and penetration of renewables (Fig. 8, bottom panels), which improve as the monsoon season unfolds. The best values for both metrics are observed in October, when the streamflow starts to decrease, a behaviour attributable to the use of water accumulated during the summer monsoon. As for the three grid configurations (Case 1, 2, and 3), the temporal patterns of reservoir re-operation are similar, although a closer inspection reveals that the reservoirs get increasingly re-operated with a larger capacity of solar in the system. With the addition of solar, the system is given the option to use an alternative non-dispatchable resource, causing a more frequent mismatch between the available and dispatched hydropower. Naturally, the deployment of more solar (Case 2/2H and 3/3H) further impacts the monthly operating costs and penetration of renewables.

Among the six reservoirs in the grid, Kamchay is re-operated the least. Spatially, the reservoirs can be divided into three main clusters, based on the interconnections of the transmission lines (refer to Fig. 3): (i) Atay, LR Chrum and Tatay, (ii) Kirirom I and Kirirom III, and (iii) Kamchay. The re-operation patterns exhibited by the reservoirs in groups (i) and (ii) are similar and well-reflected in the transmission line usage. Indeed, the increase in re-operation during the monsoon keeps the line usage within a narrow operational band (refer to Figure S10 in the SI). Since both clusters consist of multiple reservoirs grouped together, the power supply within each respective sub-grid comes primarily from hydropower, resulting in a very similar pattern between
re-operation and line usage. In contrast, Kamchay is a generating unit in a sub-grid that consists of other generating units, demand nodes, and transmission lines, all feeding Phnom Penh. The difference in response can be attributed to the complex interactions between the different components in the sub-grid. As seen in Section 4.1, different generators respond differently with the addition of the feedback mechanism. The hard coupling also influence the frequency of the N-1 violations (instances in which any of the high-voltage lines reaches 75% of its capacity) for the most congested lines (see Figure S10 in the SI). Taking the spatial and temporal variability together, the above analyses highlight that the topology of the grid, existing resources, and seasonal dynamics play a part in how the system responds to the hard-coupling. Given the role that hydropower reservoirs play in both systems, their operations have to be adapted to local conditions, an information that is not available in soft-coupling approaches.

5. Conclusions

Overall, our study contributes a novel modelling framework that co-simulates water and energy systems. The hard coupling is facilitated by a dam re-operation model able to handle both individual and multi-reservoir systems. By testing the coupled models under uncertainty in external conditions, we show robust improvements—w.r.t. a soft-coupling approach—in terms of system operating costs, CO₂ emissions, and penetration of renewables in the generation mix. This implies that through modelling the water and energy systems as a whole—instead of taking them as individual systems—more efficient operating patterns can be achieved. In particular, we show that the hard coupling works best in conditions of high water availability and low electricity demand, a condition that allows us to shift water volumes and hydropower production so as to ‘optimally’ track the intermittent production of other renewables. Because of this added flexibility, the benefits of co-simulation are expanded in cases with more solar deployed in the grid, where we see more renewable penetration in the generation mix (and more frequent periods of locational generation “oversupply”). Thus, our results suggest that co-simulation frameworks may be a useful tool to support not only operational decision-making, but also grid decarbonization efforts that are inevitably associated with major investment costs (e.g., Siala et al. [5] and Deshmukh et al. [46]). In this regard, it is worth stressing here that our results likely represent the upper bound of the benefits attainable through hard-coupling frameworks, because there might be unintended consequences of reservoir re-operations. For example, storing water during the monsoon season to track an increase of solar production may raise the frequency of spill events (see Figure S11 in the SI). Another potential consequence is hydropoeaking, as a sudden surge in electricity demand may prompt operators to respond by releasing water from hydropower reservoirs, putting pressure on riverine ecosystems [47]. To address these potential challenges, knowledge about the behaviour of individual grid components—like the one revealed by our analysis—is therefore useful. For example, one could choose which dams to re-operate, where to place renewables, or whether to narrow the re-operation space.

Looking forward, we believe modelling frameworks like this one could potentially support many other research efforts in the water-energy domain. For example, seasonal streamflow forecasts are increasingly being used by hydropower operators to plan production schedules [48], so it is important to understand and quantify the relationship between forecast accuracy and value [49]—measured in monetary terms or added hydropower production. Such efforts may benefit co-simulation environments, as they would help researchers account for all aspects (e.g., presence of other renewables, grid congestion problems) that are likely to affect the forecast value. Another example application in the hydropower industry domain are pumped hydro energy storage systems [50,51], whose operation is intrinsically connected to the state of the grid. Finally, research efforts should address the increase in computational requirements that is inevitably associated to the hard-coupling of water and energy models. If we find a solution to this issue, we then have a pathway to study the integration of renewables over larger spatial domains.

CRediT authorship contribution statement

Rachel Koh: Software, Methodology, Formal analysis, Investigation, Writing — original draft. Jordan Kern: Conceptualization, Methodology, Supervision, Writing — review & editing. Stefano Galelli: Conceptualization, Methodology, Visualization, Supervision, Writing — review & editing.

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.

Acknowledgement

Rachel Koh is supported by the President’s Graduate Fellowship from the Singapore University of Technology and Design.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.apenergy.2022.119386.

References


[34] Cook BI, Buckley BM. Objective determination of monsoon season onset, withdrawal, and length. J Geophys Res: Atmos 2009;114(D23).


