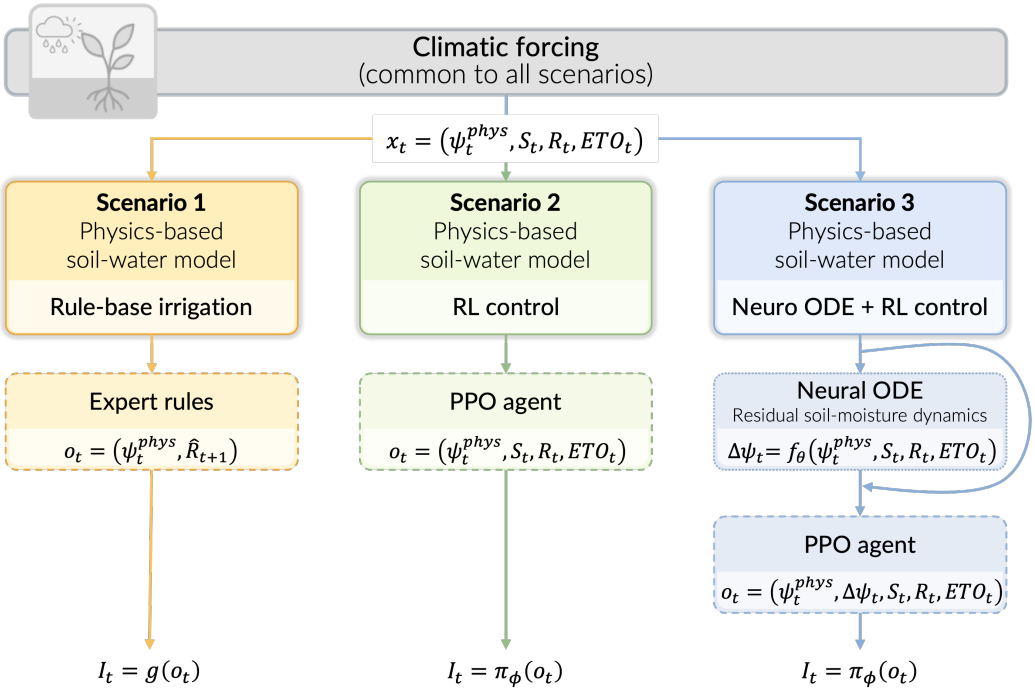
Graphical Abstract

Control-aware physics-informed reinforcement learning for adaptive irrigation under climatic uncertainty



Highlights

Control-aware physics-informed reinforcement learning for adaptive irrigation under climatic uncertainty

- A physics-informed RL framework is proposed for irrigation control.
- Rule-based, RL, and hybrid control strategies are compared under identical conditions.
- A Neural ODE-inspired residual model is integrated to correct daily prediction errors.
- Hybrid control mitigates extreme stress while preserving water-use efficiency within the tested soil and climatic parameter settings.

Control-aware physics-informed reinforcement learning for adaptive irrigation under climatic uncertainty

Abstract

Decision-making in physical systems often relies on process models. Although interpretable, these models may not accurately represent real-world dynamics. Standard reinforcement learning (RL) methods address uncertainty but can be unreliable, particularly in partially observed systems with slow response times or that are influenced by stochastic external factors. This study proposes a hybrid RL approach that integrates physics-based rules with a residual neural network to manage dynamic irrigation systems. The method explicitly distinguishes hidden physical states from sensor measurements and incorporates learned corrections into the environment model. This enables the controller to adapt when the model diverges from actual system behavior. To preserve physical accuracy and stability, RL is employed for decision-making, while learned corrections address discrepancies in system performance. Three decision-making strategies are evaluated under identical random external conditions: rule-based control, standard RL, and hybrid RL with corrections implemented via Neural Ordinary Differential Equations (ODE). The results indicate that model-based RL improves irrigation efficiency but may induce extreme soil stress. In contrast, the hybrid RL approach substantially mitigates these extreme states while retaining most of the irrigation efficiency, thus optimizing the trade-off between robustness

and efficiency. Although demonstrated on an irrigation control system, the proposed method applies to predictive maintenance, building energy management, and other partially observed systems characterized by delays or uncertainty.
Keywords:
Intelligent irrigation, Environmental decision support systems,
Control-aware modelling, Physics-informed reinforcement learning, Hybrid neuro-physical models, Adaptive control under uncertainty

---

1. Introduction

Climate change has led to more frequent and severe weather events, thereby increasing the complexity of water management in agriculture. Irrigation accounts for a significant share of global freshwater consumption and is influenced by variable rainfall, seasonal fluctuations, and complex soil–plant interactions (Siebert et al., 2010). Conventional irrigation methods often depend on fixed schedules or expert-defined rules, which lack the adaptability necessary to address changing environmental conditions (Padilla-Nates et al., 2025). Consequently, these methods may lead to inefficient water use, greater drainage losses, and increased crop water stress.

From a modeling perspective, irrigation management represents a complex sequential decision-making problem defined by uncertainty and physical constraints (Shang et al., 2018). To address these complexities, researchers have developed models that simulate soil and water processes, frequently utilizing simplified methods such as bucket models or FAO-56-type formulas (FAO, 2025; Raes et al., 2009). These models depend on assumptions and

parameterizations that may not adequately represent local soil heterogeneity or the effects of unpredictable weather (Portu et al., 2025).

Recent advances in RL have enabled the development of adaptive control methods that learn in complex, unpredictable environments (Sutton and Barto, 2018). Therefore, RL is increasingly utilized in irrigation and water management, where it has demonstrated superior water-use efficiency compared to fixed rule-based strategies (Alkaff et al., 2025). RL is particularly suitable for daily decision-making tasks that require adaptation to delayed feedback, while alternative algorithms are generally used for fixed or seasonal planning (Saikai et al., 2023). However, fully data-driven approaches still face challenges related to interpretability, compliance with physical laws, and operational safety. These issues are especially significant in environmental systems, where interventions may produce delayed and potentially irreversible effects (Liu et al., 2023).

Physics-informed learning improves data-driven models by integrating established physical principles, which enhances model stability and reliability (Willard et al., 2022). In this field, Neural Ordinary Differential Equations (Neural ODEs) provide a flexible means to model system components that are insufficiently captured by existing physical models, while preserving interpretability (Rackauckas et al., 2020). In environmental science, these hybrid models are used not to replace traditional ones, but to correct errors resulting from simplifications, parameter uncertainty, or omitted processes (Ratn et al., 2025).

Most irrigation systems operate on a daily decision-making schedule, relying on sensor data such as soil moisture and rainfall measurements (Tincani

et al., 2025). The effective use of learning-based methods necessitates synchronization among decision timing, data acquisition, and model operation (Nghiem et al., 2023). In this study, learning is implemented in two principal ways: (i) to guide irrigation decisions using RL, and (ii) to model temporal system dynamics with a neural network inspired by Neural ODEs. This methodology supports adaptive decision-making while preserving model interpretability and computational efficiency, in alignment with available data and system constraints.

Although interest in learning-based irrigation control is growing, several unresolved questions persist. These include the comparative effectiveness of RL and expert rule-based methods when combined with physical models, the capacity of learned models to overcome limitations of traditional approaches, and the influence of increased model–controller integration on interpretability, transferability, and practical deployment. The Environmental Modelling & Software (EMS) community places increasing emphasis on model transparency, explicit uncertainty management, and reproducible evaluation methodologies (Refsgaard et al., 2007; Jakeman et al., 2006). This study adheres to these principles by utilizing a structured experimental design, systematically varying soil and weather parameters, and maintaining a clear distinction between training and testing for learning-based controllers.

Accordingly, this study addresses the following research questions:

- **RQ1:** To what extent does RL enhance irrigation control performance relative to expert rule-based strategies when applied to a physics-based soil-water model?
- **RQ2:** To what extent does augmenting the physical model with learned

residual dynamics increase the stability and safety of learning-based irrigation control under stochastic climatic forcing?

- **RQ3:** What are the effects of increased model–controller integration on interpretability, transferability, and deployability in environmental decision-support systems?

The primary contributions of this study are as follows:

1. A physics-informed RL framework that formulates 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 strategies: rule-based control, RL-based control, and hybrid neuro-physical RL control, conducted within a unified physical simulation environment.
3. The integration of a discrete-time Neural ODE-inspired residual model to correct soil-water balance dynamics, thereby preserving physical consistency and interpretability.
4. A systematic evaluation under stochastic rainfall and evapotranspiration forcing, which highlights trade-offs among water-use efficiency, stress avoidance, and drainage losses as functions of soil and meteorological parameterisation.
5. A transferable modelling and control framework intended to facilitate future integration of real sensor data and extension to other environmental systems that require adaptive control under uncertainty.

The remainder of this paper is organized as follows. Section 2 reviews research on irrigation modeling, rule-based control, RL, and hybrid physics-informed approaches. Section 3 describes the materials and methods, including problem formulation, simulated data sources, the physical environment, control scenarios, and experimental design. Section 4 presents and discusses the results for each scenario, focusing on parameter sensitivity and performance trade-offs. Section 5 concludes the paper and outlines directions for future research.

2. Related Work

This section reviews literature on irrigation modelling and control, emphasizing the progression from process-based and rule-based methods to contemporary learning-based and hybrid physics-informed approaches. This study integrates neural ordinary differential equations with adaptive RL to enhance the robustness of irrigation control under changing climatic conditions, an area that has received limited attention in previous research works. The aim is to position this study within the broader field of environmental modelling and to identify the particular gaps that the proposed approach is designed to address.

2.1. Process-based modelling of irrigation systems

Process-based models are fundamental tools in irrigation modelling, representing interactions among soil, water, and plants through basic physical principles and evapotranspiration formulas (Bo et al., 2024). Examples include simple bucket models and FAO-56 methods, as implemented in tools such as AquaCrop (Raes et al., 2009). These models are transparent, require

minimal data, and are therefore valuable for scenario comparison and policy research (Starke and Lünich, 2020).

However, process-based models rely on simplified assumptions regarding soil heterogeneity, root water uptake, and boundary conditions. Their predictive accuracy declines under highly variable or extreme weather conditions, as they are unable to capture all relevant complexities and uncertainties (Verbruggen et al., 2025). Recent studies indicate that, although these physical models are essential, they are insufficient on their own for effective irrigation management in a changing climate (Fatichi et al., 2016; Seneviratne et al., 2021).

2.2. Rule-based and heuristic irrigation control

Rule-based irrigation methods, frequently derived from expert recommendations or agricultural guidelines, remain widely used. These systems employ fixed thresholds for soil moisture, crop stress, or water loss. Their simplicity, transparency, and ease of implementation make them advantageous in data-scarce environments (Fereres and Soriano, 2007).

However, these rule-based systems lack adaptability to changing climatic conditions (Padilla-Nates et al., 2025). Research demonstrated that fixed rules are often ineffective under unpredictable rainfall or when soil response is delayed, motivating the exploration of adaptive control methods (Jones et al., 2022). Consequently, rule-based control is now primarily regarded as an initial approach rather than a sustainable solution for climate-resilient irrigation.

2.3. RL for irrigation and water management

RL is a versatile framework for sequential decision-making in dynamic and uncertain environments (Sutton and Barto, 2018). In water management, RL has been applied to reservoir operation, canal control, and irrigation planning. Numerous studies have shown that RL can outperform fixed rule-based systems (Yang et al., 2021; Giuliani et al., 2021).

Despite these successes, data-driven RL methods also face significant challenges in environmental applications. They typically require extensive training data, may violate physical constraints, and can yield solutions that are difficult to interpret or trust (Li et al., 2022). Recent reviews have highlighted that ensuring physical consistency and reliability remains a major obstacle for RL in environmental systems (Rolnick et al., 2022; Reichstein et al., 2019).

In environmental modelling, adaptive control and policy search methods have been investigated to support decision-making under uncertainty and multiple objectives. Policy search approaches offer systematic means to balance robustness, efficiency, and risk in water management (Giuliani et al., 2016). Specialized algorithms, such as the Borg framework, can identify diverse optimal control strategies for multi-objective environmental problems (Hadka and Reed, 2013).

Within this context, RL is considered a valuable complement rather than a replacement for existing decision-making approaches in environmental management. RL shares conceptual similarities with policy search and multi-objective optimization but offers advantages in leveraging temporal information and managing delayed system responses (Padilla-Nates et al., 2025).

2.4. Physics-informed and hybrid learning approaches

To address the previously outlined challenges, physics-informed machine learning has become increasingly important. This approach incorporates physical knowledge into data-driven models. Research has demonstrated that integrating physical constraints enhances model performance, stability, and interpretability, particularly in data-limited scenarios (Willard et al., 2022; Karniadakis et al., 2021).

Neural ODEs provide an effective means of integrating data-driven and physical models. They enable neural networks to approximate unknown or unmodeled components in continuous-time systems (Rackauckas et al., 2020). In environmental science, Neural ODEs and related techniques are increasingly employed to augment, rather than replace, physical models, thereby maintaining interpretability while improving accuracy (Rackauckas et al., 2021; Shamekh et al., 2023).

2.5. Hybrid modelling for control under uncertainty

Physics-informed learning has been extensively studied for system identification and prediction, but its integration with RL for control in environmental modelling remains limited. Recent research on safe and model-based RL underscores the importance of incorporating physical constraints to enhance control reliability and prevent unsafe or unrealistic solutions (Berkenkamp et al., 2017; Perkins et al., 2023).

Few studies have systematically compared rule-based control, RL, and hybrid neurophysical control within a unified modelling framework. This research addresses this gap by evaluating these methods under identical physical and meteorological conditions, thereby ensuring fair and reproducible re-

sults. Each approach utilizes the same foundational data, such as evapotranspiration rates, soil type, and initial moisture, and is subjected to the same weather scenarios. The application of Neural ODEs to capture unmodeled dynamics and enhance irrigation control robustness under climate variability remains underexplored (Höge et al., 2022). Accordingly, this study investigates control strategies that increase model complexity within a physics-based irrigation context.

Policy search and multi-objective optimization methods have also been applied to irrigation management (Giuliani et al., 2016, 2021); however, this study distinguishes itself through the control setup and model construction. While most policy search approaches emphasize fixed or seasonal strategies, the present work focuses on daily control with real-time feedback and delayed system responses. The proposed framework explicitly separates the effects of learning-based control and model correction by comparing rule-based, RL, and hybrid neuro-physical methods under consistent conditions. This experimental design enables a controlled and transparent comparison of how increasing levels of learning and model integration affect irrigation system behavior. In contrast to prior studies that primarily emphasize aggregate performance gains, the proposed framework allows simultaneous assessment of performance, stability, and interpretability across scenarios under identical physical and climatic forcing.

Performance is evaluated using cumulative irrigation volumes, water-use efficiency, and drainage losses. Stability is assessed through the temporal smoothness of irrigation actions and the presence or absence of extreme soil-water tension excursions. Interpretability is examined qualitatively by com-

paring the structural complexity of the control mechanisms, ranging from explicit rule-based decisions to learned policies and residual corrections, and their consistency with known soil–water dynamics.

Together, these complementary metrics provide a coherent basis for analyzing how learning-based and hybrid approaches address key limitations of traditional rule-based irrigation control, particularly with respect to robustness under climatic variability and transparency of decision-making.(Huang et al., 2025).

3. Materials and Methods

This section details the materials and methods used in the study. The system dynamics, study scope, and data sources are first clarified. The irrigation management problem is then formulated. The physics-based soil–water environment for simulating daily dynamics is described, along with the three control scenarios: rule-based control, RL, and hybrid neuro-physical RL. The experimental design, training procedures, and evaluation protocols are provided to ensure reproducibility and enable fair comparison across control strategies.

3.1. Soil–water dynamics and physical assumptions

Figure 1 illustrates the main physical processes represented in the soil–water balance model used in all scenarios. The model treats the root zone as a control volume, accounting for water inputs from rainfall and irrigation, changes in storage, and losses due to evapotranspiration and drainage. Rainfall is treated as an external climate input, while irrigation is considered a

management action. Both inputs contribute to soil-water storage, which is constrained by the soil's physical capacity.

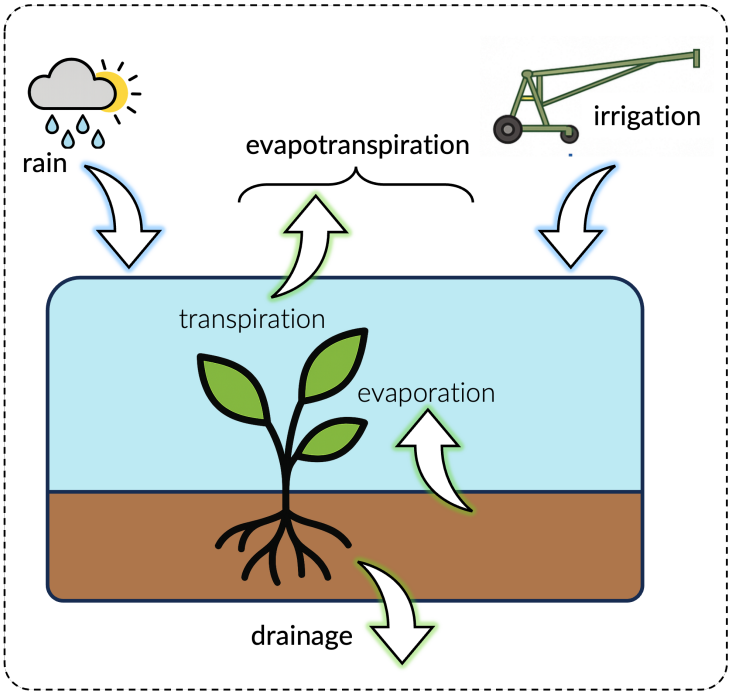


Figure 1: Schematic representation of the soil-water balance processes considered in the physical model, including water inputs from rainfall and irrigation, losses due to evapotranspiration (evaporation and transpiration), and drainage from the root zone.

Water losses are represented as evapotranspiration, which combines soil evaporation and plant transpiration into a single simplified process. This process is influenced by atmospheric demand and soil-water availability. Reference evapotranspiration is calculated from weather data and adjusted using crop coefficients and stress functions to reflect plant responses under water-limited conditions, following standard practice (Allen et al., 1998; Monteith, 1965). This methodology is widely used in irrigation models and crop

simulation tools, as it provides a straightforward yet realistic depiction of land–atmosphere exchanges (Raes et al., 2009).

Drainage occurs when soil-water storage exceeds field capacity, causing water to move out of the root zone and become unavailable to crops. Conceptual models typically represent this process using threshold-based or linear drainage functions, which capture the influence of soil properties without explicitly modeling the full vertical flow (Rodríguez-Iturbe and Porporato, 2004).

The soil is modeled as a single, uniform root-zone compartment with fixed storage capacity and a simplified relationship between soil-water storage and matric potential. This bucket-like model omits vertical heterogeneity, specialized flow paths, layered soil properties, and lateral water movement. While these simplifications improve usability, they do not capture actual soil processes that affect water retention and drainage. As a result, irrigation schedules applied to real conditions may be less accurate, since real soils are often layered and complex. Including these complexities would increase model difficulty and reduce interpretability. Therefore, a simplified model is chosen for clarity, with the understanding that this reduces realism (Vereecken et al., 2007; Fatichi et al., 2016). Previous research has shown that this approach can overestimate or underestimate water availability by 10–20% compared to more detailed models, emphasizing the need for calibration and sensitivity analysis when interpreting results.

Soil differences and nonlinear water movement create "memory effects", where past rainfall, irrigation, and evapotranspiration influence water availability over several days or weeks. These slow responses complicate estima-

tion of the soil state and irrigation control, particularly under unpredictable weather and limited sensor data (Seneviratne et al., 2010). Fully capturing these effects requires complex multi-layer or detailed ecohydrological models (Appendix A; Appendix B), which are computationally demanding and difficult to integrate with learning-based controllers.

The use of a simple model in this study is therefore a deliberate choice. By employing a straightforward, daily soil–water balance model, the performance of different control strategies and learning methods under uncertainty can be evaluated. This approach facilitates integration with RL and hybrid methods, while retaining the key water processes relevant for irrigation decisions. Consequently, all control and learning strategies are tested within the same physics-based environment, ensuring that performance differences are attributable to controller design rather than the physical model.

Appendix C provides a comparative analysis of the purposes and control compatibility of alternative soil–water modelling formulations. Appendix D demonstrates the stability and well-posedness of these formulations.

3.2. Data sources and scope of the study

The study employs data generated by a physics-based soil–water balance model driven by stochastic climate inputs. This simulation-based methodology enables controlled experimentation, systematic evaluation of management strategies, and reproducibility across diverse hydroclimatic conditions, while minimizing the influence of confounding factors such as limited data availability, measurement noise, and site-specific calibration requirements.

The simulated environment produces daily trajectories of rainfall, reference evapotranspiration, soil-water storage, and soil matric potential (ten-

sion), which together define the *state* and *observation* variables for the control algorithms. In all scenarios, irrigation decisions are evaluated using identical simulated physical dynamics and weather realizations, ensuring fair and consistent comparisons across control strategies. They defined the *action* variable.

Although the study is based on simulated data, the proposed framework is explicitly designed for transferability to real irrigation systems. In practice, soil-water storage is not directly measurable; instead, soil-water status is monitored using tensiometers that record soil matric potential ψ_t. For real-world implementation, simulated ψ_t values can be substituted with tensiometer measurements.Rainfall data can be sourced from on-site rain gauges or open-access meteorological datasets, such as those provided by national weather services. The observation structure, based on ψ_t and climatic variables, is constructed to accommodate these practical constraints.

Successful field deployment requires addressing several specific considerations: tensiometers must be calibrated and maintained to ensure accurate readings; real-time measurements should be integrated into a centralized data aggregation system accessible to the control framework. Key challenges include maintaining sensor connectivity and data accuracy across large agricultural areas. Investment in robust communication networks or cloud-based platforms may be necessary to support reliable data flow.

Potential discrepancies between simulated and real-world conditions, such as variability in soil properties and small-scale climatic differences, require site-specific tuning of control model parameters. The implementation of this study provides user-friendly interface to address this issue. Collaboration

with stakeholders is essential to address practical considerations, including equipment costs, training requirements, and sustained user participation. This research is being carried out in partnership with Rives et Eaux du Sud-Ouest (Tarbes, France), a company that focuses on the design and safety of hydraulic infrastructure to promote the long-term, equitable use of water resources. The skills of its engineers and technicians are mobilized to inform public policy and contribute to the ecological transition of local territories. This partnership include further conducting a pilot study to verify system effectiveness under actual field conditions, and using API weather service, before large-scale deployment.

3.3. Problem formulation

This subsection introduces a unified mathematical formulation of the irrigation control problem underlying all scenarios in this study. Irrigation management is formalized 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 establishes a common reference framework for rule-based control, RL, and hybrid neurophysical approaches, ensuring that performance differences arise from the control strategy rather than discrepancies in system representation.

3.3.1. System and time discretization

The analysis focuses on a single agricultural plot over a growing season of length T days, discretized into daily decision steps $t \in \{0, \ldots, T-1\}$.

Irrigation decisions are made once per day, consistent with the temporal resolution of meteorological forcing and the operational granularity targeted in this study.

Let S_t (mm) denote the *latent* root-zone soil-water storage at day t, representing the amount of plant-available water in the effective root zone. The term *latent* indicates that this variable is not directly measurable in typical operational settings. Instead, soil-water status is monitored using tensiometers that provide soil matric potential (tension) ψ_t (cbar), which serves as the primary observable variable.

The soil-water retention curve defines the link between latent storage and observed tension:

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

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 internally maintains and updates the soil-water storage S_t. In Scenarios 2 (RL) and Scenario 3 (hybrid physics-informed control), S_t may be included in the agent's observation vector to isolate the effect of the control strategy under identical physical dynamics and climatic forcing. This variable is not directly measurable in operational settings and is therefore considered simulator-accessible rather than sensor-accessible. Addressing partial observability requires exploring filtering techniques, such as Kalman filtering, or implementing recurrent policies that utilize historical observation data. Preliminary results indicate that incorporating belief-state estimation can improve control accuracy under limited data availability. However, chal-

lenges remain in designing efficient filtering strategies that operate within the constraints of real-time decision-making and field variability.

The proposed configuration offers an upper-bound assessment of learning-based control performance rather than a deployable solution. In practical applications, controllers rely on sensor-level observations such as soil-water tension, rainfall, and reference evapotranspiration. Extensions to partially observable settings, including belief-state estimation or recurrent policies, are identified as directions for future research.

3.3.2. Climate drivers

Daily climate inputs include rainfall R_t (mm), reference evapotranspiration $ET0_t$ (mmday^{-1}), and the crop coefficient Kc_t (dimensionless). We summarize these external factors as follows:

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

We treat these variables as random disturbances drawn from a distribution that may change over time:

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

3.3.3. Physics-based dynamics (mass balance)

A bucket-style soil-water mass balance model describes how water moves in the root zone:

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

Here, I_t (mm) is the irrigation depth applied as the control action, η_I (between 0 and 1) is the irrigation efficiency, S_{max} is the maximum allowed

storage, and $\dot{clip}()$ keeps values within physical limits. Drainage is calculated as:

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

Drainage usually occurs when storage goes above field capacity. Crop evapotranspiration is calculated using a method based on FAO guidelines:

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

Here, $f_{ET}(\psi_t)$ (ranging from 0 to 1) is a factor that reduces evapotranspiration based on soil tension. The next day's tension is calculated from storage using Eq. (1).

3.3.4. Sequential decision-making and objective

Each day t, the controller chooses how much water to apply for irrigation:

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

Here, $I_{\max}$ (mm) is the maximum amount of irrigation allowed per day.

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

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

Consequently, the problem remains partially observed in real-world operations. An MDP based on available observations is employed to ensure practicality and comparability. Extensions for partial observability are discussed in Section 4.

The goal of the control is to reduce crop water stress, reduce irrigation water use, and minimize water losses. We define the daily reward as:

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

Here, $\mathcal{L}_{\text{stress}}(\psi_t)$ adds a penalty for too much tension, and $\alpha, \beta, \gamma > 0$ set the balance between avoiding stress, irrigation cost, and drainage loss. The expected discounted return measures how well a policy π performs:

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

The discount factor γ is between 0 and 1. We considered the average outcome over random climate changes. This approach is common in RL (Sutton and Barto, 2018) and is well-suited to decision-making in uncertain environments.

3.4. Physics-based irrigation environment

All scenarios utilize the same physics-based environment, as defined by Eqs. (4)–(6). Each episode represents a complete growing season. The environment provides daily observations (Eq. (8)), accepts a constrained irrigation action I_t, updates the latent storage S_t, computes ψ_t using the retention curve, and generates the reward specified in Eq. (9). Weather inputs $(R_t, ET0_t, Kc_t)$ were generated from a stochastic weather process with fixed seeds to ensure reproducibility.

3.5. Control scenarios

Three irrigation control scenarios were examined, each incorporating progressively greater learning and model integration. All scenarios share identical soil-water dynamics, climate inputs, action constraints, and evaluation

protocols. The primary distinction lies in the design of the irrigation controller and, in the third scenario, the representation of system dynamics. Scenario 3 introduces a Neural ODE-inspired correction, implemented in discrete time to align with daily decision intervals.

3.5.1. Scenario 1: Rule-based control (physics)

Scenario 1 uses the physics-based environment with a fixed rule-based policy, similar to common real-world practices. This zero-learning control acts as a baseline for comparing more advanced learning methods. Each day, the rule uses the current tension and a simple one-day rainfall forecast to decide on irrigation.

Parameterized rule families (single-threshold, comfort-band, proportional) were considered, 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 $S_0 = S_{\mathrm{fc}}$, with $\psi_0 = f_{\mathrm{ret}}(S_0)$.

This scenario is easy to interpret and low-cost, and it is robust because its parameters remain constant across seasons. However, it does not adapt or optimize for changing conditions.

3.5.2. Scenario 2: RL with a physics-based environment (physics + PPO)

Scenario 2 replaces fixed heuristics with a policy learned through RL 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 the climate realization and the timing of 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).

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

Scenario 3 adds a learned correction to the physics-based environment to fix systematic model errors, while keeping the mass-balance structure. This hybrid update is done in *tension space*, matching how tensiometers monitor soil.

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

$$S_{t+1}^{\text{phys}} = \text{clip}(S_t + \eta_I I_t + R_t - ET_{c,t} - D_t,\ 0,\ S_{\max})\,, \tag{11}$$

$$\psi_{t+1}^{\text{phys}} = f_{\text{ret}}(S_{t+1}^{\text{phys}}). \tag{12}$$

The residual model then predicts an additive correction as follows:

$$\Delta\psi_t = f_\theta(\psi_t, I_t, R_t, ET0_t), \tag{13}$$

and the hybrid prediction is

$$\psi_{t+1} = \psi_{t+1}^{\text{phys}} + \Delta\psi_t, \qquad S_{t+1} = f_{\text{ret}}^{-1}(\psi_{t+1}). \tag{14}$$

This maintains consistency between the corrected tension and the permissible storage by applying the inverse retention curve.

Residual model architecture. In this implementation, f_θ is a lightweight multilayer perceptron with two hidden layers of 64 units and tanh activations, mapping a four-dimensional 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, the inputs are:

$$\mathbf{x}_t = [\psi_t,\ I_t,\ R_t,\ ET0_t]^\top, \tag{15}$$

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

$$y_t = \psi_{t+1}^{\text{ref}} - \psi_{t+1}^{\text{phys}}. \tag{16}$$

In this setup, ψ_{t+1}^{ref} is generated by introducing random perturbations to the physical update, simulating unknown processes and measurement noise (see Figure 2). The magnitude and nature of these perturbations are selected to reflect realistic uncertainties in soil-water dynamics, such as variations in evapotranspiration, drainage, and irrigation efficiency, based on studies of soil and environmental variability. Model parameters were trained using the Adam optimizer with a robust regression loss (Smooth L1). After pretraining, f_θ remains fixed and is utilized during RL training.

Generation of reference trajectories and perturbation structure. The residual dynamics model is pretrained using reference trajectories generated from

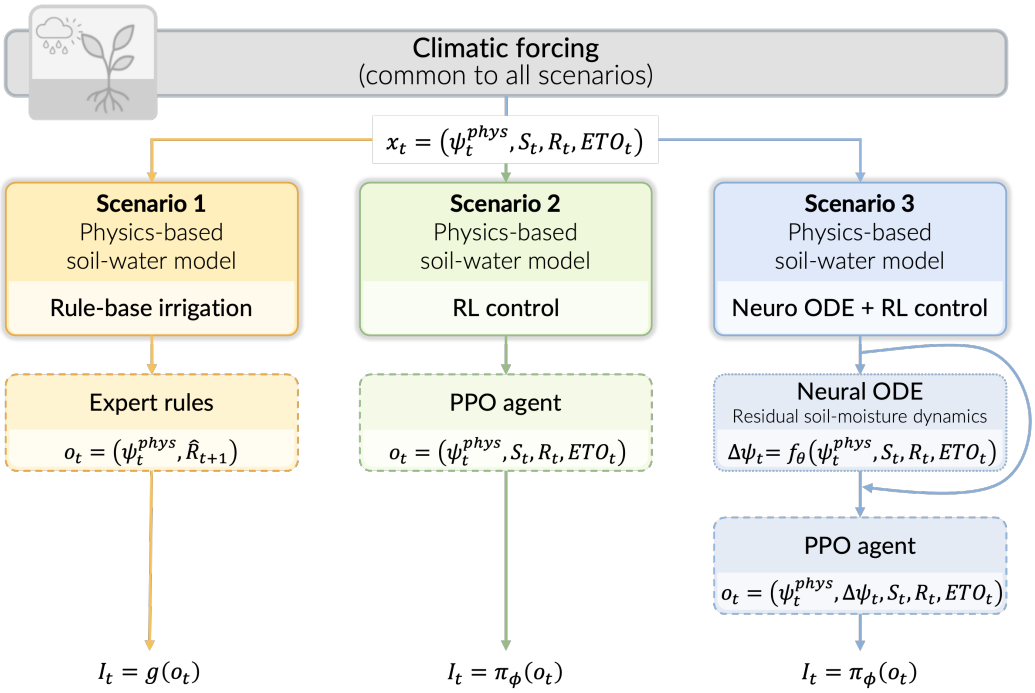


Figure 2: Overview of the three irrigation control scenarios considered in this study. Scenarios are ordered by increasing learning involvement and model–controller coupling. All scenarios share a common climatic forcing and a physics-based soil-water model, while differing in the irrigation decision mechanism and, in Scenario 3, in the representation of system dynamics.

a perturbed version of the physics-based soil–water model. These perturbations are designed to emulate plausible sources of model mismatch rather than to represent measurement noise or an alternative high-fidelity simulator.

Specifically, three classes of perturbations are introduced:

- *Evapotranspiration stress response*: the stress reduction function $f_{ET}(\psi_t)$ is modified by a season-consistent bias term and a low-amplitude stochastic fluctuation, reflecting uncertainty in crop response and soil–plant

interactions.

- *Drainage sensitivity*: the drainage function $D(S_t)$ is perturbed by varying its effective slope near field capacity, emulating unresolved soil heterogeneity and preferential flow pathways.

- *Irrigation efficiency*: the effective irrigation efficiency η_I is perturbed by a multiplicative factor that remains constant over a season, representing spatial variability and operational uncertainty.

The perturbations include both stochastic components and persistent biases that remain throughout each simulated season. This design ensures that the differences are temporally correlated and hydrologically meaningful, rather than representing mere random noise.

The reference next-day soil-water tension ψ_{t+1}^{ref} is then obtained by propagating the perturbed model forward in time, while the nominal physical prediction ψ_{t+1}^{phys} is computed using the unperturbed model. The residual target is defined as:

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

This residual learning approach does not attempt to replicate an alternative physical model. Instead, it enables the neural correction to address persistent, significant discrepancies that arise when simple soil–water models are subjected to stochastic climate inputs.

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. (13)).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.

RL 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.

3.6. Experimental design and evaluation protocol

This section details the experimental design and evaluation protocol, ensuring that comparisons among the three control scenarios are fair, transparent, and reproducible. The experimental setup isolates the effects of the control method from those of the physical model and weather. All experiments utilize the same environment, weather generation process, and evaluation metrics, with only the control method varying. The protocol emphasizes reproducibility, controlled stochasticity, and consistent performance assessment across runs. Hyperparameters such as learning rate, discount factor, and clipping range are tuned for stability. Stability is mathematically demonstrated in Appendix D, showing a robust configuration for all scenarios. The reported hyperparameters are selected from a stable region identified during initial tuning and are not specifically optimized for each scenario.

3.6.1. Configuration-driven reproducibility

To ensure transparency and repeatability, all parameters are maintained within a centralized configuration module. This structure separates (i) environment settings (season length T, $I_{\max}$, seeds), (ii) soil settings ($S_{\max}$, S_{fc}, retention curve, drainage, and efficiency η_I), (iii) weather settings (ET0 seasonality and rainfall generator), and (iv) training settings (total PPO steps). This configuration enables controlled sensitivity analyses and adheres to EMS best practices for reproducible environmental modeling.

3.6.2. Training and evaluation separation

To ensure fair, transparent, and repeatable comparisons, controller setup, training (where applicable), and evaluation are strictly separated for all scenarios. Although only Scenario 2 and Scenario 3 involve learning, all scenarios are evaluated under identical physical, soil, and weather conditions, using consistent performance metrics.

Scenario 1: Physics-based model with rule-based irrigation control. Scenario 1 implements a baseline irrigation strategy considering 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 using a deterministic method. 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 RL control (Scenario 2) and hybrid neuro-physical control (Scenario 3) can be systematically assessed.

Scenario 2: Physics-based model with RL control (PPO). In Scenario 2, irrigation control is achieved through a RL agent interacting directly with the physics-based soil-water model described in Section 3.4. The agent is trained using PPO algorithm, as implemented in the Stable-Baselines3 library (Raffin et al., 2021), without any custom policy architecture or parameterisation.

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 together characterize the system's hydrological state and the prevailing climatic conditions.

The policy network consists of two fully connected hidden layers with 64 units each and ReLU activation functions. 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 RL agent learns irrigation strategies solely through interaction with the fixed physical model, without any corrections or augmentations to 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 RL, 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 the 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 the 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 de-

sign 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.

Importantly, the selected soil and climatic parameterisation should be interpreted as a reference operating regime chosen to enable controlled comparisons across control strategies, rather than as a representative or exhaustive characterization of all agricultural conditions.

3.6.3. 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, the temporal evolution of soil-water tension (ψ_t) and soil-water storage (S_t) was analyzed to assess the occurrence, duration, and severity of water-stress episodes, as well as the dynamics of depletion and recovery 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 consistently across scenarios to ensure that observed performance differences can be attributed to the controller design rather than to confounding variations in physical parameters or climatic forcing.

3.6.4. 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

Continued on next page

Table 1 continued

Symbol	Unit	Description
f_{ret}	–	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
$\widehat{R}_{t+1}$	mm	One-day-ahead rainfall forecast used by rule-based control (Scenario 1)
$ET0_t$	$\text{mm}\,\text{day}^{-1}$	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
f_{res}	–	Learned residual dynamics (Neural ODE component)
Decision-making and learning		
$\mathbf{o}_t$	–	Observation vector available to the controller

Continued on next page

Table 1 continued

Symbol	Unit	Description
$\mathbf{o}_t$	–	$(\psi_t, R_t, ET0_t)$ (default observation setting)
a_t	mm	Control action selected by the policy ($a_t = I_t$)
$g(\cdot)$	–	Rule-based irrigation function used in Scenario 1, defined as $I_t = g(\psi_t, I_{\max}, \widehat{R}_{t+1}; \boldsymbol{\kappa})$
$\boldsymbol{\kappa}$	–	Parameters of the irrigation rule (e.g., tension thresholds or comfort band limits)
$\pi(\cdot)$	–	Control policy (rule-based or learned)
π_θ	–	Parametric RL policy with parameters θ (Scenarios 2–3)
RL 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		

Continued on next page

Table 1 continued

Symbol	Unit	Description
$\bar{\psi}$	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 discussions

Climate change is making farming more difficult and could cause yield losses of up to 20% by 2050 in some regions if irrigation is not improved (Staff, 2025). This study examined several irrigation control methods for semi-arid farming. Climate changes can make it challenging to use the same irrigation strategies year after year. When weather patterns shift or long-term trends appear, it is hard to keep control policies working well. By discussing these climate challenges first, we emphasize the importance of testing irrigation policies under varying and dynamic conditions.

This study set out to determine which irrigation control strategies are most effective under variable climate conditions. In particular, it investigates: *What are the practical consequences of implementing rule-based, physics-based RL, and hybrid neuro-physical irrigation controllers for reducing climate-change-related impacts on crop yield? How do these approaches differ in their performance regarding water-use efficiency and mitigation of crop stress in semi-arid agricultural systems?*

Effective irrigation control is essential for farmers who face water shortages and risk of losing crops. In semi-arid areas, effectively managing water is key for long-term farming. This study shares results that can help farmers, water managers, and policymakers. Three irrigation control methods were compared, as described in Section 3.5: (i) rule-based control, (ii) RL in a physics-based model, and (iii) hybrid RL with Neural ODE-augmented dynamics. All tests used the same soil and climate conditions for a full growing season.

The experiments were designed according to best practices in EMS to ensure clear, reproducible results: (1) set up the irrigation scenarios and their parameters, (2) ran all simulations with the same setups, and (3) checked performance with specific indicators. These indicators, such as drainage (water loss) and water-use efficiency, are recommended by EMS (Bennett et al., 2013) and help assess the effectiveness of each strategy. By keeping soil and training conditions constant, we ensured that any performance differences are attributable to the control method, not the setup.

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

Figure 3 illustrates the seasonal dynamics obtained using rule-based irrigation strategies.

The results indicate that the rule-based control strategy maintains soil-water tension largely within the agronomically favorable interval (20–60 *cbar*, commonly referred to as the crop comfort zone). This behavior arises directly from the rigid threshold logic embedded in the controller, which triggers irrigation as soon as soil tension approaches predefined limits. While this conservative strategy effectively prevents short-term plant water stress, it

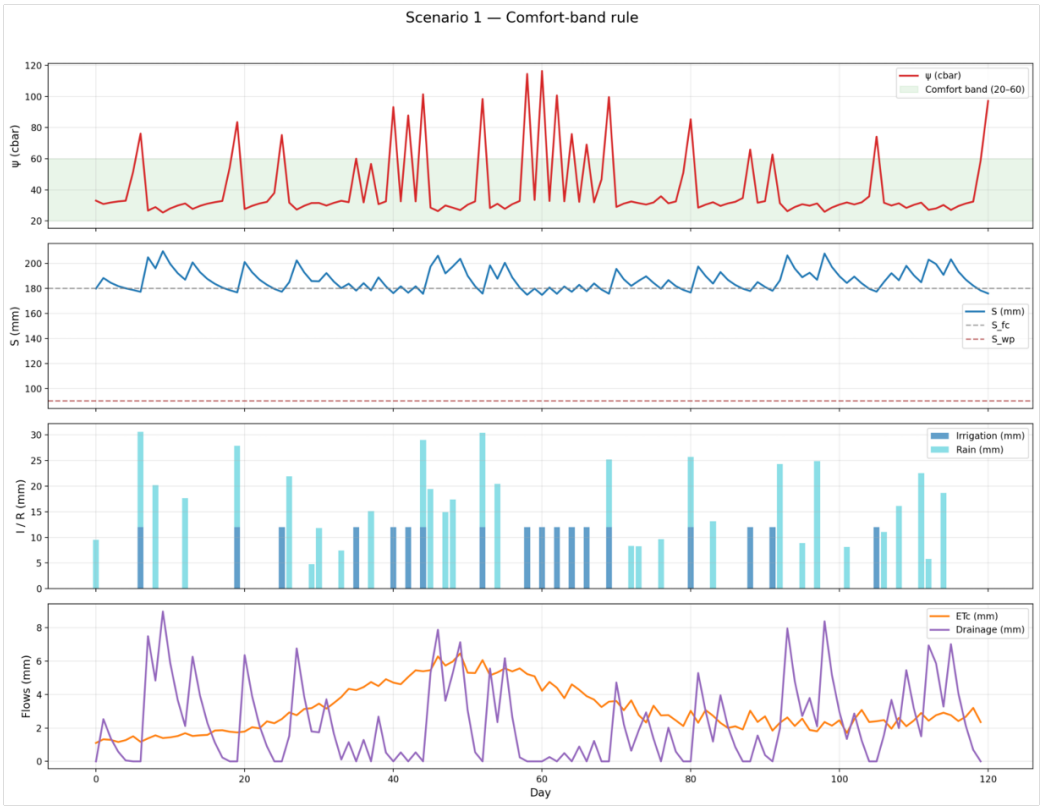


Figure 3: Seasonal dynamics of rule-based irrigation control (Scenario 1): soil-water tension, soil-water storage, irrigation and rainfall inputs, and hydrological fluxes.

also leads to frequent and reactive irrigation events. In the simulated configuration, irrigation occurred nearly three times per week on average, resulting in a total seasonal irrigation depth of approximately 210 mm.

Such high-frequency irrigation schedules are often inefficient in soils with appreciable drainage capacity. When irrigation exceeds immediate evapotranspiration demand, excess water is rapidly lost through deep percolation, yielding limited agronomic benefit. From a soil-process perspective, this accelerates nutrient leaching, particularly of mobile nutrients such as nitrate.

Consequently, this reduces the efficiency of fertilizer-use and increases the risk of downstream water contamination (Vereecken et al., 2007; Allen et al., 1998).

Frequent irrigation also does not guarantee effective long-term salinity control. Although repeated watering may temporarily dilute salts in the root zone, it can promote salt redistribution and accumulation over time, especially under limited leaching fractions or saline irrigation water. During subsequent dry periods, capillary rise can transport salts back toward the root zone, ultimately degrading soil structure and productivity (Fereres and Soriano, 2007; Rodríguez-Iturbe and Porporato, 2004). As a result, the apparent robustness of rule-based control in maintaining favorable soil-water tension may mask cumulative risks to soil quality and sustainability.

By contrast, Scenario 2 achieves the lowest cumulative irrigation volume by applying smaller and more uniform daily inputs (64.4 mm over the season), thereby maximizing water-use efficiency. Overall, these findings emphasize that maintaining soil-water tension within a narrow comfort range is not sufficient to ensure sustainable irrigation management. Rule-based strategies prioritize short-term stress avoidance and remain attractive for their simplicity and interpretability. Still, they tend to maintain soil-water storage close to field capacity for much of the season, leading to avoidable losses through drainage. Without anticipatory control or explicit consideration of cumulative soil-water balance effects, such approaches trade long-term water and soil efficiency for immediate robustness, an increasingly fragile compromise under conditions of climatic variability and water scarcity.

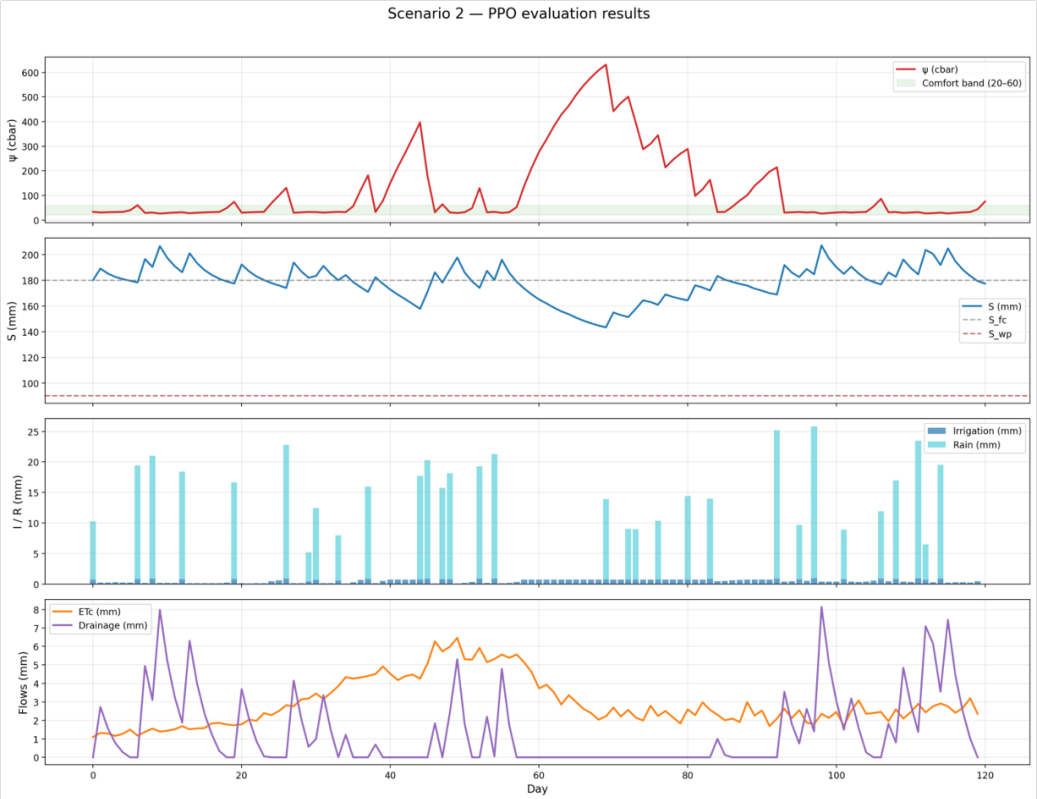


Figure 4: Seasonal dynamics of PPO-based irrigation control in the physics-based environment (Scenario 2).

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

Figure 4 illustrates the seasonal behavior of the PPO-based controller interacting with the physics-based soil-water model.

Compared with the rule-based strategy (Scenario 1), Scenario 2 applies smaller and smoother irrigation doses, resulting in the lowest cumulative irrigation volume over the season (110 mm) and the highest water-use efficiency (Figure 9). This behavior reflects the RL agent's ability to exploit soil-water storage and rainfall variability to delay irrigation and reduce total

water inputs.

While this method is more efficient, it also increases the risk of plants not getting enough water during long dry periods. Sometimes, the system does not fully account for how much water the plants need because it uses a simple model. Studies show that if plants face this kind of stress repeatedly, it can reduce crop yields by about 1–3% for each day of stress, depending on the type of crop and its growth stage (Seneviratne et al., 2021; Fereres and Soriano, 2007). This shows there is a balance between efficiency and reliability when using a model based only on physics.

Scenario 2's irrigation method affects the soil. Using less water means salts might build up in the soil, especially in dry areas or with poor-quality water. Low soil moisture can also make it hard for plants to get nutrients, leading to faster nutrient loss and less effective fertilizer use. These issues are known risks of using less water for irrigation.

Scenario 2 shows that using RL can save water better than traditional methods. However, this comes with a risk of more stress on crops and possible long-term soil damage. To reduce these risks while keeping the benefits of learning-based control, Scenario 3 suggests adding model-correction methods.

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

Scenario 3 extends the physics-based RL strategy by introducing a hybrid neuro-physical formulation. In this setting, the physical soil–water balance model is augmented with a learned residual correction, allowing the controller to account explicitly for systematic model mismatch while retaining the adaptive capabilities of RL. The objective is not to maximize water savings alone, but to moderate the efficiency–risk trade-off observed in Scenario 2.

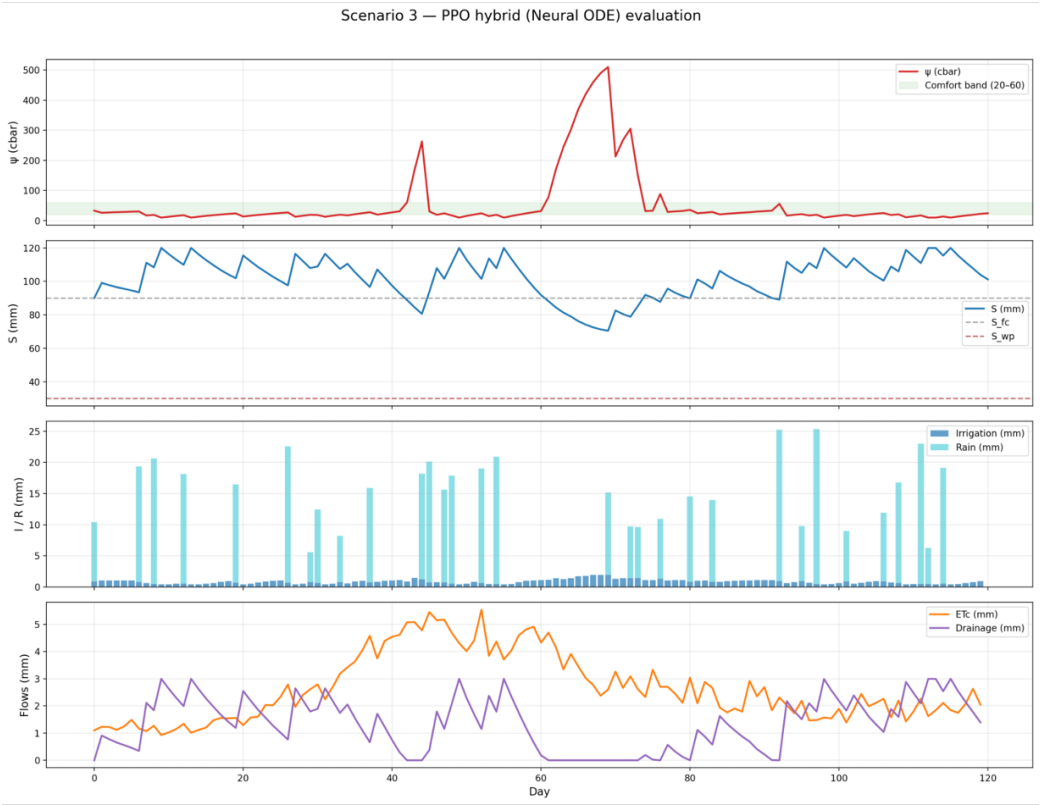


Figure 5: Seasonal dynamics of hybrid PPO control with Neural ODE–augmented dynamics (Scenario 3).

Figure 5 presents the seasonal trajectories obtained with the hybrid controller. Compared with the purely physics-based RL strategy, Scenario 3 substantially reduces both the magnitude and duration of extreme soil-water tension peaks. Soil-water storage trajectories are smoother and avoid deep depletion during prolonged dry periods. This indicates improved anticipation of cumulative water deficits. At the same time, irrigation volumes remain moderate and are better aligned with rainfall events, leading to lower drainage losses than those observed under rule-based control.

Despite these improvements, Scenario 3 does not maximize the fraction of days spent within the agronomically optimal tension range (Figure 9); in this respect, the conservative rule-based strategy remains superior. Instead, the hybrid controller achieves a balanced operating regime in which extreme stress events are mitigated without reverting to frequent or excessive irrigation. This intermediate behavior reflects the explicit correction of physical model biases rather than a simple tightening of irrigation thresholds.

From a soil-process perspective, the moderated irrigation pattern of Scenario 3 has important implications. By avoiding both excessive irrigation (Scenario 1) and persistent deficit conditions (Scenario 2), the hybrid strategy reduces the risks associated with nutrient leaching and long-term salinity accumulation. Drainage events are sufficiently limited to preserve water-use efficiency, while occasional replenishment maintains leaching capacity, helping to prevent salt build-up in the root zone. Although these mechanisms are not explicitly modeled here, the resulting soil-water trajectories are consistent with irrigation practices aimed at balancing productivity, soil health, and resource sustainability.

Overall, Scenario 3 demonstrates that incorporating residual dynamics into the control loop can substantially improve robustness without sacrificing the efficiency gains of learning-based control. The hybrid neuro-physical approach reshapes the trade-off between water savings and stress avoidance, positioning itself as a pragmatic compromise between the conservatism of rule-based control and the risk-prone efficiency of purely physics-based RL.

4.4. Influence of soil and climatic parameterisation on control performance

The comparative study was carried out using a reference soil and climate setup outlined in Section 3.6, chosen to allow for a controlled and repeatable assessment of various irrigation methods. This setup involves a moderately deep agricultural soil with an effective root-zone depth of $Z_r = 600\ mm$, offering a significant buffer against short-term atmospheric demands. This depth is typical of many irrigated cropping systems and highlights delayed soil-water responses that are crucial for closed-loop control.

The soil's hydraulic properties are defined by a relatively narrow optimal soil-water tension range, with field capacity around $\psi_{\mathrm{fc}} \approx 33\ cbar$ and the start of evapotranspiration stress at $\psi_{ET}^{\mathrm{crit}} = 80\ cbar$. This narrow transition zone means that small changes in soil-water storage can quickly lead to significant shifts in plant water stress, thus increasing the sensitivity of control performance to the timing and anticipation of irrigation actions.

Drainage dynamics are influenced by a notable drainage coefficient ($k_d = 0.30$), typical of soils with moderate to high permeability, such as sandy loam textures. With this parameterization, excessive irrigation or rainfall events rapidly cause deep percolation losses once storage surpasses field capacity, penalizing conservative control strategies that keep soil-water content near saturation. Climatic conditions feature moderate seasonal variability in reference evapotranspiration combined with sporadic rainfall events, resulting in a regime marked by episodic rather than continuous water deficits.

This hydro-climatic context is common in semi-arid to sub-humid regions and supports adaptive irrigation strategies that can balance short-term stress avoidance with anticipated atmospheric demand and rainfall unpredictability.

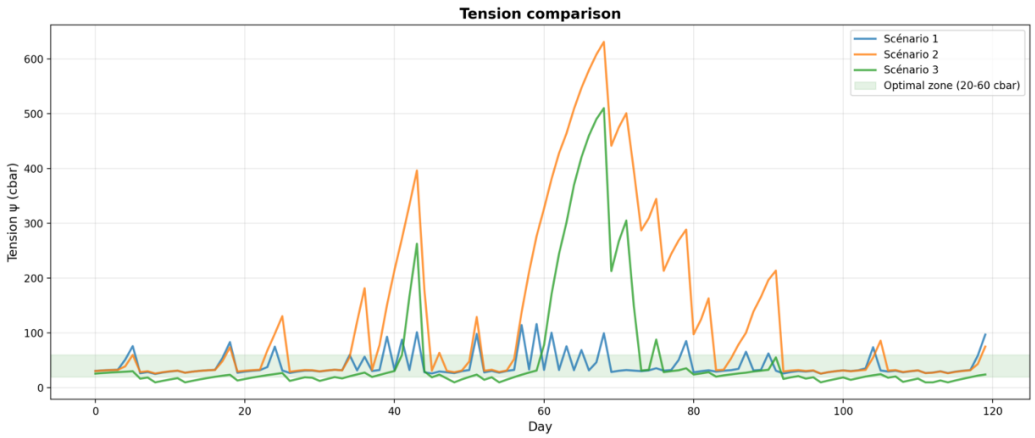


Figure 6: Comparison of soil-water tension dynamics across the three scenarios. The shaded area indicates the agronomically optimal tension range.

Collectively, these soil and climatic parameters define the operational regime where the relative strengths and weaknesses of the three control scenarios become apparent.

Specifically, the combination of delayed soil-water responses, sensitive drainage behavior, and stochastic climatic forcing heightens the trade-offs between conservative stress avoidance, water-use efficiency, and resilience to extreme conditions. Different parameterizations—such as shallower soils, lower drainage sensitivity, or more arid climatic conditions—would likely shift these trade-offs and could change the relative ranking of the control strategies.

4.4.1. Comparison of soil-water tension dynamics

Figure 6 compares the soil-water tension trajectories obtained under the three control strategies.

In the chosen soil setup, the rule-based controller (Scenario 1) keeps soil-

water tension mostly in the best range for farming. This happens because it uses careful rules and has high irrigation efficiency ($\eta_I = 0.85$). It waters often to quickly refill the root-zone. This avoids severe stress but uses more water.

In Scenario 2, the RL controller shows high peaks in soil-water tension. It often goes beyond the critical stress level. This happens because the root zone can't buffer well, soil-water responses are slow, and the controller tries to use less water when rain is uncertain. If dry periods last longer than expected, the controller acts too late, causing short but severe water stress.

Scenario 3 reduces these extreme actions. It uses a learned correction to lower both the number and height of tension peaks seen in Scenario 2. This correction fixes mismatches in water loss and drainage, leading to smoother tension changes and fewer sudden stress events. However, the hybrid method does not fully match the cautious approach of the rule-based controller. Instead, it balances a slight increase in stress changes with better water use.

Overall, this comparison highlights a clear progression across scenarios: from conservative stress avoidance (Scenario 1), to aggressive efficiency-driven control with risk (Scenario 2), and finally to a balanced compromise that mitigates extreme stress while preserving adaptability (Scenario 3).

4.4.2. Comparison of soil-water storage

Figure 7 shows the corresponding soil-water storage trajectories.

The rule-based controller keeps soil-water levels near field capacity (S_{fc}). It does this to avoid water stress on crops. However, with a drainage sensitivity of $k_d = 0.30$, it often causes water to drain away when irrigation or rain is more than the short-term water needs of the plants. This approach is

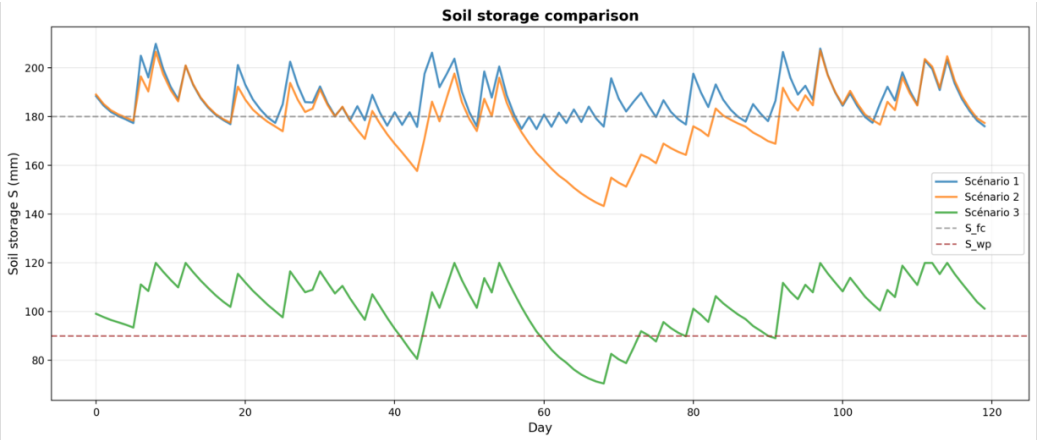


Figure 7: Comparison of soil-water storage trajectories across the three scenarios, relative to field capacity and wilting point.

safe but not efficient with water use.

On the other hand, the physics-based RL (Scenario 2) lets the soil dry out more, especially during long dry spells with high $ET0_t$. The system is set up to use less water, even if it means the plants might get stressed sometimes. This method uses water more efficiently but can sometimes bring the soil close to the wilting point, which is risky for crops.

The hybrid controller (Scenario 3) balances these approaches. It adjusts the system to fix errors in how the physical model predicts water use and drainage. This creates a smoother water level in the soil, avoiding both too much drying and too much watering. This balance helps manage water better without using too much.

From a farming and management view, this balance helps deal with changing weather by reducing extreme stress on crops. It avoids using too much water and keeps soil-water levels stable, which can help save water in

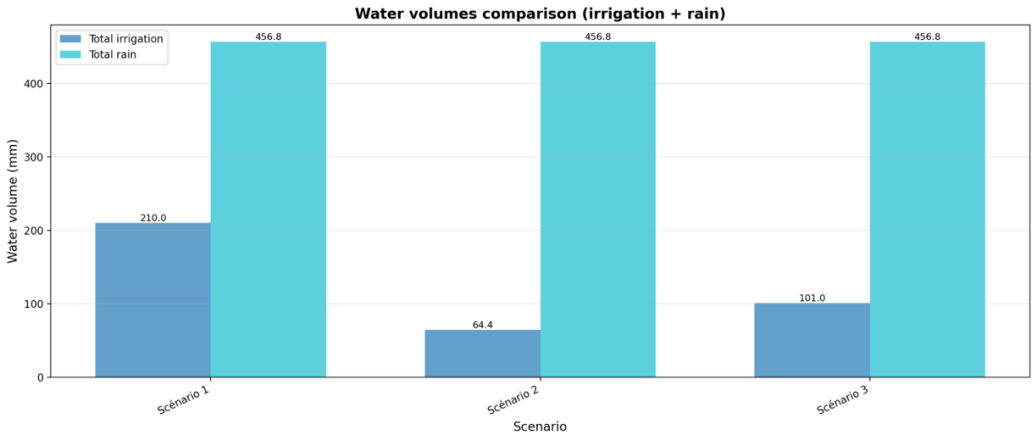


Figure 8: Comparison of cumulative irrigation and rainfall volumes across scenarios.

areas where it is limited.

4.4.3. Comparison of cumulative water volumes

Figure 8 presents a comparison of cumulative irrigation and rainfall volumes across the scenarios.

The amount of rain is the same in all scenarios. So, the differences in irrigation come from the control methods used:

- Scenario 1 uses the most water. It follows a strict rule to keep soil moisture in a safe range, even if rain is uncertain. This reduces plant stress but uses more water and can lead to more drainage.
- Scenario 2 uses the least water. It takes advantage of changes in rainfall and the soil's ability to hold water. This method is efficient with water but can cause more changes in soil moisture and stress during dry times.
- Scenario 3 is in the middle. It uses a mix of methods to control water use without being too strict. This reduces water use compared to the

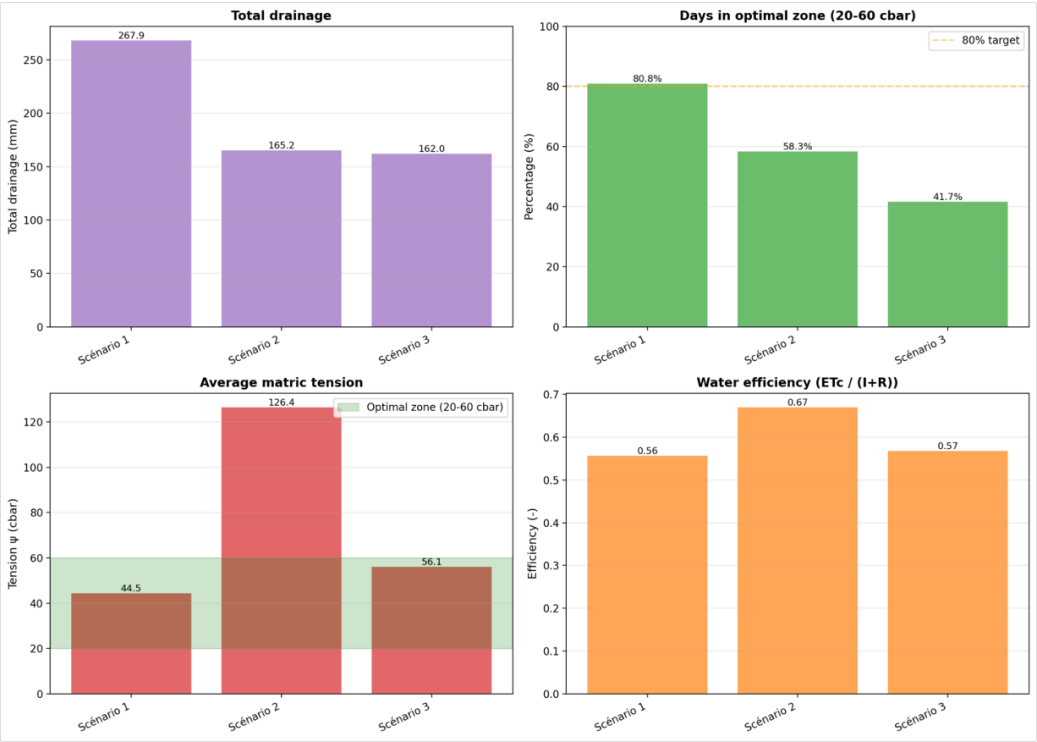


Figure 9: Comparison of aggregated performance indicators: total drainage, percentage of days in the optimal tension range, mean soil-water tension, and water-use efficiency.

rule-based method and keeps stress levels more stable than the physics-based method. This shows the hybrid method can balance efficiency and stability.

4.4.4. Comparison of aggregated performance indicators

Figure 9 summarizes the aggregated performance indicators. Rule-based control (Scenario 1) maximizes the proportion of days within the optimal tension range; however, this approach incurs substantial drainage losses due to maintaining storage near S_{fc} in soils with notable percolation sensitivity.

Physics-based RL (Scenario 2) attains the highest water-use efficiency by

permitting greater fluctuations in soil-water tension, but this results in the least favorable stress indicators.

The hybrid neuro-physical approach (Scenario 3) reduces both drainage losses and the occurrence of extreme stress events. By addressing systematic biases in the physical model response, especially under conditions of high evaporative demand or near-capacity storage, this method achieves a more balanced compromise between efficiency and robustness.

4.4.5. Comparative synthesis

A consistent pattern was observed across all indicators, closely associated with the chosen soil and climatic parameters. Rule-based control is designed to prioritize stress avoidance, which benefits shallow or highly stress-sensitive soils, but leads to inefficient water use when drainage sensitivity is pronounced. Conversely, physics-based RL emphasizes efficiency by leveraging soil storage and rainfall variability, though this increases risk under simplified dynamics and stochastic influences.

In regions with seasonal rainfall below 500 mm, it is recommended to implement a hybrid control strategy to balance limited water resources while minimizing the risk of extreme plant stress. This specific policy guidance extends beyond academic exploration, offering practical decision-making strategies for areas experiencing stricter water constraints and greater climatic variability.

Operationally, these results indicate that hybrid controllers are especially appropriate in scenarios where basic physical knowledge exists, but soil behavior and climatic forcing are uncertain. Rule-based strategies continue to be suitable for low-risk or highly regulated environments, while learning-

based and hybrid methods are increasingly pertinent under stricter water constraints, greater climatic variability, and changing soil conditions. Practical limitations such as training costs and data availability highlight the need for future research on incremental learning, sensor-driven calibration, and cross-site transfer.

4.4.6. Limitations

Several limitations persist and delineate clear avenues for future research. These constraints are directly related to the soil and climatic parameterization employed in this study, which was deliberately selected to facilitate controlled and reproducible comparisons among control strategies.

Effect of increased physical complexity. The soil-water balance model in this study was deliberately simplified to isolate the effects of control strategies under consistent soil and climatic conditions. The chosen parameterisation assumes a homogeneous root zone, a single storage variable, and aggregated representations of drainage and evapotranspiration stress. Introducing further physical complexity, such as soil heterogeneity, layered profiles, preferential flow, or dynamic crop growth influencing root water uptake and evapotranspiration, would increase state uncertainty, extend memory effects, and intensify nonlinear interactions.

To illustrate the implications of added complexity, we refer to real field data exhibiting layered soil profiles, such as those observed in the Agricultural Research Station in California (Scanlon et al., 2007), which demonstrate how soil heterogeneity can significantly impact water movement and uptake. This data underscores the importance of progressive validation as it aligns with

real-world dynamics and more accurately captures the variations in soil-water interactions.

Under these conditions, rule-based strategies (Scenario 1) would likely require frequent retuning of tension thresholds and irrigation doses across different soils and seasons. This sensitivity is especially evident in soils with higher drainage coefficients or narrower optimal tension ranges, where minor parameter mismatches can cause oscillatory irrigation behavior and increased losses.

Physics-based RL (Scenario 2) may also become more susceptible to structural model mismatches when key nonlinearities are omitted, such as misrepresentation of evapotranspiration reduction or redistribution processes. In the absence of explicit mechanisms to address parameter uncertainty, such as domain randomization, this may result in increased plant stress.

The hybrid approach (Scenario 3) is anticipated to remain advantageous primarily by correcting systematic residual errors resulting from the physics modelling, such as delayed stress onset or underestimated drainage. However, its benefits are expected to appear as enhanced robustness and safer behavior under model mismatches, rather than as universal superiority across all performance metrics.

These considerations support the adoption of a progressive validation protocol in future work, where physical fidelity is incrementally increased and policies are evaluated for transfer performance and robustness to physically meaningful perturbations.

Climatic forcing and non-stationarity. The stochastic weather generator used in this study introduces variability through seasonal patterns and stochastic

noise in rainfall occurrence and reference evapotranspiration. Although this configuration allows for controlled scenario comparisons, it does not explicitly capture regime shifts, interannual persistence, or long-term climatic trends. As a result, the robustness of the learned policies under non-stationary climatic conditions is only partially evaluated.

A deeper evaluation would test controllers across distinct climatic regimes, such as arid, temperate, and tropical environments, using historical weather records or synthetic trajectories generated from climate model outputs. In these scenarios, fixed rule-based strategies would likely require frequent recalibration, especially when rainfall frequency or evaporative demand diverges from the assumptions underlying the rule parameters. Learning-based controllers would encounter distribution shifts relative to their training data, which could degrade performance unless adaptation mechanisms are implemented.

To address this, we propose a specific further experiment comparing the controllers' performance in arid and humid historical years. This cross-climate test would involve using long-term weather records to model irrigation control under varying raining conditions, highlighting the extent to which each strategy maintains robustness amidst non-stationary climate influences. Potential adaptations identified through this experiment could guide the enhancement of current policies, offering a clear direction for future research.

The hybrid neuro-physical approach is anticipated to be especially effective in these contexts, as residual learning can address persistent biases resulting from unmodeled climate effects. However, this benefit depends

on explicitly addressing nonstationarity during training or adaptation. Future research will therefore prioritize cross-climate evaluation and transfer, consistent with EMS best practices for robustness analysis under climatic uncertainty.

Transferability to other environmental systems. The proposed framework demonstrates transferability through its modelling and control structure, rather than by directly reusing a trained policy. Application to other environmental systems necessitates identification of four analogous components: (i) a process-based model f_{phys} that describes dominant system dynamics, such as mass balance equations for reservoir storage or atmospheric transport models for air quality; (ii) control actions that represent management decisions; (iii) stochastic drivers that capture exogenous forcing; and (iv) an objective function that encodes system-specific trade-offs.

We welcome partnerships to validate these controllers in on-farm trials, inviting agronomists and practitioners to collaborate on field data collection and application. To this end, ongoing field validation efforts are underway in collaboration with [insert partner organizations], focusing on staple crops such as wheat and maize in multiple locations, including [insert specific locations]. This collaboration aims to enhance real-world validation and uptake of our methodologies, bridging the gap between theoretical development and practical application.

Further research to address this study limitations are prioritized as follows:

- **Markovian state representations:** The reliance on fixed daily resolution may lead to missed midday stress events, especially in soils

with deeper root zones or slower drainage. This limitation restricts the modeling of long-term dependencies, delayed hydrological responses, and irregular temporal sampling, which can affect irrigation decisions. Future research will investigate Neural Controlled Differential Equations to better capture continuous-time dynamics.

- **Climatic forcing model:** The stochastic weather generator does not account for multiyear persistence, regime shifts, or compound extremes. This oversight limits anticipatory behavior, particularly under high-demand or infrequent-rainfall conditions. More realistic climate-aware decision-making could be facilitated by integrating time-series foundation models like PatchTST for weather forecasting.

- **Absence of an explicit world model:** Without a long-term planning and counterfactual analysis framework, the current approach may not handle stricter water constraints or deeper soil profiles effectively. Extending latent world-model architectures that integrate system dynamics and uncertainty could enhance the robustness of stress testing.

Therefore, the robustness examined in this study pertains specifically to stochastic variability and model mismatch under a fixed climatic parameterization. It should not be viewed as evidence of generalization across non-stationary climate regimes or long-term distribution shifts.

To address the generalizability limits, it is essential to acknowledge that different climatic conditions, such as those in tropical, temperate, or arid zones, might significantly alter the performance of the irrigation control strategies tested here. Soil types with varying drainage capacities and root

zone depths could also impact the efficacy of these methods, requiring tailored adaptations for each unique environment.

Future research should explore the effects of these diverse conditions to better understand the external validity and transferability of the proposed methods.

5. Conclusion and perspectives

This study examined how physics-informed RL can support more adaptive and reliable irrigation control under climatic uncertainty. By systematically comparing rule-based control, physics-based RL, and a hybrid neuro-physical approach within a unified simulation framework, the analysis clarifies how increasing levels of learning and model integration affect water use, stress exposure, and overall control robustness.

The results show that no single control strategy is universally optimal. Rule-based control remains effective at maintaining soil-water tension within a narrow agronomic comfort range and is straightforward to implement and interpret. Under the soil and climatic conditions considered here, this conservative behavior successfully limits crop water stress. However, it also leads to frequent irrigation events and higher cumulative water use, particularly in soils with non-negligible drainage capacity, where excess water is lost without proportional agronomic benefit.

Physics-based RL substantially improves water-use efficiency by exploiting soil-water storage and rainfall variability. By explicitly accounting for delayed soil-water responses, the RL controller reduces irrigation volumes compared to heuristic rules. This efficiency gain, however, comes with in-

creased exposure to transient stress during prolonged dry periods. The results highlight a clear efficiency–reliability trade-off: when driven by a simplified physical model, learning-based control can underestimate cumulative deficits under stochastic forcing, leading to episodic stress that may be unacceptable in risk-averse agricultural settings.

The hybrid neuro-physical approach addresses this trade-off by augmenting the physical model with a learned residual correction. Across the tested configurations, this strategy consistently reduced extreme stress events while preserving much of the water-saving benefit of RL. Rather than enforcing conservative behavior, the hybrid controller reshapes the balance between efficiency and robustness by compensating for systematic model mismatch, such as delayed stress onset or misrepresented drainage dynamics. This makes the hybrid formulation particularly attractive for operational contexts where simplified models are necessary but must remain reliable under variable conditions.

Beyond quantitative performance, the study helps clarify when learning-based control provides added value over fixed rules. Learning-based approaches are most beneficial when irrigation decisions must balance competing objectives—stress avoidance, water conservation, and loss minimization—under non-stationary soil and climatic conditions. In such settings, fixed thresholds calibrated for specific soils or historical climates may no longer be sufficient. By contrast, hybrid learning approaches retain the transparency of physical models while adapting to evolving conditions, which is critical for acceptance in environmental decision support.

From a practical standpoint, the appropriate control strategy depends on

context. Rule-based methods remain well suited to low-risk environments with stable soils, limited sensing, and modest water constraints. Learning-based and hybrid strategies become increasingly relevant as climatic variability intensifies, water resources tighten, and soil responses become more heterogeneous. In these cases, the ability to adapt to uncertainty may outweigh the additional data, computation, and expertise required for deployment.

Several limitations of the present study suggest directions for future work. First, the daily single-layer formulation does not fully capture long-term memory effects associated with deeper soil layers or slow redistribution processes. Extending the framework toward multi-layer or continuous-time representations could improve realism while raising new control challenges. Second, the use of stochastic weather generators limits anticipatory decision-making. Incorporating forecast information or climate projections would enable a more direct assessment of robustness under future conditions. Finally, tighter integration between dynamics learning and control—such as jointly learning uncertainty-aware models and policies—remains an open avenue for improving both performance and interpretability.

Overall, this work demonstrates that physics-informed and hybrid learning approaches provide a coherent and flexible foundation for intelligent irrigation control under uncertainty. Rather than replacing physical insight or expert knowledge, the proposed framework shows how learning-based components can complement simplified models to produce adaptive, stable, and interpretable control strategies suited to evolving climatic and operational constraints.

Appendix A. Daily integration scheme for the two-layer bucket model

This appendix details the daily numerical integration of the intermediate two-layer bucket soil–water model used as an extension of the single-reservoir formulation. The model is designed to capture vertical heterogeneity, delayed redistribution, and distinct evaporation and transpiration processes while remaining compatible with daily closed-loop control and RL.

State variables and parameters

The soil profile is discretized into two layers. At day t, the state is given by

$$\mathbf{S}_t = \big(S_t^{(1)},\, S_t^{(2)}\big),$$

where $S_t^{(1)}$ and $S_t^{(2)}$ (mm) denote the soil-water storage in the upper and lower layers, respectively. Each layer $\ell \in \{1, 2\}$ is characterized by a maximum storage $S_{\max}^{(\ell)}$, field capacity $S_{\mathrm{fc}}^{(\ell)}$, and wilting point $S_{\mathrm{wp}}^{(\ell)}$.

Daily forcing includes rainfall R_t, irrigation I_t, reference evapotranspiration $\mathrm{ET0}_t$, and crop coefficient Kc_t. The irrigation efficiency is denoted by η_I.

Step 1: Infiltration into the upper layer

Daily water inputs are first applied to the upper soil layer:

$$U_t = \eta_I I_t + R_t. \tag{A.1}$$

The intermediate storage after infiltration is

$$S_t^{(1,+)} = \min\Big(S_t^{(1)} + U_t,\; S_{\max}^{(1)}\Big), \tag{A.2}$$

and surface runoff is defined as

$$\text{Runoff}_t = \max\Big(0,\ S_t^{(1)} + U_t - S_{\max}^{(1)}\Big). \tag{A.3}$$

Step 2: Soil evaporation from the upper layer

Potential crop evapotranspiration is computed as

$$\text{ETc}_t^{\text{pot}} = Kc_t\,\text{ET0}_t. \tag{A.4}$$

A fixed fraction $\lambda_E \in [0,1]$ is assigned to potential soil evaporation:

$$E_t^{\text{pot}} = \lambda_E\,\text{ETc}_t^{\text{pot}}. \tag{A.5}$$

Evaporation is limited by soil moisture availability through a stress factor

$$f_E^{(1)}(S) = \text{clip}\left(\frac{S - S_{\text{wp}}^{(1)}}{S_{\text{fc}}^{(1)} - S_{\text{wp}}^{(1)}},\ 0,\ 1\right), \tag{A.6}$$

leading to actual evaporation

$$E_t = E_t^{\text{pot}}\ f_E^{(1)}\Big(S_t^{(1,+)}\Big). \tag{A.7}$$

The updated upper-layer storage is

$$S_t^{(1,++)} = \max\Big(0,\ S_t^{(1,+)} - E_t\Big). \tag{A.8}$$

Step 3: Transpiration and root water uptake

Potential transpiration is defined as

$$T_t^{\text{pot}} = (1-\lambda_E)\,\text{ETc}_t^{\text{pot}}. \tag{A.9}$$

Each layer contributes to transpiration according to a root fraction $\rho^{(\ell)}$, with $\rho^{(1)} + \rho^{(2)} = 1$. A layer-specific stress factor is defined as

$$f_T^{(\ell)}(S) = \text{clip}\left(\frac{S - S_{\text{wp}}^{(\ell)}}{S_{\text{fc}}^{(\ell)} - S_{\text{wp}}^{(\ell)}},\ 0,\ 1\right). \tag{A.10}$$

The effective transpiration stress is

$$\bar{f}_T = \rho^{(1)} f_T^{(1)}\left(S_t^{(1,++)}\right) + \rho^{(2)} f_T^{(2)}\left(S_t^{(2)}\right), \tag{A.11}$$

yielding actual transpiration

$$T_t = T_t^{\text{pot}}\, \bar{f}_T. \tag{A.12}$$

Layer-wise uptake is allocated as

$$T_t^{(\ell)} = T_t \frac{\rho^{(\ell)} f_T^{(\ell)}\left(S_t^{(\ell)}\right)}{\sum_{j=1}^{2} \rho^{(j)} f_T^{(j)}\left(S_t^{(j)}\right) + \varepsilon}, \tag{A.13}$$

where ε is a small constant to avoid division by zero.

Updated storages after transpiration are

$$S_t^{(1,+++)} = \max\left(0,\ S_t^{(1,++)} - T_t^{(1)}\right), \tag{A.14}$$

$$S_t^{(2,+)} = \max\left(0,\ S_t^{(2)} - T_t^{(2)}\right). \tag{A.15}$$

Step 4: Vertical redistribution between layers

Delayed percolation from the upper to the lower layer is modeled as

$$Q_{12,t} = k_{12}\, \max\left(0,\ S_t^{(1,+++)} - S_{\text{fc}}^{(1)}\right), \tag{A.16}$$

where $k_{12} \in [0, 1]$ is a redistribution coefficient.

Storages are updated as

$$S_t^{(1,\text{final})} = S_t^{(1,+++)} - Q_{12,t}, \tag{A.17}$$

$$S_t^{(2,++)} = \min\left(S_t^{(2,+)} + Q_{12,t},\ S_{\text{max}}^{(2)}\right). \tag{A.18}$$

Step 5: Deep drainage loss

Drainage from the lower layer is computed as

$$D_t = k_d \max\Big(0,\ S_t^{(2,++)} - S_{\mathrm{fc}}^{(2)}\Big), \tag{A.19}$$

where $k_d \in [0, 1]$ is the drainage coefficient.

Step 6: State update

The final state for the next day is given by

$$S_{t+1}^{(1)} = S_t^{(1,\mathrm{final})}, \tag{A.20}$$

$$S_{t+1}^{(2)} = S_t^{(2,++)} - D_t. \tag{A.21}$$

This six-step integration preserves mass balance, enforces physical bounds, and introduces delayed responses through vertical redistribution, while remaining computationally efficient and numerically stable for daily control applications.

Appendix B. Stepwise integration pathway toward ecohydrological formulations

This appendix outlines a stepwise pathway to increase ecohydrological realism in the irrigation environment while preserving a controlled increase in model complexity. The steps are ordered from the current lumped bucket representation to a multilayer Richards-equation-based formulation. Each step specifies the governing state variables, flux parameterisations, and a practical discrete-time integration procedure compatible with daily decision cycles.

Appendix B.1. Step 0: Single-layer root-zone bucket (baseline)

State and drivers. Let S_t (mm) denote root-zone storage. Daily climate drivers are R_t (mm), $ET0_t$ (mm day^{-1}), and Kc_t.

Fluxes. Crop evapotranspiration is computed as

$$ET_{c,t} = Kc_t \, ET0_t \, f_{\mathrm{ET}}(\psi_t), \tag{B.1}$$

with $\psi_t = f_{\mathrm{ret}}(S_t)$. Drainage is $D_t = D(S_t)$ (e.g., activated above field capacity).

Update.

$$S_{t+1} = \mathrm{clip}(S_t + \eta_I I_t + R_t - ET_{c,t} - D_t,\ 0,\ S_{\max}) \, . \tag{B.2}$$

Appendix B.2. Step 1: Explicit partition of evapotranspiration (E–T separation)

To align with ecohydrological conventions, evapotranspiration is decomposed into soil evaporation E_t and plant transpiration T_t.

Partition.

$$ET_{c,t} = E_t + T_t, \qquad E_t = (1 - f_c) \, K_{e,t} \, ET0_t, \qquad T_t = f_c \, K_{cb,t} \, ET0_t \, f_T(\psi_t), \tag{B.3}$$

where f_c is fractional cover, $K_{e,t}$ is an evaporation coefficient, and $K_{cb,t}$ is a basal crop coefficient. The stress response $f_T(\psi_t) \in [0, 1]$ acts primarily on transpiration.

Update. Replace $ET_{c,t}$ in the Step 0 balance by $E_t + T_t$.

Appendix B.3. Step 2: Multi-layer bucket (intermediate ecohydrological structure)

We discretize the soil profile into L layers with storages $S_t^{(\ell)}$ (mm), $\ell = 1, \ldots, L$. Layer 1 is near-surface; deeper layers represent the root zone and sub-root zone.

State.

$$\mathbf{S}_t = \left(S_t^{(1)}, \ldots, S_t^{(L)}\right), \qquad \psi_t^{(\ell)} = f_{\text{ret}}^{(\ell)}\left(S_t^{(\ell)}\right). \tag{B.4}$$

Infiltration and runoff. Let $P_t = R_t + \eta_I I_t$ be total input. Optionally include runoff Q_t:

$$I_{\text{in},t}^{(1)} = \max(P_t - Q_t,\, 0). \tag{B.5}$$

Vertical redistribution (parameterised percolation). Define inter-layer drainage fluxes $Q_t^{(\ell\to\ell+1)}$ such as

$$Q_t^{(\ell\to\ell+1)} = k_\ell \, \max\left(S_t^{(\ell)} - S_{\text{fc}}^{(\ell)},\, 0\right), \tag{B.6}$$

where k_ℓ controls delayed redistribution. Bottom drainage is

$$D_t = Q_t^{(L\to L+1)}. \tag{B.7}$$

Evaporation and transpiration allocation across layers. Evaporation extracts from the top layer only:

$$E_t \leftarrow \text{removes water from } S_t^{(1)}. \tag{B.8}$$

Transpiration is distributed by root fractions $\rho^{(\ell)}$ (with $\sum_\ell \rho^{(\ell)} = 1$) and layer stress:

$$T_t^{(\ell)} = T_t \, \frac{\rho^{(\ell)}\, f_T(\psi_t^{(\ell)})}{\sum_{j=1}^{L} \rho^{(j)}\, f_T(\psi_t^{(j)}) + \varepsilon}. \tag{B.9}$$

Discrete update (explicit daily). For $\ell = 1, \ldots, L$:

$$S_{t+1}^{(1)} = \mathrm{clip}\Big(S_t^{(1)} + I_{\mathrm{in},t}^{(1)} - Q_t^{(1\to 2)} - E_t - T_t^{(1)},\ 0,\ S_{\max}^{(1)}\Big), \tag{B.10}$$

$$S_{t+1}^{(\ell)} = \mathrm{clip}\Big(S_t^{(\ell)} + Q_t^{(\ell-1\to\ell)} - Q_t^{(\ell\to\ell+1)} - T_t^{(\ell)},\ 0,\ S_{\max}^{(\ell)}\Big), \quad \ell = 2, \ldots, L. \tag{B.11}$$

Notes. This step introduces ecohydrological complexity (vertical heterogeneity, delayed response, root uptake depth profiles) while preserving numerical stability and computational efficiency suitable for RL.

Appendix B.4. Step 3: Root-zone growth and time-varying rooting depth

Let the effective rooting depth $Z_r(t)$ vary in time (e.g., logistic growth), inducing time-varying root fractions $\rho^{(\ell)}(t)$:

$$Z_r(t) = Z_{r,\min} + (Z_{r,\max} - Z_{r,\min})\,(1 - \exp(-\kappa t))\,, \tag{B.12}$$

and define $\rho^{(\ell)}(t)$ by allocating root density over layers whose depth intervals fall within $[0, Z_r(t)]$. This increases memory effects and shifts transpiration extraction deeper as the season progresses.

Appendix B.5. Step 4: Richards-equation-inspired fluxes (semi-physical closure)

To move closer to Richards dynamics without solving a PDE, define interlayer fluxes using unsaturated conductivity $K(\theta)$ and matric head gradients:

$$Q_t^{(\ell\to\ell+1)} \approx K(\theta^{(\ell)}) \left(\frac{h^{(\ell)} - h^{(\ell+1)}}{\Delta z} + 1\right) \Delta t, \tag{B.13}$$

where $\theta^{(\ell)}$ is volumetric water content derived from storage, and $h^{(\ell)}$ is pressure head. This introduces physically interpretable gradients while retaining a layered ODE-like update.

Appendix B.6. Step 5: Full multilayer Richards equation (reference ecohydrological model)

A full ecohydrological reference can be defined by the 1D Richards equation with a root-uptake sink:

$$\frac{\partial \theta(z,t)}{\partial t} = \frac{\partial}{\partial z}\left[K(\theta)\left(\frac{\partial h}{\partial z} + 1\right)\right] - S_{\text{root}}(h, z, t). \tag{B.14}$$

Discretisation (outline). Using L layers, define $\theta_t^{(\ell)}$ and $h_t^{(\ell)}$ per layer. Compute conductivities $K^{(\ell)}$ and interface fluxes $q^{(\ell+1/2)}$ using hydraulic functions (e.g., van Genuchten–Mualem). Update via an implicit scheme for stability:

$$\boldsymbol{\theta}_{t+1} \text{ solves } \boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_t = \Delta t\, \mathbf{F}(\boldsymbol{\theta}_{t+1}) - \Delta t\, \mathbf{S}_{\text{root}}(\boldsymbol{\theta}_{t+1}), \tag{B.15}$$

typically requiring Newton iterations. Boundary conditions represent infiltration at the surface and free drainage or fixed head at the bottom.

Use in the present framework. This step is most suitable as a high-fidelity "reference simulator" for offline benchmarking, stress testing, and residual-learning targets, rather than for large-scale RL training loops.

Appendix B.7. Summary

The described pathway enables a systematic transition from a single-bucket representation (Step 0) to an ecohydrologically grounded, layered model that incorporates delayed responses and heterogeneity (Steps 2 to 4), culminating in a Richards-equation-based reference (Step 5). This approach facilitates transparent model development, controlled increases in complexity, and reproducible assessment of controller robustness as physical realism increases.

Appendix C. Purpose and control coupling of alternative soil–water modelling formulations

This appendix clarifies the conceptual purpose, physical realism, and degree of control coupling for the various soil–water modelling formulations examined in this study. The aim is to position the proposed model in relation to classical Richards-equation-based and ecohydrological formulations, and to highlight the respective advantages and limitations of each approach within closed-loop irrigation control frameworks.

Appendix C.1. Proposed physics-based and hybrid control-aware formulation

The soil–water model used in this study is a reduced, discrete-time representation of root-zone water balance, specifically designed for closed-loop irrigation control. It integrates mass conservation, physically interpretable fluxes such as evapotranspiration, drainage, and storage, and incorporates stress-modulated plant response within a daily decision-making framework consistent with operational irrigation practices.

A key characteristic of this formulation is its control awareness. Rather than providing a high-fidelity description of soil hydrodynamics at fine spatial or temporal scales, the model is designed to: (i) remain numerically stable under repeated policy exploration, (ii) support large numbers of simulation episodes required by RL, (iii) expose meaningful and actionable state variables to the controller, and (iv) allow systematic isolation of control effects from physical modelling assumptions.

The hybrid extension augments this structure with a learned residual correction that compensates for systematic modelling errors while maintain-

ing the underlying physical structure. This design allows learning-based controllers to explicitly address model discrepancies, thereby enhancing robustness under stochastic climatic conditions without compromising interpretability or computational efficiency.

Advantages.

- The model aligns closely with daily irrigation decision cycles and the availability of sensor data.
- It demonstrates computational efficiency that is suitable for RL applications and sensitivity analyses.
- There is an explicit separation among physical dynamics, control logic, and learned correction components. The approach maintains stable behavior under stochastic forcing and during exploratory control actions.

Limitations.

- Vertical soil heterogeneity and fine-scale redistribution processes are represented through parameterization rather than explicit resolution.
- Memory effects extending beyond the daily time step are approximated using storage dynamics instead of continuous-time processes.
- Predictive fidelity is reduced under conditions dominated by strong vertical gradients or preferential flow.

Appendix C.2. Richards-equation-based soil–water models

Richards-equation-based models offer a physically rigorous framework for describing unsaturated flow in soils by resolving pressure head gradients, hydraulic conductivity, and moisture redistribution in continuous space and time. These models are extensively used in hydrology and soil physics to investigate infiltration, redistribution, and drainage under well-defined boundary and initial conditions.

However, Richards-based formulations are primarily descriptive rather than control-oriented. Their numerical complexity, stiffness, and sensitivity to parameter uncertainty present significant challenges for direct integration into closed-loop control or RL frameworks.

Advantages.

- These models exhibit high physical fidelity and provide explicit representation of vertical soil processes.
- They are capable of capturing delayed responses, sharp wetting fronts, and nonlinear hydraulic behavior. These models are suitable for use as reference models in benchmarking or offline validation studies.

Limitations for control.

- They incur high computational costs and impose restrictive time-step requirements.
- Numerical instability may occur under exploratory or extreme control actions.

- There is limited compatibility with large-scale policy learning or online decision-making processes.
- It is challenging to provide interpretable, low-dimensional state representations for use by controllers.

Consequently, within the present framework, Richards-equation-based models are most appropriately regarded as high-fidelity reference simulators or as sources of training data for residual learning, rather than as operational control environments.

Appendix C.3. Classical Ecohydrological Models

Ecohydrological models serve as an intermediate approach between fully physical Richards-equation-based formulations and simplified bucket-type models. These models typically represent the soil–plant–atmosphere continuum using multiple soil layers, explicit root water uptake profiles, evapotranspiration partitioning, and simplified vertical redistribution processes. They are often formulated as systems of coupled ordinary differential equations.

These models are well suited for long-term ecosystem analysis, water balance studies, and investigations of climate–vegetation interactions, where the primary objective is to understand system behavior rather than to optimize control actions. Consequently, they have traditionally been developed as descriptive or predictive tools, not as components of closed-loop decision-making systems.

Advantages.

- Explicit and physically meaningful representation of soil–plant–atmosphere interactions.

- Improved handling of vertical heterogeneity, delayed redistribution, and depth-dependent root water uptake compared to single-layer bucket models.
- Enhanced interpretability through process-level decomposition, enabling diagnostic insight into the relative contributions of evaporation, transpiration, and layer-wise water uptake.
- Strong grounding in ecohydrological theory and empirical observations, facilitating scientific analysis and model validation.

Limitations for control integration.

- Increased state dimensionality raises the complexity of policy learning and state-space exploration for RL, even when individual state variables remain physically interpretable.
- Longer memory effects and internal redistribution dynamics reduce controller responsiveness at daily decision scales and complicate credit assignment.
- Computational cost grows rapidly with the number of layers and nonlinear process parameterizations, limiting suitability for large-scale policy training and repeated stochastic evaluation.
- Limited native support for tight coupling with control objectives, uncertainty-aware exploration, and reproducible benchmarking under exploratory control actions.

Appendix C.4. Comparative perspective and rationale for model choice

The modelling approach adopted in this study represents a deliberate balance between physical realism and compatibility with control frameworks. Although Richards-equation-based and ecohydrological models provide enhanced descriptive capabilities, their direct integration with RL frameworks is frequently impractical.

In contrast, the proposed formulation preserves key ecohydrological mechanisms while maintaining explicit compatibility with closed-loop control and learning. Embedding physical constraints within a control-oriented structure and incorporating learned residual dynamics enables systematic exploration of irrigation strategies under uncertainty, while retaining numerical stability and interpretability.

Therefore, the proposed model serves as a complementary decision-support abstraction, specifically designed for adaptive irrigation control, benchmarking, and policy evaluation under stochastic climatic conditions, rather than as a substitute for detailed ecohydrological simulation.

Appendix D. Stability and well-posedness of the control-aware soil–water model

This appendix presents a mathematical analysis of the proposed daily soil–water dynamics, establishing the following properties:

(i) *Well-posedness (existence and uniqueness)*: for any admissible initial condition and any bounded climatic forcing and irrigation input, the state update equations admit a *unique* solution at each daily time step,

yielding a well-defined state trajectory over the entire simulation horizon.

(ii) *Positive invariance of the physical domain*: the soil–water storage remains within a physically meaningful bounded set for all time, i.e., if the initial storage lies within the admissible domain $[0, S_{\max}]$, then all subsequent states generated by the update equations remain in this domain, ensuring mass conservation and physical consistency.

(iii) *Bounded-input bounded-state stability (BIBS)*: under bounded climatic forcing and bounded irrigation actions, the resulting soil-water state remains bounded for all time. Specifically, the system dynamics exhibit an input-to-state-like stability property, such that variations in external forcing result in proportionate and bounded variations in soil-water storage, thereby preventing unphysical divergence or numerical instability.

These properties justify the model's use as a numerically stable closed-loop environment for control and RL.

Appendix D.1. Model definition

We consider the daily storage update (Eq. (4)):

$$S_{t+1} = \mathrm{clip}(S_t + U_t - L_t(S_t),\ 0,\ S_{\max})\,, \tag{D.1}$$

where

$$U_t := \eta_I I_t + R_t \geq 0$$

is the total daily input (irrigation + rainfall), and

$$L_t(S_t) := ET_{c,t}(S_t) + D(S_t) \geq 0$$

is the total loss, combining evapotranspiration and drainage. Evapotranspiration is defined as

$$ET_{c,t}(S_t) = Kc_t\, ET0_t\, f_{\mathrm{ET}}\big(\psi_t\big), \qquad \psi_t = f_{\mathrm{ret}}(S_t). \tag{D.2}$$

In the hybrid setting (Scenario 3), the simulator maintains S_t and $\psi_t = f_{\mathrm{ret}}(S_t)$, and applies a residual correction in tension space:

$$\psi_{t+1} = \psi_{t+1}^{\mathrm{phys}} + \Delta\psi_t, \qquad S_{t+1} = f_{\mathrm{ret}}^{-1}(\psi_{t+1}), \tag{D.3}$$

with $\Delta\psi_t = f_\theta(\psi_t, I_t, R_t, ET0_t)$.

Appendix D.2. Assumptions

We state standard boundedness and regularity assumptions consistent with operational agro-hydrological practice.

1. **Bounded inputs.** There exist finite constants $I_{\max}$, $R_{\max}$, $ET0_{\max}$ and $Kc_{\max}$ such that

   $$0 \le I_t \le I_{\max}, \quad 0 \le R_t \le R_{\max}, \quad 0 \le ET0_t \le ET0_{\max}, \quad 0 \le Kc_t \le Kc_{\max}.$$

2. **Stress factor boundedness.** The evapotranspiration stress reduction satisfies

   $$0 \le f_{\mathrm{ET}}(\psi) \le 1 \quad \text{for all admissible } \psi.$$

3. **Drainage nonnegativity.** The drainage function satisfies $D(S) \ge 0$ for all $S \in [0, S_{\max}]$.
4. **Retention curve properties.** The mapping $f_{\mathrm{ret}} : [0, S_{\max}] \to [\psi_{\min}, \psi_{\max}]$ is continuous, strictly monotone, and invertible on its image. Moreover, both f_{ret} and f_{ret}^{-1} are Lipschitz on their respective domains with constants L_{ret} and L_{ret}^{-1}.

5. **Residual boundedness (hybrid only).** The learned residual is bounded:

$$|\Delta\psi_t| \leq \Delta\psi_{\max} \quad \text{for all admissible inputs.}$$

In practice this is enforced by architecture choice (e.g., tanh output) and/or explicit clipping.

Existence and uniqueness follow directly from the explicit nature of the update equations. The state transition is defined as a deterministic, single-valued mapping $S_{t+1} = F(S_t, I_t, d_t)$ composed of continuous elementary functions (addition, multiplication, thresholding, and clipping). As no implicit equation or fixed-point problem is involved, a unique next-state value exists for every admissible input and initial condition.

Appendix D.3. Positive invariance and boundedness

Theorem 1 (Positive invariance of the storage domain). *For any initial condition $S_0 \in [0, S_{\max}]$ and any admissible input sequence, the trajectory generated by* (D.1) *satisfies*

$$S_t \in [0, S_{\max}] \quad \textit{for all } t \geq 0.$$

Proof 1. *By definition,* $\mathrm{clip}(x, 0, S_{\max}) \in [0, S_{\max}]$ *for any real* x*. Therefore, if* $S_t \in [0, S_{\max}]$*, then* S_{t+1} *defined by* (D.1) *belongs to* $[0, S_{\max}]$*. Since* $S_0 \in [0, S_{\max}]$*, the result follows by induction.*

Interpretation. Theorem 1 formalizes that the simulator is *physically admissible by construction*: storage cannot become negative and cannot exceed the prescribed capacity even under exploratory actions, which is essential for stable RL training.

Appendix D.4. Bounded-input bounded-state stability (BIBS)

Theorem 2 (BIBS stability of the physical update). *Under Assumptions 1–3, the storage trajectory of* (D.1) *is bounded for any bounded input sequence. In particular,*

$$0 \le S_t \le S_{\max} \quad \forall t,$$

and the losses satisfy the uniform bound

$$0 \le ET_{c,t}(S_t) \le Kc_{\max} ET0_{\max}, \qquad 0 \le D(S_t) \le \sup_{S \in [0, S_{\max}]} D(S).$$

Proof 2. *The state bound is Theorem 1. For evapotranspiration, by* (D.2) *and Assumption 2,*

$$ET_{c,t}(S_t) = Kc_t ET0_t f_{\mathrm{ET}}(\psi_t) \le Kc_{\max} ET0_{\max}.$$

Drainage is nonnegative by Assumption 3, and bounded on the compact interval $[0, S_{\max}]$ *if* D *is continuous (standard for threshold/linear forms used here).*

Appendix D.5. Stability of the hybrid residual coupling

We now show that the hybrid correction in tension space preserves boundedness and well-posedness, provided the residual is bounded and the retention mapping is invertible on a bounded interval.

Theorem 3 (Well-posedness and boundedness of the hybrid update). *Assume Theorem 1 holds for the physical predictor and Assumptions 4–5 hold. If* ψ_{t+1} *in* (D.3) *is projected onto* $[\psi_{\min}, \psi_{\max}]$ *(equivalently, if* $\Delta\psi_t$ *is bounded tightly enough such that* $\psi_{t+1} \in [\psi_{\min}, \psi_{\max}]$*), then the hybrid update is well-defined and yields*

$$S_{t+1} \in [0, S_{\max}] \quad \forall t.$$

Proof 3. *By Theorem 1, the physical predictor yields* $S_{t+1}^{\text{phys}} \in [0, S_{\max}]$*, hence* $\psi_{t+1}^{\text{phys}} = f_{\text{ret}}(S_{t+1}^{\text{phys}}) \in [\psi_{\min}, \psi_{\max}]$ *by Assumption 4. If* ψ_{t+1} *is ensured to remain in* $[\psi_{\min}, \psi_{\max}]$*, then* $S_{t+1} = f_{\text{ret}}^{-1}(\psi_{t+1})$ *is well-defined and belongs to* $[0, S_{\max}]$ *by invertibility of* f_{ret} *on this interval.*

Practical remark (implementation detail).. In a control-aware simulator, it is appropriate to enforce the projection

$$\psi_{t+1} \leftarrow \text{clip}(\psi_{t+1}^{\text{phys}} + \Delta\psi_t,\ \psi_{\min}, \psi_{\max})$$

before applying f_{ret}^{-1}. This preserves physical admissibility without altering the conceptual role of the residual.

Appendix D.6. Incremental stability (optional Lipschitz bound)

To quantify sensitivity, we provide a simple incremental bound. Let S_t and $\tilde{S}_t$ be two trajectories driven by the same exogenous forcing but possibly different actions, and define $\delta_t := |S_t - \tilde{S}_t|$.

Proposition 1 (One-step incremental bound (physical update)). *Assume* $D(\cdot)$ *and* $f_{\text{ET}}(f_{\text{ret}}(\cdot))$ *are Lipschitz on* $[0, S_{\max}]$ *with constants* L_D *and* L_{ET} *(uniformly in* t *through bounded* Kc_tET0_t*). Then the pre-clipped map satisfies*

$$|S_{t+1}^{\text{raw}} - \tilde{S}_{t+1}^{\text{raw}}| \leq (1 + L_D + L_{ET})\,|S_t - \tilde{S}_t| + \eta_I |I_t - \tilde{I}_t|,$$

and after clipping,

$$|S_{t+1} - \tilde{S}_{t+1}| \leq |S_{t+1}^{\text{raw}} - \tilde{S}_{t+1}^{\text{raw}}|,$$

since $\text{clip}(\cdot, 0, S_{\max})$ *is non-expansive.*

Proof 4. *Triangle inequality gives the bound for the raw update. Non-expansiveness of clipping follows from the fact that projection onto a convex interval is 1-Lipschitz.*

Interpretation. Proposition 1 formalizes the numerical stability of the environment under exploratory control. The one-step map is Lipschitz, and the clipping operation is non-expansive, which together prevent uncontrolled amplification of small perturbations.

Appendix D.7. Summary

The proposed environment demonstrates the stability necessary for closed-loop benchmarking. The soil–water state remains physically bounded for all admissible actions and climatic forcing, ensuring positive invariance. Hydrological fluxes remain bounded under bounded drivers, reflecting bounded-input bounded-state (BIBS) stability. Additionally, the hybrid residual coupling preserves well-posedness under mild boundedness and invertibility assumptions.

Furthermore, the stability and well-posedness analysis establishes continuous dependence of the soil–water state on initial conditions, control inputs, and climatic forcing. This property is a necessary prerequisite for sensitivity analysis. Explicit analytical parametric sensitivity metrics are not derived in this appendix. Instead, sensitivity is investigated empirically through controlled scenario analyses, consistent with the control-oriented and benchmarking-focused scope of the model.

Collectively, these properties support the control-aware purpose of the formulation. The focus is on ensuring stable, interpretable, and reproducible

closed-loop simulations under uncertainty, rather than replicating all fine-scale soil hydrodynamics.

References

Alkaff, M., Basuhail, A., Sari, Y., 2025. Optimizing water use in maize irrigation with reinforcement learning. Mathematics 13, 595. doi:10.3390/math13040595.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. Number 56 in FAO Irrigation and Drainage Paper, Food and Agriculture Organization of the United Nations, Rome, Italy.

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., Turchetta, M., Schoellig, A., Krause, A., 2017. Safe model-based reinforcement learning with stability guarantees. Advances in neural information processing systems 30, 908–918.

Bo, Y., Liang, H., Li, T., Zhou, F., 2024. Process-based modeling framework for sustainable irrigation management at the regional scale: Integrating rice production, water use, and greenhouse gas emissions. Geoscientific Model Development Discussions 2024, 1–29.

FAO, 2025. Fao-56 soil water accounting. https://swb.hydroagrinexus.com/. FAO-56 Soil Water Simulator.

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.

Fereres, E., Soriano, M.A., 2007. Deficit irrigation: A global perspective. Agricultural Water Management 80, 1–22. doi:10.1016/j.agwat.2005.07.014.

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.

Höge, M., Scheidegger, A., Baity-Jesi, M., Albert, C., Fenicia, F., 2022. Improving hydrologic models for predictions and process understanding using neural ODEs. Hydrology and Earth System Sciences 26, 5085–5106. doi:10.5194/hess-26-5085-2022.

Huang, Z., Wang, Y., Hui, C., XiaoCheng, 2025. An intelligent water-saving irrigation system based on multi-sensor fusion and visual servoing control. arXiv preprint arXiv:2510.23003. doi:10.48550/arXiv.2510.23003.

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.

Li, Z., Zhang, Y., Wang, J., Chen, X., Li, J., 2022. Deep learning for intelligent irrigation decision-making: A review. Agricultural Water Management 265, 107561. doi:10.1016/j.agwat.2022.107561.

Liu, G., Amini, A., Pandey, V., Motee, N., 2023. Data-driven distributionally robust mitigation of risk of cascading failures. arXiv preprint arXiv:2310.12021. doi:10.48550/arXiv.2310.12021.

Monteith, J.L., 1965. Evaporation and environment. Symposia of the Society for Experimental Biology 19, 205–234.

Nghiem, T.X., Drgoňa, J., Jones, C., Nagy, Z., Schwan, R., Dey, B., Chakrabarty, A., Di Cairano, S., Paulson, J.A., Carron, A., Zeilinger,

M.N., Cortez, W.S., Vrabie, D.L., 2023. Physics-informed machine learning for modeling and control of dynamical systems. arXiv preprint arXiv:2306.13867. doi:10.48550/arXiv.2306.13867.

Padilla-Nates, J.P., Garcia, L.D., Lozoya, C., Orona, L., Cortes-Perez, A., 2025. Greenhouse irrigation control based on reinforcement learning. Agronomy 15, 2781. doi:10.3390/agronomy15122781.

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.

Portu, J.C., Delle Femine, C., Muro, K.S., Quartulli, M., Restelli, M., 2025. Limitations of physics-informed neural networks: a study on smart grid surrogation, in: 2025 IEEE Kiel PowerTech, IEEE. pp. 1–6.

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.

Ratn, S., Rampriyan, S., Ray, B., 2025. Hybrid physics–ml model for forward osmosis flux with complete uncertainty quantification. arXiv preprint arXiv:2512.10457. doi:10.48550/arXiv.2512.10457.

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.

Rodríguez-Iturbe, I., Porporato, A., 2004. Ecohydrology of Water-Controlled Ecosystems: Soil Moisture and Plant Dynamics. Cambridge University Press, Cambridge, UK.

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

Saikai, Y., Peake, A., Chenu, K., 2023. Deep reinforcement learning for irrigation scheduling using high-dimensional sensor feedback. PLOS Water 2, e0000169. doi:10.1371/journal.pwat.0000169.

Scanlon, B.R., Jolly, I., Sophocleous, M., Zhang, L., 2007. Global impacts of conversions from natural to agricultural ecosystems on water resources: Quantity versus quality. Water resources research 43. doi:doi:10.1029/2006WR005486.

Seneviratne, S.I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A.D., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., 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.

Seneviratne, S.I., et al., 2010. Investigating soil moisture–climate interactions in a changing climate. Earth-Science Reviews 99, 125–161.

Shamekh, S., Lamb, K.D., Huang, Y., Gentine, P., 2023. Implicit learning of convective organization explains precipitation stochasticity. Proceedings of the National Academy of Sciences 120, e2216158120.

Shang, C., Chen, W., Stroock, A.D., You, F., 2018. Robust model predictive control of irrigation systems with active uncertainty learning and data analytics. arXiv preprint arXiv:1810.05947. doi:10.48550/arXiv.1810.05947.

Siebert, S., Burke, J., Faures, J.M., Frenken, K., Hoogeveen, J., Döll, P., Portmann, F.T., 2010. Groundwater use for irrigation – a global inventory. Hydrology and Earth System Sciences 14, 1863–1880. doi:10.5194/hess-14-1863-2010.

Staff, S.N., 2025. Climate change reduces crop yields worldwide even with adaptation.

https://news.stanford.edu/stories/2025/06/climate-change-cuts-global-crop-yields Stanford Report.

Starke, C., Lünich, M., 2020. Artificial intelligence for political decision-making in the european union: Effects on citizens' perceptions of input, throughput, and output legitimacy. Data & Policy 2, e16. doi:10.1017/dap.2020.19.

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

Tincani, M., Kerouch, K., Garlando, U., Barezzi, M., Sanginario, A., Indiveri, G., Del Luca, C., 2025. A neuromorphic continuous soil monitoring system for precision irrigation. arXiv preprint arXiv:2509.14066. doi:10.48550/arXiv.2509.14066.

Verbruggen, W., Wårlind, D., Horion, S., Meunier, F., Verbeeck, H., Schurgers, G., 2025. Implementing a process-based representation of soil water movement in a second-generation dynamic vegetation model: application to dryland ecosystems (LPJ-GUESS-RE v1.0). EGUsphere , 1–28doi:10.5194/egusphere-2025-1259.

Vereecken, H., et al., 2007. Modeling soil processes: Review, key challenges, and new perspectives. Vadose Zone Journal 6, 749–762.

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.