Graphical Abstract

A physics-informed reinforcement learning framework for robust irrigation control under climatic uncertainty

Raymond Houé Ngouna, Philippe Berton, Fabien Dauriac

Highlights

A physics-informed reinforcement learning framework for robust irrigation control under climatic uncertainty

Raymond Houé Ngouna, Philippe Berton, Fabien Dauriac

- A physics-informed reinforcement learning framework is proposed for irrigation control
- Rule-based, reinforcement learning, and hybrid neuro-physical controllers are compared
- A soil–water balance model is augmented with learned residual dynamics
- Hybrid control mitigates extreme stress while preserving water-use efficiency across soil and climatic parameter settings
- The framework supports adaptive decision-making under climatic uncertainty

A physics-informed reinforcement learning framework for robust irrigation control under climatic uncertainty

Raymond Houé Ngouna[a], Philippe Berton[b], Fabien Dauriac[b]

[a] *Université de Technologie de Tarbes Occitanie Pyrénées, Laboratoire Génie de Production, 47 Avenue d'Azereix, Tarbes, 65000, France*
[b] *Rives et Eaux du Sud-Ouest, BP 449 Chemin de Lalette, Tarbes, 65000, France*

Abstract

Climate change is increasing hydro-climatic variability, challenging conventional irrigation strategies that rely on static rules and fixed decision thresholds. From a modelling perspective, irrigation management can be formulated as a dynamic environmental control problem operating under uncertainty while constrained by soil–water physics. This study proposes a physics-informed reinforcement learning framework for irrigation control that combines mechanistic soil–water modelling with adaptive decision-making. A physically inspired irrigation environment is constructed using a daily soil–water balance model subject to stochastic rainfall and evapotranspiration forcing. Within this environment, three control strategies of increasing modelling and control complexity are systematically compared: (i) a deterministic rule-based controller derived from expert-defined soil water tension thresholds; (ii) a reinforcement learning controller trained using Proximal Policy Optimization (PPO) interacting directly with the physical model; and (iii) a hybrid neuro-physical formulation in which systematic model mismatch is corrected using a Neural ODE-inspired residual model integrated in discrete time. This progressive formulation enables a controlled comparison between heuristic control, learning-based control, and physics-augmented learning under identical soil and meteorological parameterizations. Simulation results indicate that reinforcement learning improves water-use efficiency relative to rule-based control but may induce increased stress variability under simplified physical modelling assumptions. Augmenting the physical model with learned residual dynamics attenuates extreme stress responses and reduces drainage losses while preserving efficiency gains, thereby reshaping the trade-off between robustness and water savings within the considered soil and climatic parameterizations. Beyond irrigation, the proposed framework illustrates a transferable decision-support paradigm for environmental systems in which simplified process-based models are combined with adaptive control under stochastic forcing. The emphasis is on robustness to model mismatch and parameter uncertainty rather than on the direct transfer of trained policies across domains.

Keywords:
Intelligent irrigation, Environmental decision support, Physics-informed reinforcement learning, Hybrid neuro-physical modelling, Soil–water dynamics, Adaptive control under uncertainty

1. Introduction

Climate change is increasing the frequency and intensity of hydro-climatic variability, placing growing pressure on agricultural water management systems. Irrigation, which accounts for a major share of global freshwater withdrawals, is particularly sensitive to rainfall uncertainty, seasonal variability, and soil–plant interactions. Conventional irrigation strategies, typically based on fixed schedules or expert-defined thresholds, often lack the flexibility required to adapt to non-stationary environmental conditions, leading to inefficient water use, excessive drainage losses, and increased crop water stress.

From a modelling perspective, irrigation management can be formulated as a dynamic environmental control problem in which decisions are made sequentially under uncertainty while respecting physical constraints. Process-based soil–water models have long been used to represent irrigation dynamics and support decision-making, offering interpretability and physical consistency through simplified representations such as bucket models or FAO-56-based approaches (Raes et al., 2009). However, these models rely on structural assumptions and parameterizations that may not fully capture context-dependent responses, soil heterogeneity, or the cumulative effects of stochastic climatic forcing.

Recent advances in reinforcement learning (RL) have enabled the development of adaptive control policies that can exploit temporal dependencies and delayed system responses in complex and stochastic environments (Sutton and Barto, 2018). RL-based approaches have been increasingly explored for irrigation and water management applications, demonstrating potential improvements over static rule-based strategies in terms of water-use efficiency. Nevertheless, purely data-driven control raises well-known concerns regarding interpretability, physical plausibility, and safety, particularly in environmental systems where decisions have delayed and potentially irreversible consequences. These limitations have motivated growing interest in physics-informed and hybrid modelling approaches that combine mechanistic understanding with learning-based adaptability.

Physics-informed learning provides a principled means of constraining data-driven models using known system structures, thereby improving stability and generalization (Willard et al., 2022). In this context, Neural Ordinary Differential Equations (Neural ODEs) have emerged as a flexible framework for learning residual dynamics that augment existing physical models while preserving interpretability (Rackauckas et al., 2020). In environmental applications, such hybrid formulations are increasingly viewed not as replacements for process-based models, but as corrective mechanisms that compensate for systematic modelling errors arising from simplifications, parameter uncertainty, or unresolved processes.

Importantly, many operational irrigation systems operate on discrete daily decision cycles and rely on sensor-level observations such as soil matric potential and rainfall measurements. Consequently, the practical integration of learning-based methods requires careful alignment between control decisions, temporal discretization, and physical modelling assumptions. In this study, learning is introduced at two distinct levels: (i) at the policy level, through reinforcement learning with continuous irrigation actions, and (ii) at the dynamics level, through a discrete-time residual neural model inspired by Neural ODE formulations. This separation enables adaptive decision-making while maintaining a physically interpretable and computationally efficient simulation framework aligned with available data and operational constraints.

Despite increasing interest in learning-based irrigation control, several open questions remain. These include the relative benefits of reinforcement learning compared to expert rule-based strategies when interacting with simplified physical models, the extent to which learned residual dynamics can mitigate structural model mismatch, and the implications of tighter model–controller integration for interpretability, transferability, and practical deployment. Addressing these questions requires controlled comparative studies that explicitly examine modelling assumptions, parameter settings, and performance trade-offs under identical physical and climatic forcing.

Within the Environmental Modelling & Software community, there is growing emphasis on transparent model design, explicit uncertainty handling, and reproducible evaluation workflows (Refsgaard et al., 2007; Jakeman et al., 2006). This study aligns with these principles by adopting a configuration-driven experimental framework, systematically varying soil and meteorological parameters, and clearly separating training and evaluation phases for learning-based controllers.

Accordingly, this study addresses the following research questions:

- **RQ1:** To what extent does reinforcement learning improve irrigation control performance compared to expert rule-based strategies when interacting with a physics-based soil–water model?
- **RQ2:** Does augmenting the physical model with learned residual dynamics improve the stability and safety of learning-based irrigation control under stochastic climatic forcing?
- **RQ3:** How does increasing model–controller integration affect interpretability, reproducibility, and transferability in environmental decision-support systems?

The main contributions of this study are as follows:

1. A physics-informed reinforcement learning formulation that models irrigation as a finite-horizon sequential decision-making problem under climatic uncertainty, explicitly distinguishing latent physical states from sensor-level observations.
2. A structured and reproducible comparison of three irrigation control paradigms—rule-based control, reinforcement learning, and hybrid neuro-physical reinforcement learning—within a unified physical simulation environment.
3. The integration of a discrete-time Neural ODE-inspired residual model to correct simplified soil–water balance dynamics while preserving physical consistency and interpretability.
4. A systematic evaluation under stochastic rainfall and evapotranspiration forcing, highlighting trade-offs between water-use efficiency, stress avoidance, and drainage losses as a function of soil and meteorological parameterization.

5. A transferable modelling and control framework designed to support future integration of real sensor data and extension to other environmental systems requiring adaptive control under uncertainty.

The remainder of this paper is organised as follows. Section 2 reviews related work on irrigation modelling, rule-based control, reinforcement learning, and hybrid physics-informed approaches. Section 3 presents the materials and methods, including the problem formulation, simulated data sources, physical environment, control scenarios, and experimental design. Section 4 reports and discusses the results across scenarios, with emphasis on parameter sensitivity and performance trade-offs. Finally, Section 5 concludes the paper and outlines perspectives for future research.

2. Related Work

This section reviews the relevant literature on irrigation modelling and control, focusing on the transition from process- and rule-based approaches to learning-based and hybrid physics-informed methods. The aim was to position the present work within established environmental modelling paradigms and identify the methodological gaps addressed by the proposed framework.

2.1. Process-based modelling of irrigation systems

Process-based models remain a cornerstone of irrigation modelling, providing mechanistic representations of soil–water–plant interactions based on mass balance principles and evapotranspiration formulations. Classical approaches range from conceptual bucket models to FAO-56-based methods, as implemented in decision-support tools such as AquaCrop (Raes et al., 2009). These models offer interpretability, transparency, and relatively low data requirements, which are essential for scenario analyses and policy-oriented studies.

Nevertheless, process-based models rely on simplified assumptions regarding soil heterogeneity, root water uptake, and boundary conditions. Their performance may degrade under highly variable or extreme climatic conditions, where unmodeled dynamics and parameter uncertainty become dominant. Recent reviews highlight that although physically based models are indispensable, their standalone use may be insufficient for adaptive irrigation management under climate change (Fatichi et al., 2016; Seneviratne et al., 2021).

2.2. Rule-based and heuristic irrigation control

Rule-based irrigation strategies, often derived from expert knowledge or agronomic guidelines, remain widely practiced. These approaches typically rely on fixed thresholds for soil moisture, crop water stress indices, or accumulated evapotranspiration deficits. Their appeal lies in their simplicity, interpretability, and ease of deployment, particularly in data-scarce environments.

However, heuristic controllers are inherently static and may struggle to adapt to nonstationary climatic conditions. Recent comparative studies emphasize that threshold-based rules can perform suboptimally when rainfall variability or delayed soil responses are significant, motivating the exploration of adaptive control strategies (Jones et al., 2022). Consequently, rule-based control is increasingly regarded as a baseline rather than a long-term solution for climate-resilient irrigation.

2.3. Reinforcement learning for irrigation and water management

Reinforcement learning (RL) provides a flexible framework for sequential decision-making in uncertain and dynamic environments (Sutton and Barto, 2018). In the context of water management, RL has been applied to reservoir operation, canal regulation, and irrigation scheduling, with several studies reporting improved performance compared to static policies (Yang et al., 2021; Giuliani et al., 2021).

Despite these advances, purely data-driven RL approaches face significant challenges in environmental applications. They often require extensive training data, may violate physical constraints, and can produce policies that are difficult to interpret or to trust. Recent reviews have stressed that the lack of physical consistency and robustness remains a major barrier to the operational adoption of RL in environmental systems (Rolnick et al., 2022; Reichstein et al., 2019).

Within the environmental modelling community, adaptive control and policy search methods have long been investigated as means to support decision-making under deep uncertainty and competing objectives. In particular, policy search approaches have been shown to provide a principled framework for exploring the trade-offs between robustness, efficiency, and risk in water management systems (Giuliani et al., 2016). Similarly, many-objective evolutionary algorithms, such as the Borg framework, have been widely adopted to identify diverse and non-dominated control strategies across complex environmental objectives (Hadka and Reed, 2013).

In this context, reinforcement learning can be viewed not as a replacement for existing environmental decision-support methodologies but as a complementary adaptive control paradigm that shares conceptual foundations with policy search and many-objective optimization while offering an enhanced capacity to exploit temporal structure and delayed system responses.

2.4. Physics-informed and hybrid learning approaches

To address these limitations, physics-informed machine learning has emerged as a promising paradigm that embeds mechanistic knowledge into data-driven modelling. Comprehensive surveys have highlighted that constraining learning with a physical structure improves generalization, stability, and interpretability, particularly in data-limited regimes (Willard et al., 2022; Karniadakis et al., 2021).

Neural Ordinary Differential Equations (Neural ODEs) provide a natural framework for hybrid modelling by enabling continuous-time representations in which neural networks parametrize unknown or residual dynamics (Rackauckas et al., 2020). In environmental and Earth system sciences, Neural ODEs and related universal differential equation frameworks have been increasingly used to augment physical models rather than replace them, preserving interpretability while improving fidelity (Rackauckas et al., 2021; Beucler et al., 2021).

2.5. Hybrid modelling for control under uncertainty

Although physics-informed learning has been extensively studied for system identification and forecasting, its integration with reinforcement learning for control remains comparatively underexplored in environmental modelling. Recent work in safe and model-based RL emphasizes the importance of incorporating physical structures to improve robustness and avoid unsafe or unrealistic control policies (Berkenkamp et al., 2017; Perkins et al., 2023).

Few studies have conducted systematic comparisons between rule-based control, reinforcement learning, and hybrid neurophysical control within a unified environmental modelling framework. In particular, the potential of learning residual dynamics via Neural ODEs to enhance irrigation control robustness under climatic variability has received limited attention. This gap motivates the present study, which investigates control strategies for increasing modelling complexity within a physics-based irrigation environment.

3. Materials and Methods

This section describes the materials and methods adopted in this study. We begin by clarifying the scope and data sources of the work.We then formulated irrigation management.Next, we present the physics-based soil–water environment used to simulate daily dynamics, followed by the definition of the three control scenarios considered: rule-based control, reinforcement learning, and hybrid neuro-physical reinforcement learning. Finally, we detail the experimental design, training procedures, and evaluation protocols used to ensure reproducibility and fair comparison across the control strategies.

3.1. Data sources and scope of the study

This study was conducted using simulated data generated by a physics-based soil–water balance model driven by stochastic climatic forcing. The use of simulations enables controlled experimentation, systematic comparison of control strategies, and reproducibility under a wide range of hydroclimatic conditions while avoiding confounding effects related to data scarcity, sensor noise, or site-specific calibration.

The simulated environment produces daily trajectories of rainfall, reference evapotranspiration, soil water storage, and soil matric potential (tension), which together define the state and observation variables used by the control algorithms. In all scenarios, irrigation decisions were evaluated against identical simulated physical dynamics and weather realizations, ensuring fair and consistent comparisons across the control strategies.

Although the present study focuses on simulated data, the proposed framework is explicitly designed to be transferable to real-world irrigation systems. In operational settings, soil water storage is typically not directly observable; instead, the soil water status is monitored using tensiometers that measure the soil matric potential ψ_t. Rainfall inputs can be obtained from on-site rain gauges or open-access meteorological datasets (e.g., national weather services or global reanalysis products). The observation structure adopted in this study, based on ψ_t and climatic variables, directly reflects these practical constraints.

3.2. Problem formulation

This subsection establishes a unified mathematical formulation of the irrigation control problem that underpins all the scenarios considered in this study. We formalize irrigation management as a finite-horizon sequential decision-making process, explicitly distinguishing between (i) the latent physical state of the soil–water system, (ii) the observable variables available to the controller, and (iii) the control actions constrained by agronomic and operational limits. This formulation provides a common reference framework for rule-based control, reinforcement learning, and hybrid neurophysical approaches, ensuring that performance differences arise from the control strategy rather than discrepancies in system representation.

3.2.1. System and time discretization

We consider a single agricultural plot evolving over a growing season of length T days, which is discretized into daily decision steps $t \in \{0, \ldots, T-1\}$. Irrigation decisions were made once per day, consistent with the temporal resolution of the meteorological forcing and the operational granularity targeted in this study.

3.2.2. Latent state, observations, and retention relationship

Let S_t (mm) denote the *latent* root-zone soil water storage at day t, which represents the amount of plant-available water in the effective root zone of the plant. In operational settings, S_t is not generally measured directly. Instead, soil water status is monitored using tensiometers that provide soil matric potential (tension) ψ_t (cbar), which is treated as the primary observable variable.

The link between latent storage and observed tension is defined by the soil water retention curve:

$$\psi_t = f_{\mathrm{ret}}(S_t), \qquad S_t = f_{\mathrm{ret}}^{-1}(\psi_t), \tag{1}$$

where $f_{\mathrm{ret}}(\cdot)$ is a soil-specific monotonic mapping determined by the hydraulic properties.

Simulator-accessible state. For controlled benchmarking, the simulation environment maintains and updates S_t internally. In Scenario 2 (RL), we optionally provide S_t to the agent as part of the observation vector to isolate the effect of the control strategy under identical dynamics and forcing. Importantly, this does not contradict the operational statement above: S_t is *simulator-accessible* rather than *sensor-accessible*. Partial-observability extensions that rely only on ψ_t and exogenous drivers are discussed for future work (Section 4.5.6).

3.2.3. Climate drivers

Daily climate forcing is represented by rainfall R_t (mm), reference evapotranspiration $ET0_t$ (mmday^{-1}), and crop coefficient Kc_t (dimensionless). We denote exogenous drivers compactly as

$$d_t := (R_t, ET0_t, Kc_t), \tag{2}$$

and treat them as stochastic disturbances sampled from a (possibly nonstationary) distribution:

$$d_t \sim \mathcal{P}_d. \tag{3}$$

3.2.4. Physics-based dynamics (mass balance)

Root zone dynamics are governed by a bucket-type soil–water mass balance:

$$S_{t+1} = \mathrm{clip}(S_t + \eta_I I_t + R_t - ET_{c,t} - D_t,\ 0,\ S_{\max}), \tag{4}$$

where I_t (mm) is the applied irrigation depth (control action), $\eta_I \in (0,1]$ is the irrigation efficiency, $S_{\max}$ is the maximum admissible storage, and $\mathrm{clip}(\cdot)$ enforces physical bounds. Drainage is computed as

$$D_t := D(S_t), \tag{5}$$

It is typically activated when the storage exceeds the field capacity. Crop evapotranspiration was computed following an FAO-inspired structure:

$$ET_{c,t} = Kc_t\, ET0_t\, f_{ET}(\psi_t), \tag{6}$$

where $f_{ET}(\psi_t) \in [0,1]$ is the stress reduction factor driven by soil tension. The next-day tension is obtained from storage using Eq. (1).

3.2.5. Sequential decision-making and objective

At each day t, the controller selects an irrigation action

$$I_t \in [0, I_{\max}], \tag{7}$$

where $I_{\max}$ (mm) is the maximum allowable daily irrigation depth.

The controller receives an observation vector $\mathbf{o}_t$. In this study, we used a compact, physically meaningful representation:

$$\mathbf{o}_t = \begin{cases} (\psi_t,\ R_t,\ ET0_t) & \text{(sensor-level baseline)}, \\ (\psi_t,\ S_t,\ R_t,\ ET0_t) & \text{(simulator-accessible benchmarking, used in Scenario 2)}. \end{cases} \tag{8}$$

Thus, the problem is partially observed in operational settings. Nonetheless, we adopt an MDP-over-observations formulation for tractability and reproducible comparisons, and explicitly discuss partial-observability extensions in Section 4.

The control objective is to limit crop water stress while reducing irrigation water usage and hydrological losses. We define a daily reward as

$$r_t = -\Big(\alpha\,\mathcal{L}_{\text{stress}}(\psi_t) + \beta\,I_t + \gamma\,D_t\Big), \tag{9}$$

where $\mathcal{L}_{\text{stress}}(\psi_t)$ penalizes excessive tension, and $\alpha, \beta, \gamma > 0$ weight the tradeoffs between stress avoidance, irrigation cost, and drainage loss, respectively. A policy π is evaluated by the expected discounted return:

$$J(\pi) = \mathbb{E}_\pi\left[\sum_{t=0}^{T-1} \gamma^t r_t\right], \tag{10}$$

with a discount factor $\gamma \in (0, 1]$ and expectation taken over stochastic climate forcing. This formulation is standard in reinforcement learning (Sutton and Barto, 2018) and aligns with the environmental decision-making under uncertainty.

3.3. Physics-based irrigation environment

All scenarios interact with the same physics-based environment, implementing Eqs. (4)–(6). Each episode corresponded to a full growing season. The environment returns daily observations (Eq. (8)), accepts a bounded irrigation action I_t, updates the latent storage S_t, computes ψ_t using the retention curve, and provides the reward defined in Eq. (9). Meteorological forcing $(R_t, ET0_t, Kc_t)$ was generated from a stochastic weather process with explicit seeds to ensure reproducibility.

3.4. Control scenarios

We considered three irrigation control scenarios that represented increasing levels of learning and model integration. All scenarios shared the same physics, forcing protocol, action bounds, and evaluation metrics; only the controller design differed.

3.4.1. Scenario 1: Rule-based irrigation control (physics + heuristics)

Scenario 1 couples the physics-based environment with a deterministic rule-based policy representative of common operational practices. At each day t, the rule maps the current observed tension and a simple one-day-ahead rainfall forecast to an irrigation decision:

$$I_t = \pi_{\text{rule}}(\psi_t, \widehat{R}_{t+1}), \tag{11}$$

where $\widehat{R}_{t+1}$ is a naive forecast (available in simulation; in practice, replaced by weather forecasts). We considered parameterized rule families (single-threshold, comfort-band, proportional), each defined by fixed thresholds and dose parameters. The season is simulated by iterating the physics-based update (Eq. (4)) from an initial condition at field capacity $S_0 = S_{\text{fc}}$, with $\psi_0 = f_{\text{ret}}(S_0)$.

This scenario provides an interpretable, low-cost baseline that is robust by construction but non-adaptive; parameters remain fixed across seasons and cannot optimize trade-offs under variability.

3.4.2. Scenario 2: Reinforcement learning with a physics-based environment (physics + PPO)

Scenario 2 replaces fixed heuristics with a policy learned through reinforcement learning via direct interaction with the physics-based environment. The agent observes $\mathbf{o}_t$ (Eq. (8)), selects a continuous irrigation depth $I_t \in [0, I_{\max}]$, and receives a reward r_t (Eq. (9)) and experience transitions induced by the process model (Eq. (4)).

Learning algorithm. We used Proximal Policy Optimization (PPO), an on-policy policy-gradient method with clipped updates and generalized advantage estimation. Both the policy and value functions were parameterized as multilayer perceptrons. Training proceeds over many simulated seasons under stochastic forcing, with explicit random seeds controlling both climate realization and learning initiation. Scenario 2 serves as a learning-based baseline that isolates the effect of RL when the environment dynamics are purely physics-based (no learned correction).

3.4.3. Scenario 3: Hybrid neuro-physical control (physics + residual correction + PPO)

Scenario 3 augments the physics-based environment with a learned residual correction that compensates for systematic model mismatches while preserving the mass-balance structure. The hybrid update was implemented in the *tension space*, consistent with tensiometer-driven monitoring.

Hybrid transition with residual correction. First, the physical model computes the nominal next-day storage and tension as follows:

$$S_{t+1}^{\text{phys}} = \text{clip}(S_t + \eta_I I_t + R_t - ET_{c,t} - D_t,\ 0,\ S_{\max}), \tag{12}$$

$$\psi_{t+1}^{\text{phys}} = f_{\text{ret}}(S_{t+1}^{\text{phys}}). \tag{13}$$

The residual model then predicts an additive correction as follows:

$$\Delta\psi_t = f_\theta(\psi_t, I_t, R_t, ET0_t), \tag{14}$$

and the hybrid prediction is

$$\psi_{t+1} = \psi_{t+1}^{\text{phys}} + \Delta\psi_t, \qquad S_{t+1} = f_{\text{ret}}^{-1}(\psi_{t+1}). \tag{15}$$

This ensures consistency between the corrected tension and physically admissible storage via the inverse retention curve.

Residual model architecture. In our implementation, f_θ is a lightweight MLP with two hidden layers of 64 units and tanh activations, mapping a 4D input $[\psi_t, I_t, R_t, ET0_t]^\top$ to a scalar output $\Delta\psi_t$.

Residual training (pretraining). The residual model was pretrained using supervised regression on simulated trajectories. For each sample, inputs are

$$\mathbf{x}_t = [\psi_t,\ I_t,\ R_t,\ ET0_t]^\top, \tag{16}$$

and targets are defined as the discrepancy between a perturbed "reference" next-day tension and the nominal physical prediction:

$$y_t = \psi_{t+1}^{\text{ref}} - \psi_{t+1}^{\text{phys}}. \tag{17}$$

In the current implementation, ψ_{t+1}^{ref} is generated by injecting stochastic perturbations into the physical update (emulating unresolved processes and observation noises). The parameters were optimized using Adam with a robust regression loss (Smooth L1). After pretraining, f_θ is fixed and used in the inference mode during the RL training.

Discrete-time integration choice. Although we refer to this module as a "Neural ODE" for consistency with project terminology, the implemented residual correction is *discrete-time*: the network predicts the one-day correction $\Delta\psi_t$ directly (Eq. (14)).This choice (i) aligns with the daily forcing and decision frequency, (ii) reduces the computational overhead to a single forward pass per day, and (iii) avoids solver-induced numerical issues. Continuous-time residual formulations and higher-frequency data assimilation are considered extensions of this approach.

Reinforcement learning on the hybrid environment. A PPO agent is then trained on the hybrid environment using the same reward structure as Scenario 2 (Eq. (9)). This isolates the benefit of correcting the dynamics-level mismatch while keeping the policy-learning mechanism unchanged.

3.5. Experimental design and evaluation protocol

This subsection describes the experimental design and evaluation protocol adopted to ensure fair, transparent, and reproducible comparisons among the three control scenarios. Particular attention was given to isolating the effects of the control strategy from those of the physical model and climatic forcing. To this end, all experiments shared identical environment dynamics, weather generation procedures, and evaluation metrics, while differing only in the control formulation. The protocol further emphasizes configuration-driven reproducibility, controlled randomness, and consistent performance assessments across independent runs. All reported hyperparameters correspond to a stable region identified through preliminary tuning and are not optimized per scenario.

3.5.1. Configuration-driven reproducibility

To ensure transparent and reproducible experiments, parameters are centralized in a configuration module separating (i) environment parameters (season length T, $I_{\max}$, seeds), (ii) soil parameters ($S_{\max}$, S_{fc}, retention curve parameters, drainage and efficiency η_I), (iii) weather parameters (ET0 seasonality and rainfall generator), and (iv) training parameters (total interaction budget for PPO). This design supports controlled sensitivity analyses and aligns with the EMS best practices for reproducible assessments in Environmental Modelling and Software.

3.5.2. Training and evaluation separation

To ensure a fair, interpretable, and reproducible comparison across the control paradigms, a strict separation was enforced between the controller configuration, training (when applicable), and evaluation for all scenarios. Although only Scenarios 2 and 3 involved explicit learning, all scenarios were evaluated under identical physical, soil, and meteorological conditions using the same performance indicators.

Scenario 1: Physics-based model with rule-based irrigation control. Scenario 1 implements a baseline irrigation strategy based on expert-defined rules interacting with a physics-based soil–water bucket model. This scenario serves as a reference case, isolating the effect of heuristic control without any learning or adaptive policy optimization.

The simulation is defined over a growing season of fixed length T (days), with climatic forcing generated deterministically from a prescribed random seed. Daily weather inputs include rainfall R_t, reference evapotranspiration $\mathrm{ET0}_t$, and crop coefficient $K_c(t)$. The soil system is represented by a conceptual bucket model parameterized by a water retention curve, a drainage function, and an irrigation efficiency coefficient η_I. Unless specified otherwise, default soil parameters are used.

At the beginning of the season, soil water storage is initialized at field capacity,

$$S_0 = S_{\mathrm{fc}}, \tag{18}$$

and converted to soil water tension via the retention relationship $\psi_0 = S \to \psi(S_0)$.

At each day t, irrigation is determined by a predefined rule function

$$I_t = g(\psi_t, I_{\max}, \widehat{R}_{t+1}), \tag{19}$$

where ψ_t denotes the current soil water tension, $I_{\max}$ is the maximum admissible daily irrigation depth, and $\widehat{R}_{t+1}$ is a one-day-ahead rainfall forecast (when available). The rule function may implement a single tension threshold or a comfort-band strategy, and internally clips the action to the feasible range $I_t \in [0, I_{\max}]$.

The physical soil–water dynamics are then updated deterministically. Crop evapotranspiration is computed as

$$\mathrm{ETc}_t = K_c(t)\,\mathrm{ET0}_t\, f_{\mathrm{ET}}(\psi_t), \tag{20}$$

where $f_{\mathrm{ET}}(\psi_t)$ is a stress reduction factor derived from the soil model. Drainage losses D_t occur when soil storage exceeds field capacity. The daily water balance is given by

$$S_{t+1} = \mathrm{clip}(S_t + \eta_I I_t + R_t - \mathrm{ETc}_t - D_t,\ 0,\ S_{\max}), \tag{21}$$

with clip($\cdot$) enforcing physical bounds on soil water storage. The updated soil water tension is obtained via the inverse retention relation $\psi_{t+1} = S \to \psi(S_{t+1})$.

This procedure is repeated sequentially for $t = 0, \ldots, T-1$, producing time series of soil storage, soil tension, irrigation, evapotranspiration, and drainage. All actions are fully determined by the irrigation rule and the current system state; no policy network, learning mechanism, or optimization procedure is involved.

Scenario 1 therefore provides a transparent and interpretable benchmark that reflects common rule-based irrigation practices, against which the benefits of reinforcement learning control (Scenario 2) and hybrid neuro-physical control (Scenario 3) can be systematically assessed.

Scenario 2: Physics-based model with reinforcement learning control (PPO). In Scenario 2, irrigation control is achieved through a reinforcement learning agent interacting directly with the physics-based soil–water model described in Section 3.3. The agent is trained using Proximal Policy Optimization (PPO), as implemented in the Stable-Baselines3 library (Raffin et al., 2021), without any custom policy architecture or parameterization.

The control policy is represented by the standard `MlpPolicy` provided by Stable-Baselines3. At each decision step t, the agent receives a continuous-valued observation vector

$$\mathbf{o}_t = [\psi_t,\ S_t,\ R_t,\ \mathrm{ET0}_t], \tag{22}$$

where ψ_t denotes soil water tension, S_t the soil water storage, R_t the rainfall input, and $\mathrm{ET0}_t$ the reference evapotranspiration. These variables jointly characterize the hydrological state of the system and the prevailing climatic conditions.

The policy network consists of two fully connected hidden layers with 64 units each and ReLU activation functions, corresponding to the default PPO configuration in Stable-Baselines3. A shared feature extractor feeds two output heads: a policy head, which outputs the mean of a Gaussian distribution over the one-dimensional continuous action space, and a value head, which estimates the scalar state-value function $V(\mathbf{o}_t)$. In addition, PPO maintains a learnable log-standard deviation parameter for the action distribution.

The irrigation action I_t is sampled from the Gaussian policy, squashed through a hyperbolic tangent function, and rescaled to satisfy the operational constraints of the irrigation system:

$$I_t \in [0, I_{\max}], \tag{23}$$

resulting in a continuous irrigation dose expressed in millimeters. This action is then applied to the physics-based soil–water model, which updates the system state according to the water balance equations.

In this scenario, the reinforcement learning agent learns irrigation strategies solely through interaction with the fixed physical model, without any correction or augmentation of the underlying system dynamics. Scenario 2 therefore isolates the contribution of learning-based control, providing a principled comparison with the rule-based strategy of Scenario 1 and the hybrid neuro-physical formulation introduced in Scenario 3.

Scenario 3: Hybrid environment with Neural ODE residual and PPO. Scenario 3 follows a two-stage learning protocol that explicitly separates model identification from policy optimization.

Stage 1: Pretraining of the Neural ODE residual model. Prior to reinforcement learning, the residual dynamics model was pre-trained in a supervised manner using simulated trajectories generated from a physics-based environment. The Neural ODE (implemented here as a discrete-time residual model) learns to predict a one-day correction $\Delta\psi_t$ to the soil water tension based on inputs $(\psi_t, I_t, R_t, ET0_t)$. Training uses a fixed number of trajectories (typically 32), over 50 epochs, with a batch size of 256 and a learning rate of 10^{-3}. The objective is to minimize the discrepancy between the physical prediction and perturbed reference trajectory, yielding a stable residual model prior to control learning.

Once pretrained, the Neural ODE parameters are frozen and embedded within the environmental dynamics.

Stage 2: PPO training on the hybrid environment. The PPO agent is then trained in the hybrid environment (physics + Neural ODE correction) using the same algorithmic structure as in Scenario 2. The training budget is again defined by a fixed number of interaction steps (e.g., 50,000 timesteps), and the policy architecture remains a multilayer perceptron with continuous outputs. The PPO hyperparameters (learning rate, discount factor, clipping range, and GAE parameters) were kept consistent with Scenario 2 to isolate the effect of the hybrid dynamics.

Evaluation protocol. For all three scenarios, an evaluation was conducted after configuration or training using identical soil parameters, weather realizations, and initial conditions. No learning, adaptation, or parameter tuning was performed during the evaluation. Performance metrics, including soil water tension dynamics, irrigation volumes, drainage losses, and aggregated efficiency indicators, were computed over full growing seasons.

This strict separation between configuration, training, and evaluation ensures that the observed performance differences arise from the controller design and system representation (rule-based, physical RL, or hybrid), rather than from stochastic variability, online adaptation, or unequal exposure to environmental conditions. This also reflects realistic deployment settings, where irrigation policies are typically calibrated or trained offline and then applied operationally without continuous retraining.

3.5.3. Performance indicators

The model performance was evaluated using both trajectory-level and aggregated indicators derived from the seasonal simulations. All indicators are consistently defined across scenarios and are directly linked to the notation summarized in Table 1.

At the trajectory level, we analyzed the temporal evolution of soil water tension ψ_t and soil water storage S_t to assess the occurrence, duration, and severity of water stress episodes, as well as depletion and recovery dynamics within the root zone. These trajectories provide insights into the controllers' ability to regulate soil water status under stochastic climatic forcing.

At the aggregated level, several seasonal performance metrics were computed as follows: (i) The mean soil matric potential $\bar{\psi}$ summarizing the overall stress conditions; (ii) The fraction of days spent within an agronomically optimal tension range, denoted τ_{opt}; (iii) Total irrigation volume $I_{\text{tot}} = \sum_t I_t$; (iv) Cumulative drainage losses $D_{\text{tot}} = \sum_t D(S_t)$; (v) and Water-use efficiency metric Eff defined as the ratio between productive evapotranspiration and total water inputs.

Together, these indicators capture the key trade-offs between stress avoidance, water-use efficiency, and hydrological loss. They are used uniformly across scenarios to ensure that the observed performance differences can be attributed to the controller design rather than confounding variations in physical parameters or climatic forcing.

3.5.4. Notation summary

To avoid ambiguity across the modelling, control, and learning components, Table 1 summarizes the notation used consistently throughout Section 3.

Table 1: Notation used throughout Section 3.

Symbol	**Unit**	**Description**
t	day	Discrete time index ($t = 0, \ldots, T-1$)
T	day	Length of the growing season (time horizon)
Soil–water state variables		
S_t	mm	Soil water storage in the root zone (latent physical state)
$S_{\max}$	mm	Maximum soil water storage (soil capacity)
S_{fc}	mm	Soil water storage at field capacity
ψ_t	cbar	Soil matric potential (tension), observable via tensiometers
$f_{\mathrm{retention}}$	–	Soil water retention function linking $S_t \leftrightarrow \psi_t$
Hydrological fluxes		
I_t	mm	Irrigation depth applied at day t (control action)
$I_{\max}$	mm	Maximum allowable daily irrigation depth
R_t	mm	Rainfall at day t
$ET0_t$	$\mathrm{mm\,day^{-1}}$	Reference evapotranspiration at day t
Kc_t	–	Crop coefficient at day t
ETc_t	mm	Crop evapotranspiration ($ETc_t = Kc_t \cdot ET0_t \cdot f_{ET}(\psi_t)$)
$f_{ET}(\psi_t)$	–	Water-stress reduction factor for evapotranspiration
$D(S_t)$	mm	Drainage loss as a function of soil water storage
Mass balance and dynamics		
η_I	–	Irrigation efficiency coefficient
f_{phys}	–	Physics-based soil–water balance model

Continued on next page

Table 1 continued

Symbol	Unit	Description
f_{res}	–	Learned residual dynamics (Neural ODE component)
Decision-making and learning		
$\mathbf{o}_t$	–	Observation vector available to the controller
$\mathbf{o}_t$	–	$(\psi_t, R_t, ET0_t)$ (default observation setting)
a_t	mm	Control action selected by the policy ($a_t = I_t$)
$\pi(\cdot)$	–	Control policy (rule-based or learned)
π_{rule}	–	Deterministic rule-based irrigation policy (Scenario 1)
π_θ	–	Parametric RL policy with parameters θ (Scenarios 2–3)
Reinforcement learning formulation		
r_t	–	Immediate reward at day t
γ	–	Discount factor for future rewards
$V(\mathbf{o}_t)$	–	State-value function approximation
$\hat{A}_t$	–	Advantage estimate (GAE)
$J(\theta)$	–	Expected cumulative return optimized by PPO
Neural ODE residual model (Scenario 3)		
f_θ	–	Neural network parameterizing residual correction
$\Delta\psi_t$	cbar	Residual correction to soil water tension
ψ_{t+1}^{phys}	cbar	Physical model prediction of soil tension
ψ_{t+1}	cbar	Hybrid prediction: $\psi_{t+1}^{\text{phys}} + \Delta\psi_t$
Performance indicators)		
$\bar{\psi}$	cbar	Mean soil matric potential over the season
τ_{opt}	%	Fraction of days within optimal tension range
I_{tot}	mm	Total irrigation volume over the season
D_{tot}	mm	Total drainage loss over the season
Eff	–	Water-use efficiency metric $(ETc/(I + R))$

4. Results and discussion

This section presents and discusses the results obtained for the three irrigation control scenarios described in Section 3.4: (i) rule-based control,

(ii) reinforcement learning in a physics-based environment, and (iii) hybrid reinforcement learning with Neural ODE–augmented dynamics. All simulations were conducted over a full growing season under identical physical configurations and stochastic climatic forcing.

The experimental design and evaluation protocol followed the established best practices in EMS for transparent, reproducible, and interpretable model assessment. In particular, the model performance is characterized using complementary indicators that capture both system-level behavior and decision-relevant outcomes, in line with EMS recommendations for the evaluation of environmental models (Bennett et al., 2013). The use of fixed physical configurations, controlled stochastic forcing, and consistent training and evaluation settings across scenarios ensured that the observed performance differences could be attributed to the control strategy rather than experimental artifacts.

4.1. Scenario 1: Rule-based control — conservative stability

Figure 1 illustrates the seasonal dynamics obtained using rule-based irrigation strategies.

The results show that rule-based control maintains soil water tension predominantly within or close to the agronomically optimal range for the crop. This conservative behavior results in a high proportion of days spent in the comfort zone, reflecting the explicit enforcement of predefined thresholds or bands. However, irrigation actions are triggered reactively and often abruptly, leading to frequent water application.

Soil water storage remains close to field capacity for most of the season, but this stability is achieved at the expense of substantial drainage losses, particularly when irrigation coincides with rainfall events. These results indicate that although rule-based control is robust in terms of stress avoidance and interpretability, it suffers from inefficient water use under stochastic climatic conditions.

4.2. Scenario 2: Physics-based reinforcement learning — efficiency with risk

Figure 2 presents the seasonal trajectories obtained using the PPO controller interacting with the physics-based environment.

Compared with Scenario 1, the irrigation actions were smoother and significantly reduced in magnitude, leading to the lowest cumulative irrigation volume among the three scenarios. This resulted in the highest water-use efficiency. However, the soil water tension time series exhibited pronounced and persistent peaks during dry periods, indicating severe stress episodes.

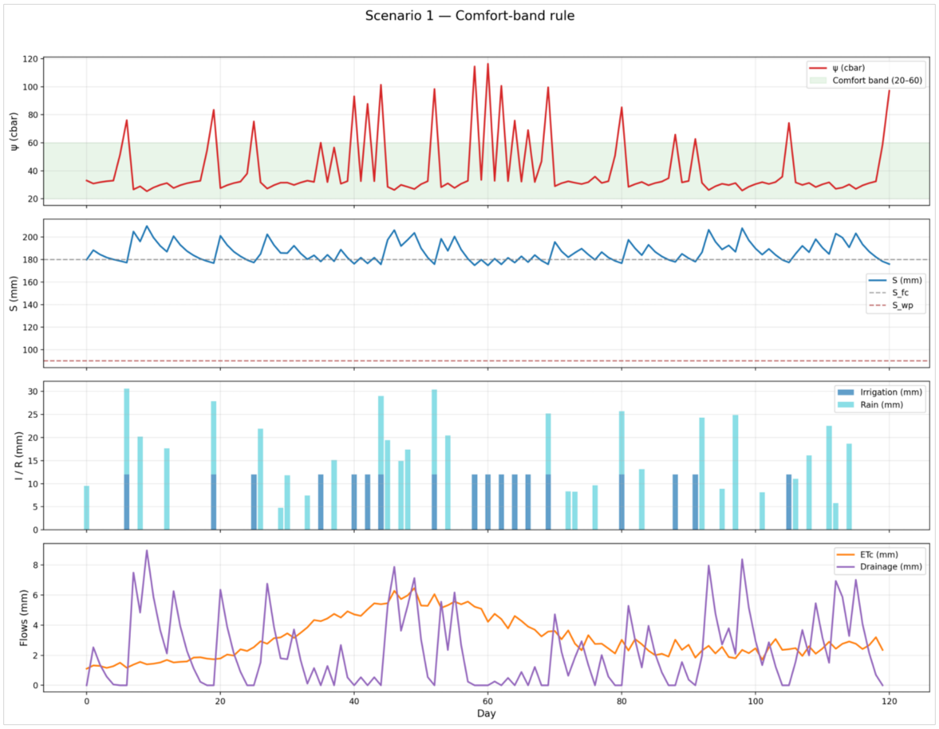

Figure 1: Seasonal dynamics of rule-based irrigation control (Scenario 1): soil water tension, soil water storage, irrigation and rainfall inputs, and hydrological fluxes.

These extreme tension values reveal the limitations of the controller's anticipation of cumulative water deficits when relying solely on a simplified physical model. Although drainage losses are reduced relative to the rule-based baseline, the occurrence of extreme stress raises concerns regarding agronomic reliability.

4.3. Scenario 3: Hybrid reinforcement learning with Neural ODE — moderated trade-offs

These extreme tension values reveal the limitations of the controller's anticipation of cumulative water deficits when relying solely on a simplified physical model. Although drainage losses are reduced relative to the rule-based baseline, the occurrence of extreme stress raises concerns regarding agronomic reliability.

Figure 3 shows the results obtained using the hybrid neuro-physical control strategy.

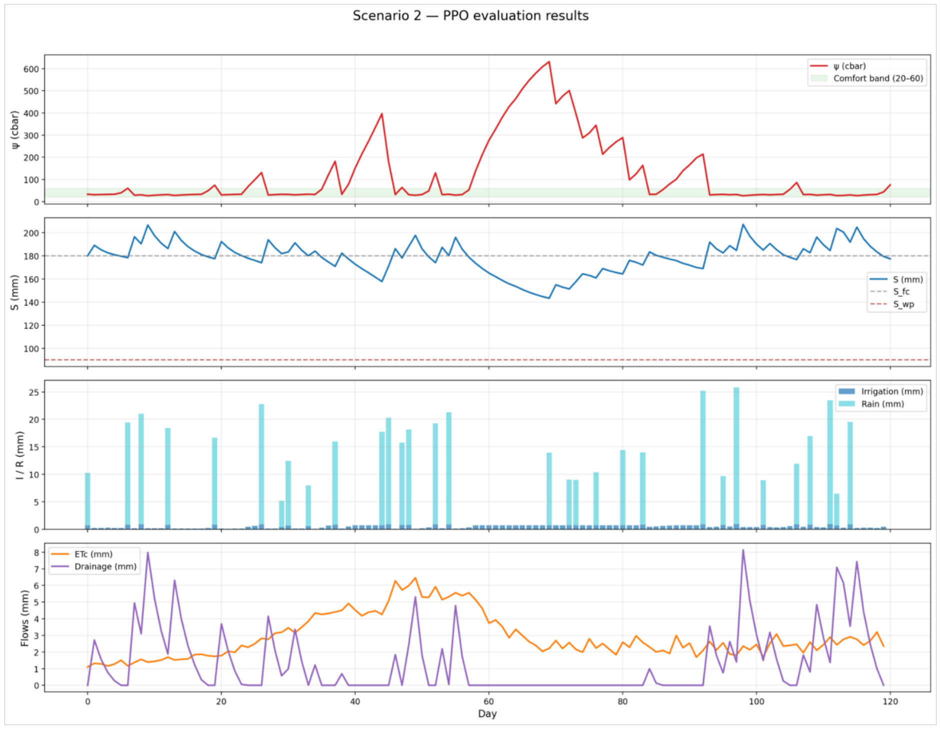

Figure 2: Seasonal dynamics of PPO-based irrigation control using the physics-based environment (Scenario 2).

The hybrid approach significantly reduced the severity and duration of the extreme soil water tension peaks observed in Scenario 2. Soil water storage trajectories are smoother and avoid deep depletion during prolonged periods of drought. Irrigation actions remain moderate and are better coordinated with rainfall events, resulting in reduced drainage losses.

Nevertheless, Scenario 3 did not maximize the time spent in the optimal tension range, which remained lower than that achieved by the rule-based strategy. Instead, the hybrid controller achieved a compromise between stress mitigation and water-use efficiency.

4.4. Soil and climatic parameterization on control performance

All reported results were obtained under a reference soil–climate configuration representative of moderately deep agricultural soil and semi-arid seasonal forcing. In particular, the root zone depth (Z_r = 600mm) and field capacity and wilting point values define a relatively large buffering ca-

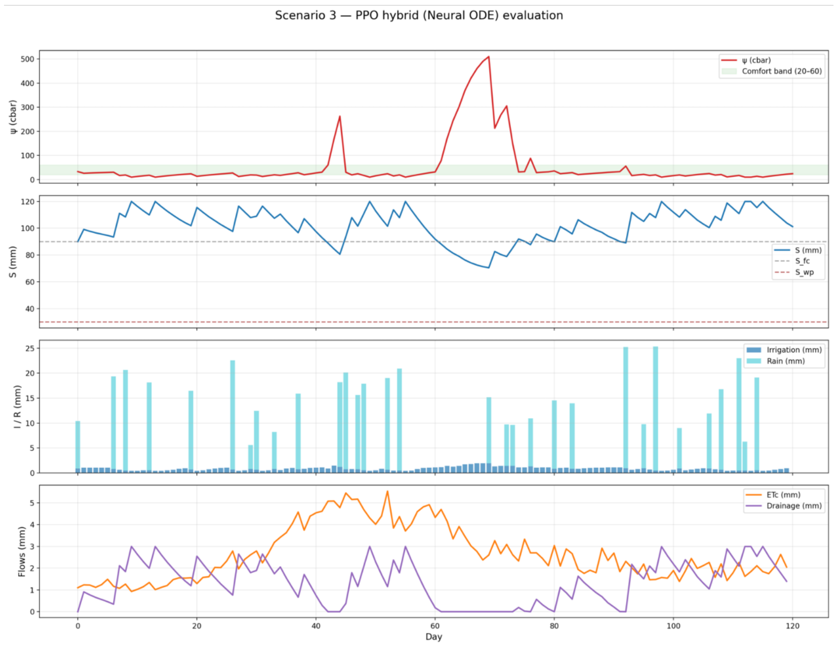

Figure 3: Seasonal dynamics of hybrid PPO control with Neural ODE–augmented dynamics (Scenario 3).

pacity, which favors control strategies capable of anticipating delayed stress responses. Under this configuration, excessive irrigation primarily manifests as drainage losses governed by the drainage coefficient ($k_d = 0.30$), which explains why aggressive rule-based strategies incur higher penalties without proportional stress reduction.

The stress threshold for evapotranstration reduction ($\psi_{\mathrm{ET}}^{crit} = 80$ cbar) further shapes the controller behavior by defining a narrow transition between non-stress and stress regimes, within which learning-based controllers exploit fine-grained irrigation adjustments. From a climatic perspective, the combination of moderate seasonal variability in reference evapotranspiration and stochastic rainfall probabilities across the growing season induces intermittent rather than persistent water deficits. This favors adaptive policies that balance short-term irrigation against the expected atmospheric demand rather than fixed threshold responses.

Therefore, these soil and climatic parameters define the operating regime in which performance differences between the control scenarios are observed and delimit the scope of generalization of the results. Alternative configurations, such as shallower soils, higher drainage capacities, or more arid rainfall regimes, may alter the relative advantages of rule-based, reinforcement learning, and hybrid neurophysical controllers.

Interactive configuration and reproducibility. In addition to the experimental results reported in this study, the proposed framework was implemented within a user-friendly interactive interface based on a Streamlit web application. This interface allows users to modify soil, climatic, and control parameters in real time, enabling the rapid exploration of alternative configurations and sensitivity analyses without altering the underlying codebase. Such an implementation facilitates reproducibility, supports scenario-based experimentation, and provides a practical bridge between the methodological contributions of this study and their potential use in decision support and educational contexts.

These effects are consistent with the soil and climatic sensitivity discussed in Section 4.4.

4.5. Comparative analysis across scenarios

To consolidate the scenario-wise analysis, we examined the comparative indicators derived from the seasonal simulations under the soil and climatic parameterization described in Section 3.5. In particular, the results reflected a moderately deep root zone ($Z_r = 600$ mm), a relatively narrow optimal soil water tension window ($\psi_{\text{fc}} \approx 33$ cbar, $\psi_{ET}^{\text{crit}} = 80$ cbar), and a non-negligible drainage sensitivity controlled by the coefficient $k_d = 0.30$. Climatic forcing combines moderate reference evapotranspiration variability with intermittent rainfall events that are representative of semi-arid to sub-humid growing conditions.

4.5.1. Comparison of soil water tension dynamics

Figure 4 compares soil water tension trajectories across the three scenarios.

Under the selected soil parameters, the rule-based controller (Scenario 1) maintained the soil water tension predominantly within the optimal range. This behavior is a direct consequence of conservative threshold settings relative to ψ_{ET}^{crit} and the relatively high irrigation efficiency ($\eta_I = 0.85$), which jointly favor frequent replenishment of root-zone storage.

In contrast, Scenario 2 exhibited pronounced stress peaks, with soil water tension occasionally exceeding the critical stress threshold. These extremes

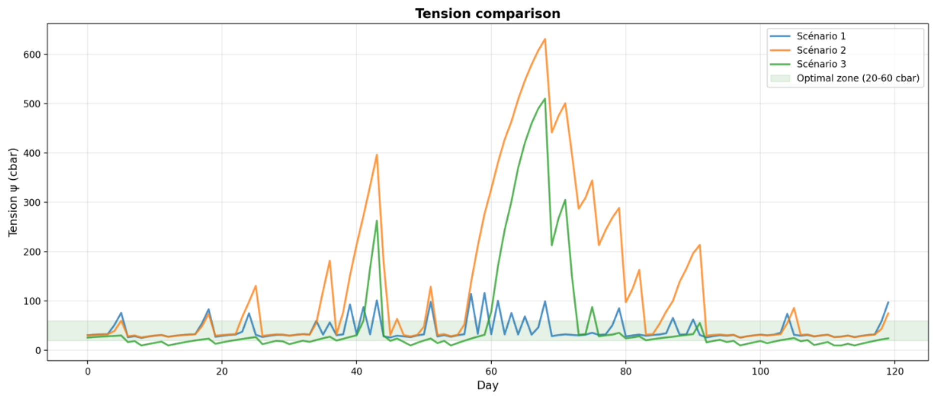

Figure 4: Comparison of soil water tension dynamics across the three scenarios. The shaded area indicates the agronomically optimal tension range.

arise from the interaction between (i) the finite root zone capacity, (ii) delayed hydrological responses embedded in the soil retention curve, and (iii) the reinforcement learning agent's incentive to minimize irrigation under stochastic rainfall. When rainfall realizations deviate from the expected patterns, the simplified physical model underestimates the risk of cumulative depletion, leading to delayed corrective actions.

Scenario 3 attenuates these extreme tension excursions relative to Scenario 2. Residual neural correction partially compensates for structural mismatches in evapotranspiration reduction and drainage response, reducing abrupt transitions into high-stress regimes. However, the hybrid controller does not fully reproduce the conservative behavior of Scenario 1, reflecting its objective of reshaping rather than eliminating the efficiency–robustness trade-off.

4.5.2. Comparison of soil water storage

Figure 5 shows the corresponding soil water storage trajectories.

The rule-based controller maintains storage close to field capacity (S_{fc}), which is consistent with its conservative irrigation logic and explains its limited exposure to stress. However, given the drainage coefficient $k_d = 0.30$, this operating regime also induces systematic deep percolation losses whenever rainfall or irrigation exceeds the short-term evapotranspiration demand.

Physics-based reinforcement learning (Scenario 2) allows for deeper soil water depletion, particularly during extended dry spells characterized by elevated $ET0_t$. This behavior was encouraged by the reward structure, which

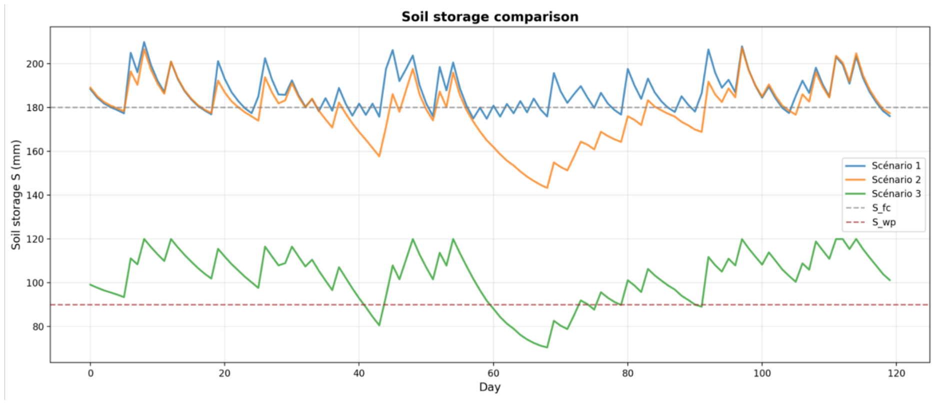

Figure 5: Comparison of soil water storage trajectories across the three scenarios, relative to field capacity and wilting point.

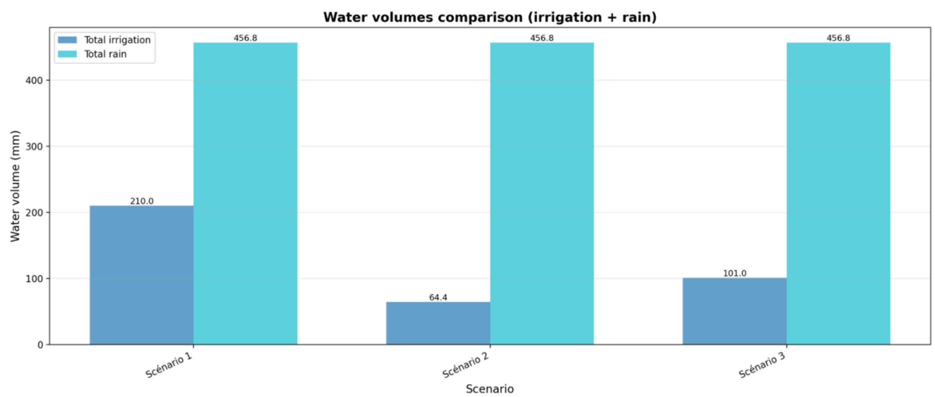

Figure 6: Comparison of cumulative irrigation and rainfall volumes across scenarios.

penalized irrigation volumes more strongly than transient stress. Consequently, the agent exploits the full dynamic range of the soil reservoir, but at the cost of occasional excursions toward the wilting point.

The hybrid controller (Scenario 3) moderates these dynamics. Residual corrections adjust the effective response of the physical model to storage depletion and drainage, leading to smoother trajectories that avoid excessive depletion and systematic over-irrigation.

4.5.3. Comparison of cumulative water volumes

Figure 6 compares the cumulative irrigation and rainfall volumes.

The rainfall contributions were identical across the scenarios. Therefore,

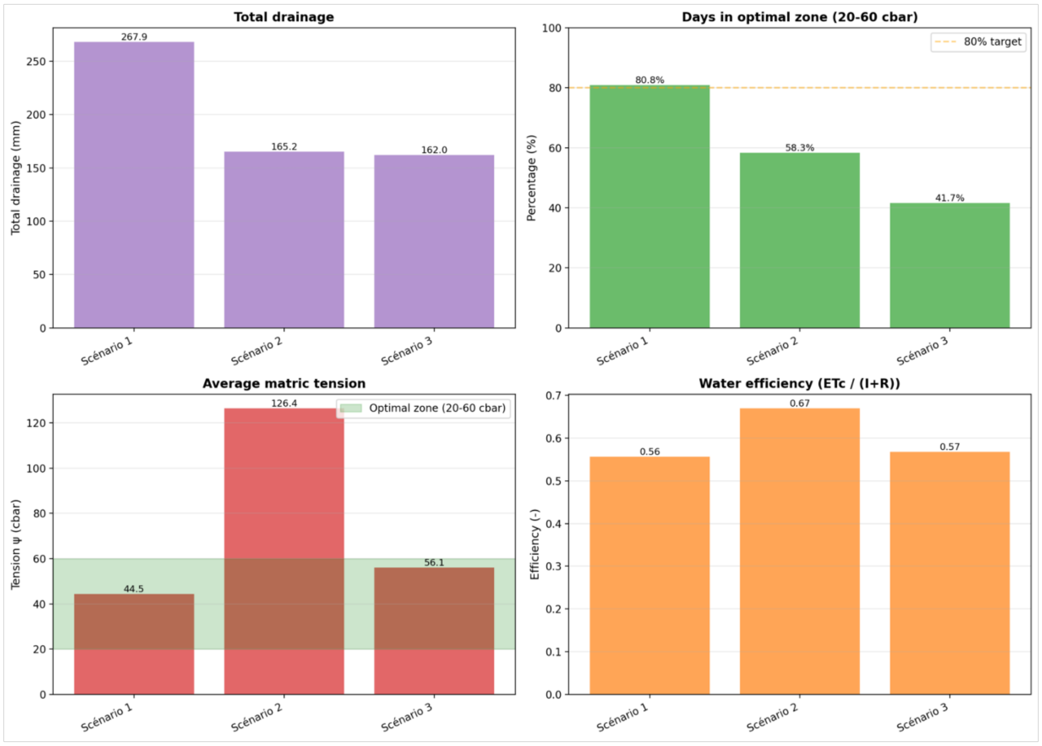

Figure 7: Comparison of aggregated performance indicators: total drainage, percentage of days in the optimal tension range, mean soil water tension, and water-use efficiency.

differences in cumulative irrigation volumes reflect purely control-induced behavior. Scenario 1 applied the largest irrigation depth, consistent with its conservative threshold logic and implicit prioritization of stress avoidance under uncertain rainfall.

Scenario 2 achieved the lowest irrigation volume by exploiting rainfall variability and soil storage capacity to maximize water-use efficiency. However, under the selected parameterization, this strategy exposes the system to a higher stress variability. Scenario 3 occupies an intermediate position, confirming that the hybrid correction reshapes irrigation demand without reverting to an overly conservative behavior.

4.5.4. Comparison of aggregated performance indicators

Figure 7 summarizes the aggregated performance

Rule-based control (Scenario 1) maximizes the fraction of days within the optimal tension range, but at the expense of high drainage losses, a direct consequence of maintaining storage near S_{fc} in a soil with non-negligible percolation sensitivity. Physics-based reinforcement learning (Scenario 2) achieved the highest water-use efficiency by tolerating larger excursions in soil water tension but consequently exhibited the poorest stress indicators.

The hybrid neuro-physical approach (Scenario 3) reduces both drainage losses and extreme stress events. By correcting systematic biases in the physical model response, particularly under high evaporative demand or near-capacity storage, a more balanced compromise between efficiency and robustness is achieved.

4.5.5. Comparative synthesis

Across all indicators, a consistent pattern emerged that was tightly linked to the chosen soil and climatic parameters. Rule-based control prioritizes stress avoidance by design, which is advantageous in shallow or highly stress-sensitive soils, but leads to inefficient water use when drainage sensitivity is significant. Physics-based reinforcement learning prioritizes efficiency by exploiting soil storage and rainfall variability but is risk-prone under simplified dynamics and stochastic forcing.

Hybrid neuro-physical reinforcement learning reshapes this compromise by correcting structural mismatches in the physical model rather than enforcing conservative control. Importantly, the Neural ODE component does not aim to maximize the time spent in agronomic comfort zones; instead, it reduces the delayed or exaggerated responses that arise from simplified evapotranspiration and drainage representations.

From an operational perspective, these results suggest that hybrid controllers are particularly well suited to contexts where basic physical knowledge is available, but soil behavior and climatic forcing remain unknown. Rule-based strategies remain attractive for low-risk or highly regulated settings, whereas learning-based and hybrid approaches are becoming increasingly relevant under tighter water constraints, higher climatic variability, and evolving soil conditions. Training costs and data availability remain practical limitations, motivating future studies on incremental learning, sensor-driven calibration, and transfer across sites.

4.5.6. Limitations

Several limitations remain and define clear directions for future research. These limitations are closely linked to the soil and climatic parameterization adopted in this study, which was intentionally chosen to enable controlled and reproducible comparisons across control strategies.

First, the controllers rely on Markovian state representations with a fixed daily resolution. Although this choice is consistent with the availability of irrigation decisions and meteorological inputs at the daily scale, it limits the representation of longer-term dependencies, delayed hydrological responses, and irregular temporal sampling. This limitation is particularly relevant for soil configurations with larger root-zone depths or slower drainage dynamics, where the impact of past deficits may propagate over multiple days. Therefore, future work will explore Neural Controlled Differential Equations to better capture the continuous-time dynamics driven by sparse or irregular sensor observations.

Second, climatic forcing is represented by a simplified stochastic weather generator parameterized by seasonal reference evapotranspiration patterns and rainfall occurrence probabilities. Although this design allows systematic stress testing under controlled variability, it does not capture multiyear persistence, regime shifts, or compound extremes. Consequently, anticipatory behavior remains limited, particularly under parameter settings with high evaporative demand or low rainfall frequency. Integrating time-series foundation models, such as PatchTST, for weather forecasting would enable more realistic climate-aware decision-making and support evaluations under non-stationary forcing.

Finally, the absence of an explicit world model constrains long-horizon planning and counterfactual analyses. Although the current formulation focuses on reactive control under uncertainty, more advanced planning capabilities could become critical under tighter water constraints, stronger drainage sensitivity, or deeper soil profiles. Future extensions of latent world-model architectures that jointly learn system dynamics, uncertainty, and control are therefore envisioned. These extensions are intended to increase the realism and robustness of stress testing rather than guarantee systematic performance improvements.

Effect of increased physical complexity. The soil–water balance model used in this study was intentionally simplified to isolate the effects of the control strategies under identical soil and climatic forcing. The selected parameterization assumes a homogeneous root zone, a single storage variable, and lumped representations of the drainage and evapotranspiration stress. Incorporating additional physical complexity, such as soil heterogeneity, layered profiles, preferential flow, or dynamic crop growth affecting root water uptake and evapotranspiration, would increase state uncertainty, introduce longer memory effects, and amplify nonlinear interactions.

Under these conditions, rule-based strategies (Scenario 1) would likely require repeated retuning of tension thresholds and irrigation doses across soils and seasons. This sensitivity is particularly pronounced in soils with higher drainage coefficients or narrower optimal tension ranges, where small parameter mismatches can induce oscillatory irrigation behavior and increased losses. Physics-based reinforcement learning (Scenario 2) may become more sensitive to structural model mismatches when key nonlinearities are omitted, for instance, when evapotranspiration reduction or redistribution processes are misrepresented. Without explicit mechanisms to handle parameter uncertainty (e.g., domain randomization), this could lead to increased stress.

The hybrid approach (Scenario 3) is expected to remain advantageous primarily by correcting systematic residual errors induced by simplified physics, such as delayed stress onset or underestimated drainage. However, its benefits are likely to manifest as improved robustness and safer behavior under mismatches rather than universal dominance across all performance metrics. These considerations motivate a progressive validation protocol in future work, in which physical fidelity is increased stepwise (layered soils, heterogeneous fields, crop modules), and policies are evaluated for transfer performance and robustness to physically meaningful perturbations.

Climatic forcing and non-stationarity. The stochastic weather generator employed in this study introduced variability through seasonal patterns and noise in rainfall occurrence and reference evapotranspiration. While this setup enables controlled comparisons across scenarios, it does not explicitly represent regime shifts, interannual persistence, or long-term trends. Consequently, the robustness of the learned policies to non-stationary climatic conditions remains only partially assessed.

A more stringent evaluation would involve testing controllers under distinct climatic regimes (e.g., arid, temperate, and tropical) using historical weather records or synthetic trajectories derived from climate model outputs. Under such conditions, fixed rule-based strategies would likely require repeated recalibration, particularly when rainfall frequency or evaporative demand departs from the assumptions embedded in rule parameters. Learning-based controllers would face distribution shifts relative to their training data, potentially degrading their performance unless adaptation mechanisms are introduced.

The hybrid neuro-physical approach is expected to be particularly relevant in these settings, as residual learning can correct persistent biases induced by unmodeled climate effects. However, this advantage is contingent on explicitly accounting for nonstationarity during training or adaptation. Therefore, future work will focus on cross-climate evaluation and transfer, in line with EMS best practices for robustness analysis under climatic uncertainty.

Residual neural correction. The residual neural component is used to model discrepancies that are not captured by simplified soil–water balance equations. In the present implementation, this component is trained in a supervised manner prior to reinforcement learning, using simulated trajectories that inject stochastic perturbations into the physical model to emulate the unmodeled effects and observation noise. The residual model was then kept fixed during policy training and applied in the inference mode within the environment.

The residual function f_{res} is parameterized as a lightweight multilayer perceptron that operates in discrete time. Its inputs consist of the current soil water tension, control action, rainfall, and reference evapotranspiration, and its output corresponds to an additive correction of the physical prediction. Although this design preserves physical interpretability and computational efficiency, it does not explicitly model uncertainty in the correction itself, nor does it adapt online to changing conditions.

Future extensions will explore residual learning strategies that incorporate uncertainty quantification, online adaptation, and continuous-time formulations, particularly when higher-frequency sensor data become available.

Transferability to other environmental systems. The proposed framework is transferable in terms of its modelling and control structure, rather than through the direct reuse of a trained policy. Applying it to another environmental system requires identifying four analogous components: (i) a process-based model f_{phys} describing dominant system dynamics (e.g., mass balance equations for reservoir storage or atmospheric transport models for air quality); (ii) control actions u_t representing management decisions; (iii) stochastic drivers d_t capturing exogenous forcing; and (iv) an objective function encoding system-specific trade-offs.

Within this structure, learning-based or hybrid controllers can be used to reshape trade-offs under uncertainty by complementing simplified physics with data-driven adaptation. Transferability therefore lies in the general decision-support paradigm—combining process knowledge, stochastic forcing, and adaptive control—rather than in assuming that a specific irrigation policy or neural architecture can be directly deployed across domains.

5. Conclusion and perspectives

This study investigated the potential of physics-informed reinforcement learning for intelligent irrigation control under stochastic climatic forcing. Three control paradigms were systematically compared within a common physics-based soil–water environment: a rule-based heuristic strategy, reinforcement learning interacting directly with a simplified physical model, and a hybrid neuro-physical approach in which the physical dynamics were augmented by a learned residual correction. The comparison was conducted under identical soil parameterizations and meteorological forcing, enabling a controlled assessment of how increasing levels of learning affect the irrigation performance.

The results confirmed that no single strategy dominated all performance criteria. Rule-based control provides strong guarantees in terms of stress avoidance and interpretability, particularly for soil configurations with moderate storage capacities and conservative tension thresholds. However, its reactive nature leads to excessive irrigation and drainage losses, particularly when drainage sensitivity or rainfall variability increases. Physics-based reinforcement learning substantially improves water-use efficiency by exploiting the delayed hydrological responses and cumulative effects encoded in the soil–water balance. However, under the chosen soil and meteorological parameters, this efficiency gain comes at the cost of increased vulnerability to extreme stress episodes, revealing sensitivity to structural model mismatch and limited anticipation of climatic variability.

The hybrid neuro-physical approach reshapes this trade-off by attenuating extreme stress events and reducing drainage losses while preserving most of the efficiency gains achieved by the learning-based control. Importantly, this improvement emerged across the explored ranges of soil storage, drainage behavior, and stochastic meteorological forcing, indicating enhanced robustness within the considered parameterization rather than universal dominance. By correcting the systematic discrepancies induced by simplified physics, such as delayed stress onset or underestimated losses, the residual neural component complements, rather than replaces, the physical model.

Beyond quantitative performance, a key contribution of this study lies in clarifying *when* and *why* learning-based controllers outperform heuristic rules. Learning becomes advantageous when irrigation objectives extend beyond strict constraint satisfaction and require balancing competing objectives, such as stress mitigation, water savings, and loss reduction, under variable climatic forcing. In such contexts, fixed thresholds tuned for specific soil or weather conditions become brittle, whereas learning-based policies can adapt decisions to evolving system trajectories shaped by soil properties and meteorological demands. The role of the Neural ODE is shown to be corrective rather than substitutive: by learning residual dynamics on top of an interpretable physical core, the hybrid controller improves robustness to modelling simplifications while maintaining physical consistency and transparency.

From an operational perspective, the results suggest that different control paradigms may be appropriate depending on soil characteristics, climatic variability, and the availability of data. Rule-based strategies remain attractive in low-risk, low-complexity settings with stable soil and limited sensing infrastructure. Learning-based and hybrid approaches are becoming increasingly relevant as evaporative demand increases, rainfall becomes more variable, and soils exhibit stronger nonlinear responses through drainage and stress functions. In such settings, the ability to adapt to parameter uncertainty and stochastic forcing outweighs the additional computational and data requirements, particularly when incremental training or transfer-learning strategies can be employed.

Several limitations of the present study provide clear directions for future research. First, the current daily Markovian state representation does not fully capture the long-term dependencies introduced by deeper soils, layered profiles, or delayed redistribution processes. The integration of neural-controlled differential equations offers a principled pathway for modeling continuous-time dynamics driven by sparse and asynchronous sensor observations. Second, reliance on simplified stochastic weather generators constrains anticipatory capabilities; coupling the framework with data-driven weather forecasting models or climate model outputs would enable a more stringent robustness assessment under non-stationary forcing. Third, the absence of an explicit world model limits planning and counterfactual analysis, motivating future extensions toward latent world-model architectures that jointly learn the dynamics, uncertainty, and control.

In conclusion, this study demonstrates that physics-informed learning provides a powerful and flexible framework for navigating the trade-offs inherent in intelligent irrigation under uncertainty. Rather than replacing physical modelling or agronomic expertise, hybrid neuro-physical approaches offer a complementary pathway toward adaptive, robust, and interpretable decision-support systems for climate-resilient water management, grounded in explicit soil and meteorological parameterization.

References

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environmental Modelling & Software 40, 1–20. doi:10.1016/j.envsoft.2012.09.011.

Berkenkamp, F., Schoellig, A.P., Krause, A., 2017. Safe model-based reinforcement learning with stability guarantees. Advances in Neural Information Processing Systems 30, 908–918.

Beucler, T., et al., 2021. Implicit learning of convective organization explains precipitation stochasticity. Nature 597, 672–677. doi:10.1038/s41586-021-03860-w.

Fatichi, S., et al., 2016. Ecosystem and land surface modelling in a changing climate. Hydrology and Earth System Sciences 20, 455–478. doi:10.5194/hess-20-455-2016.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., Reed, P.M., 2016. Coping with deep uncertainty in water management: Policy search under uncertainty. Environmental Modelling & Software 81, 60–74. doi:10.1016/j.envsoft.2016.02.006.

Giuliani, M., et al., 2021. Reinforcement learning and control of water systems: An overview. Environmental Modelling & Software 141, 105045. doi:10.1016/j.envsoft.2021.105045.

Hadka, D., Reed, P.M., 2013. Borg: An auto-adaptive many-objective evolutionary computing framework. Environmental Modelling & Software 37, 97–111. doi:10.1016/j.envsoft.2012.07.004.

Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software 21, 602–614. doi:10.1016/j.envsoft.2006.01.004.

Jones, H.G., et al., 2022. Smart irrigation systems: A review of control strategies and technologies. Agricultural Water Management 260, 107300. doi:10.1016/j.agwat.2021.107300.

Karniadakis, G.E., et al., 2021. Physics-informed machine learning. Nature Reviews Physics 3, 422–440. doi:10.1038/s42254-021-00314-5.

Perkins, S., et al., 2023. Safe reinforcement learning for real-world control systems: A survey. IEEE Transactions on Artificial Intelligence 4, 1–18. doi:10.1109/TAI.2022.3220730.

Rackauckas, C., Ma, Y., Martensen, J., Warner, P., Zubov, K., Supekar, S., Skinner, D., Ramadhan, A., Edelman, A., Perdikaris, P., 2020. Universal differential equations for scientific machine learning. Proceedings of the National Academy of Sciences 117, 29041–29048. doi:10.1073/pnas.2001336117.

Rackauckas, C., et al., 2021. Scientific machine learning through physics-informed neural networks and universal differential equations. Computing in Science & Engineering 23, 18–31. doi:10.1109/MCSE.2020.3042241.

Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2009. AquaCrop—The FAO Crop Model to Simulate Yield Response to Water. FAO, Rome.

Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., Dormann, N., 2021. Stable-baselines3: Reliable reinforcement learning implementations. Journal of Machine Learning Research 22, 1–8. URL: `https://www.jmlr.org/papers/v22/20-1364.html`.

Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.L., Vanrolleghem, P.A., 2007. Uncertainty in the environmental modelling process – a framework and guidance. Environmental Modelling & Software 22, 1543–1556. doi:10.1016/j.envsoft.2007.02.004.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven earth system science. Nature 566, 195–204. doi:10.1038/s41586-019-0912-1.

Rolnick, D., et al., 2022. Tackling climate change with machine learning. ACM Computing Surveys 55, 1–96. doi:10.1145/3485128.

Seneviratne, S.I., et al., 2021. Weather and climate extreme events in a changing climate. Nature Climate Change 11, 964–974. doi:10.1038/s41558-021-01092-9.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction. 2 ed., MIT Press, Cambridge, MA.

Willard, J., Jia, X., Xu, S., Steinbach, M., Kumar, V., 2022. Integrating physics-based modeling with machine learning: A survey. Nature Reviews Physics 4, 366–382. doi:10.1038/s42254-022-00437-4.

Yang, T., et al., 2021. Reinforcement learning for water resources management: A review. Water Resources Research 57, e2020WR028838. doi:10.1029/2020WR028838.