ORIGINAL RESEARCH Open access

Evaluating Deep Reinforcement Learning for Irrigation Scheduling: Evidence from Long-Term Weather Variability

Zsolt Kovács[1*], Katalin Szabo[1], Gabor Nagy[2], Eszter Toth[3]

Abstract

With rising worldwide pressure from water shortages, scientists, government officials, and businesses are exploring new ways to boost water efficiency in farming. Drawing on recent breakthroughs in handling complex decisions, Deep Reinforcement Learning (DRL) is now being proposed as a tool to refine irrigation plans. Preliminary work has indicated potential financial benefits from DRL compared with traditional rule-based methods such as soil-moisture triggers or set timetables. Still, no thorough review exists that assesses DRL's value for planning irrigation, accounting for regional weather differences and water-use constraints. To close this research gap, we built Aquacrop-Gym. This publicly available Python platform enables experts to develop and test personalized irrigation plans using the AquaCrop-OSPy crop-water simulation. For this investigation, Aquacrop-Gym helped determine DRL's advantages compared with standard irrigation planning methods (such as an optimized soil-moisture heuristic) for growing maize in a high-irrigation zone of the central United States. The DRL models and heuristic methods were both trained on 70 years of synthetic weather records generated by the LARS-WG, then evaluated against 30 separate unseen validation years of similarly generated weather. Outcomes reveal that when rainfall fluctuates strongly, DRL fails to exceed well-tuned conventional heuristics. Yet, if rainfall is removed entirely, DRL delivers superior earnings on the unseen validation years. In parallel, DRL surpasses optimized heuristics under tight water-use limits. Overall, this work highlights DRL as a valuable technique for managing irrigation, especially in situations where growers face major physical or policy-driven water shortages.

Keywords Deep learning, Irrigation, Water, Optimisation, Economic, Scheduling

*Correspondence:
Zsolt Kovács
zsolt.kovacs@szentistvan.hu
[1] Department of AI for Sweet Corn and Pepper Ripeness Detection, Faculty of Agriculture, Szent István University, Budapest, Hungary
[2] Department of Machine Learning for Light Interception and Canopy Structure, Faculty of Agriculture, University of Debrecen, Debrecen, Hungary
[3] Department of Smart Pesticide Reduction Strategy, Faculty of Agriculture, University of Pécs, Pécs, Hungary

Introduction

Farming accounts for the largest share of freshwater consumption in regions facing water limits, making it both a driver and a casualty of water scarcity. Raising water productivity in agriculture thus forms a core part of worldwide strategies to combat water shortages and resolve food supply problems [1, 2].

Improving agricultural water productivity emphasizes matching irrigation volumes to actual crop demand and reducing wasteful water losses. Timely irrigation planning throughout the crop cycle, combined with upgrades such as modern equipment or improved field practices, remains essential for meaningful productivity gains. Typically, growers rely on basic guidelines or heuristics tied to factors such as soil water status or plant health [3]. These rules might remain

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

unchanged throughout the season or adjust with crop development phases. Even though such straightforward rule-driven strategies feel natural and align with good farming principles, they risk leading to poor choices when the original assumptions that underpin them no longer hold due to surprises. A sudden heavy downpour or prolonged dry spell, for instance, can turn a static irrigation rule wasteful for that season. Likewise, sudden shifts in water supply — perhaps due to broken machinery or urgent drought policies — can leave existing heuristics mismatched with the new constraints.

A potential fix involves updating the heuristics mid-season by recalibrating rules and decision points once more seasonal weather patterns emerge. Prior studies confirm that linking crop models to on-site measurements, forecasts, and extra inputs can enable such adjustments, leading to better harvests, water savings, and income [4-7]. That said, this process demands significant computing resources as the season unfolds, especially with repeated recalibrations, driving up costs for farmers or advisory services. When handling vast areas with tens of thousands of unique crop setups, performing real-time recalibrations often proves too costly or impractical. On top of in-season adjustment hurdles, integrating expanding data streams available to farmers would require many layered, interdependent rules that are tough to craft and fine-tune. Deciding on the right structure for a heuristic that uses short-range forecasts (for example, apply water only if five rain-free days lie ahead) and effectively merges it with soil or canopy readings is far from clear.

To address these limitations, AI methods, including Deep Reinforcement Learning (DRL) [8], are recommended for dynamically and intelligently updating choices throughout the season using available signals. DRL has already exceeded human expert performance in demanding domains such as digital games and strategy boards [9-12]. Lately, several projects have tested DRL's suitability for irrigation timing [13-15]. Compared with unrefined plans and basic heuristics, DRL produced notable advances in both water efficiency and returns, suggesting its practical promise. However, every referenced study tested the trained DRL agent on very few unseen seasons (under 3 in each case), a too-narrow sample to judge reliability amid the wide range of unpredictable weather farmers actually experience. Moreover, comparisons were made only against researcher-selected fixed schedules (chosen dates and amounts), rigid intervals (such as 25 mm every N days), or basic single-threshold rules (apply water at a set soil moisture drop), all picked manually rather than optimized to the local setting. These limited reference points hinder a clear view of DRL's real edge over the refined heuristics standard in current irrigation studies [16-18].

Here, we conduct a more rigorous examination of DRL's potential benefits for irrigation planning relative to standard and optimized heuristics across different climate settings. We investigate how choices in system design — including the DRL agent's decision frequency and timing, and varying water restriction levels — shape projected gains in productivity and earnings. These evaluations supply updated guidance on properly testing DRL's usefulness in irrigation and clarify the specific environments where it can strongly support higher water productivity.

Materials and Methods

This section describes the system for training and testing DRL agents for irrigation management. We begin with a summary of DRL (Section 2.1), then outline its application to irrigation decisions inside a crop simulation environment (Section 2.2). Next, the case study description and experimental setup (Section 2.3), a review of overfitting risks and best practices for presenting DRL results (Section 2.4), and, finally, an explanation of the optimized heuristic used as a baseline to gauge DRL's additional contribution (Section 2.5).

Deep reinforcement learning

Reinforcement learning is typically structured around an agent interacting with its environment **(Figure 1)**. At each step, the agent observes the environment (O) and chooses an action (A). The chosen action alters the surroundings, prompting a return of reward R together with the following observation O. This process repeats until the surroundings hit an ending condition, concluding the episode. The reinforcement learning agent seeks to develop a function that maps observations to actions to maximize the total future reward [8].

© 2024 The Author(s). | Published by Agro Intelligence Research Press

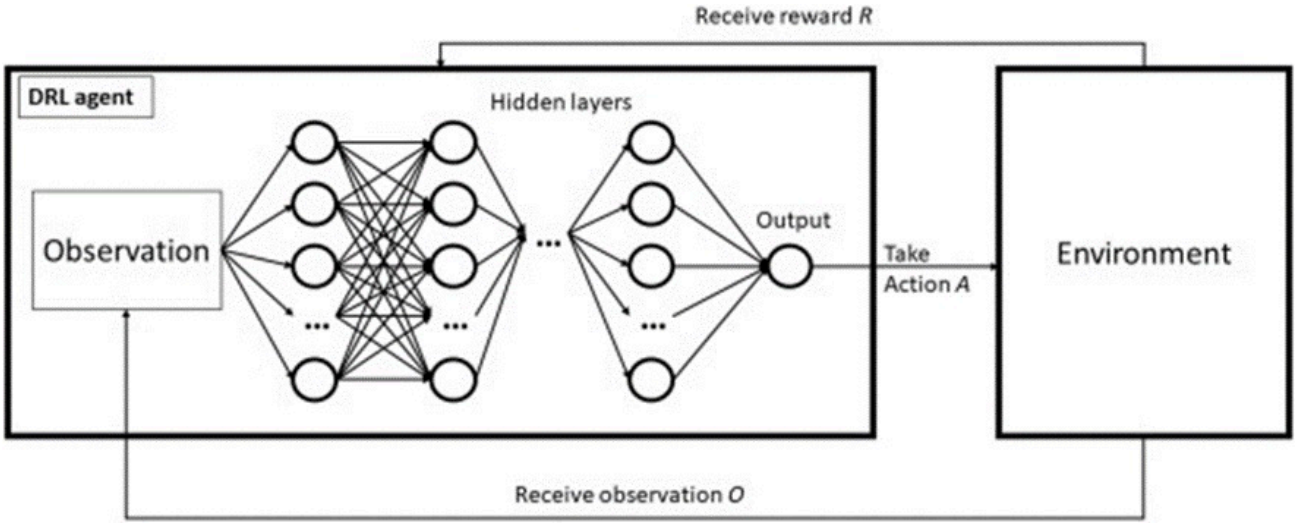


Figure 1. Typical framing for a reinforcement learning problem, where a DRL agent aims to maximize the sum of future rewards.

In Deep Reinforcement Learning (DRL), the connection from observations to actions is implemented using an artificial neural network, whose parameters are adjusted based on rewards collected by the agent. This examination adopts Proximal Policy Optimization (PPO) [19], one of the leading DRL methods applied across numerous domains, ranging from robotics [20] and video games [10] to flow control [21]. A key benefit of PPO is its seamless support for both continuous and discrete action spaces, which enables us to explore multiple irrigation approach formats. Preliminary trials confirmed that PPO yielded the fastest and most reliable learning curves among the methods examined. Implementation of PPO draws on the Ray RLLib Python package [22], a robust, actively maintained industrial tool now in use across various fields (https://docs.ray.io/en/latest/rllib/index.html).

Applying DRL to irrigation scheduling

Developing DRL agents capable of making irrigation decisions requires a crop modeling setup that can swiftly evaluate various strategies. Additionally, the setup must respect the Observation–Action–Reward–next Observation (O–A–R–O) sequence noted before. In the current work, AquaCrop-OSPy [16] models the agricultural system and serves as an interactive platform that enables the DRL agent to select irrigation actions and learn their impact on seasonal earnings. AquaCrop-OSPy is the open-source Python implementation of the AquaCrop crop-water productivity simulator [23] developed by the UN-FAO. This model stands out as one of the leading crop simulators, cited in over 334 scholarly papers in the 10 years following its debut [24]. It delivers robust predictions of crop responses to water limitations while keeping data requirements modest, making it suitable for areas with sparse information. AquaCrop-OSPy enables external control of intricate irrigation decisions, easing integration with DRL packages or search routines. We built the open-source aquacrop-gym Python toolkit (https://github.com/aquacropos/aquacrop-gym) to support seamless exchanges between the DRL agent and AquaCrop-OSPy.

Training in the Aquacrop-Gym begins by launching a DRL agent with arbitrary neural network parameters. The agent then runs several training episodes to accumulate experience **(Figure 2)**. Each episode randomly selects a weather year from the training collection and runs an AquaCrop-OSPy simulation. Every N days during the run, AquaCrop-OSPy delivers observations to the agent (e.g., the current soil water level). The agent interprets them and proposes a move (for instance, water application volume). The run continues for another N days, allowing the agent to issue subsequent moves based on fresh observations. When the season wraps up, the final reward (seasonal earnings) is calculated and returned to the agent.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

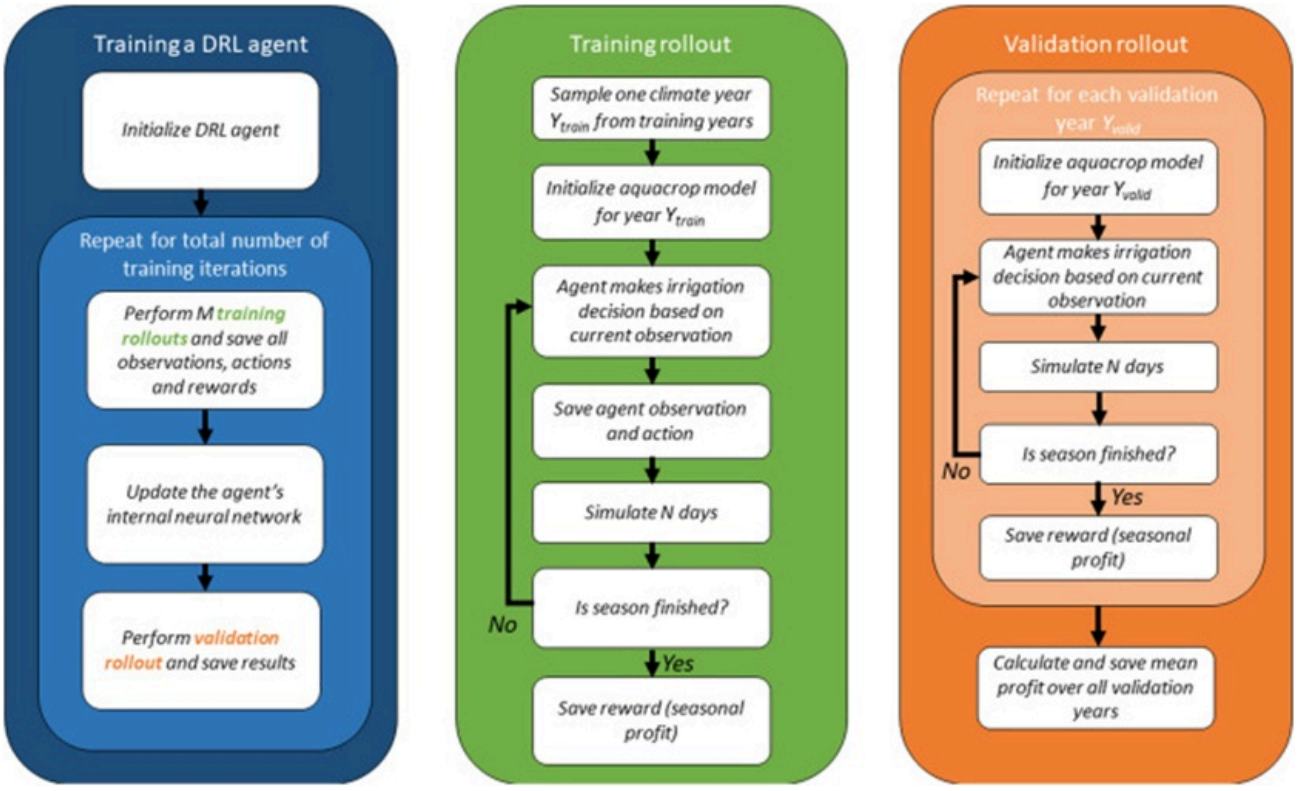


Figure 2. Flow diagram detailing how the DRL agent collects training data, which is then used to update the agent's decision-making. After updating, the new strategy is evaluated on the validation set.

Following completion of a training episode, the recorded (O–A–R–O) steps adjust the internal neural network parameters. With the policy refreshed, validation is performed through a separate rollout **(Figure 2)**. Validation tests the agent on each unseen validation year and computes the average profit level. Because no (O–A–R–O) steps are kept during validation, learning remains inactive, and only the average reward tracks progress. The cycle of assembling training experience, refining the policy, and checking performance repeats across a defined number of cycles or until halted by hand.

Case study, DRL design choices, and experimental details

This examination models irrigated maize cultivation in southwest Nebraska via AquaCrop-OSPy. Maize ranks as the top irrigated crop across Nebraska, the leading state by irrigated land area nationwide [3]. Most irrigation draws from the High Plains aquifer, which has shown sharp reductions in reserves due to prolonged heavy extraction [25, 26]. In parts of Nebraska, local authorities cap farmers' water withdrawals to limit aquifer depletion and its side effects, including lower river flows [27].

Model preparation and validation rely on past weather records gathered at a Champion, Nebraska, station [28] spanning 37 years (1982–2018), including daily low and high temperatures, rainfall, and solar radiation. These inputs drive the LARS-WG weather generator [29] to create an extended 100-year artificial dataset. The generator extracts statistical traits from the supplied series and produces new daily sequences, maintaining equivalent variation levels within and across years. Recent research verifies LARS-WG reliably mirrors important weather metrics (monthly averages and spreads for temperature or rainfall) from the supplied series [30, 31]. The technique boosts available years to 70 for training and 30 for validation, all while closely matching real historical patterns. Remaining inputs for AquaCrop runs, covering crop and soil settings, are based on published works on maize in the central United States [18, 32, 33]. Profit computations at season close adopt a grain value of $180 per tonne [34], water application expenses of $1 per ha-mm [35], plus steady other production expenses of $1728 ha−1 [35]. Entire code, model, and parameter information hosted at

© 2024 The Author(s). | Published by Agro Intelligence Research Press

(https://github.com/aquacropos/aquacrop-gym). Seasonal profit serves as the core success measure because it captures the central driver for operators in efficient, modern farming regions like the central United States. Furthermore, focusing on profit widens the scope for revealing DRL gains by exposing the balance between water spending and revenue from harvests.

Selecting suitable observations for the agent, allowable moves, and targeted reward forms a vital setup decision shaped by the modeled cropping system. Reviewing every possible mix lies beyond this study's reach. We instead probe several common and practical mixes that fit the central United States context. **Table 1** presents the action categories tested. Under depth-discrete, the agent selects a water volume for the next day from the listed options (0, 2.5, 5, …, 22.5, 25 mm ha−1). Depth continuous permits any volume inside the limits [0, 25 mm ha−1] for the next day. Binary requires the agent to decide whether to apply or skip 25 mm the next day. Soil-moisture threshold lets the agent pick a moisture cutoff that remains in effect until the next choice; once moisture falls below the cutoff, watering starts to restore field capacity (limited to 25 mm ha−1). For each action category, testing spans four time intervals (1, 3, 5, and 7 days).

Table 1. The action sets this case study apart.

Action set	Example action	Bounds
Discreet depth	0, 2.5, … 22.5, 25 mm ha^{-1}	[0,25] mm ha^{-1}
Continuous depth	8.43 mm ha^{-1}	[0,25] mm ha^{-1}
Binary	0, 25 mm ha^{-1}	[0,25] mm ha^{-1}
Soil moisture threshold	63.2 %TAW	[0,100]% TAW

In every instance, the information given to the agent ahead of each irrigation choice includes: mean daily reference evapotranspiration and precipitation totals for the current season up to that point; along with the following variables limited to the most recent 7 days: current root-zone soil-moisture content, canopy cover, biomass growth, cumulative irrigation applied, water supply remaining, crop growth stage, and the total number of days elapsed since the season started. These variables come from either the input weather time series or outputs generated by the AquaCrop-OSPy model. This selection of inputs covers a broad set of weather, soil, and crop state details that farmers could realistically track or gather to enable a DRL-driven scheduling method, and it can be expanded or simplified based on the monitoring resources available in the specific production system under consideration.

As noted in the prior section, DRL agents will be trained on 70 years of data and assessed on 30 years of validation data. The mean and standard deviation of seasonal rainfall (May–September) in the training and validation sets were very similar, at 309.6 ± 93.7 and 306 ± 93.2 mm ha−1, respectively. Even though both datasets originate from the same statistical distribution, the specific time series differ, offering a means to evaluate how effectively the agent would perform in real-world conditions with unpredictable weather. To examine the effect of this local variability on DRL outcomes, an additional experiment was conducted in which rainfall was removed from both the training and validation sets. We hypothesise that eliminating this major element of weather variability, thereby lessening the discrepancy between training and validation time series, will lead to better DRL performance than when rainfall data is included.

Earlier studies on DRL for irrigation scheduling have assumed unlimited water availability, enabling the DRL agent to apply irrigation without restrictions. However, this does not reflect conditions in our case study region of Nebraska, nor in many other parts of the world where water use is constrained by regulatory measures such as annual or multi-year quotas and allocations, or by physical supply and infrastructure limitations [27, 36]. Since the greatest marginal benefits from additional water occur during periods of shortage, we hypothesise that advanced approaches such as DRL could offer greater benefits when water resources are constrained. To investigate the influence of water restrictions on the usefulness of DRL, the existing framework is applied again with seasonal water-use caps of 75, 100, and 125 mm ha−1

© 2024 The Author(s). | Published by Agro Intelligence Research Press

enforced across all simulation years. Once this limit is reached in any growing season, no additional irrigation is permitted.

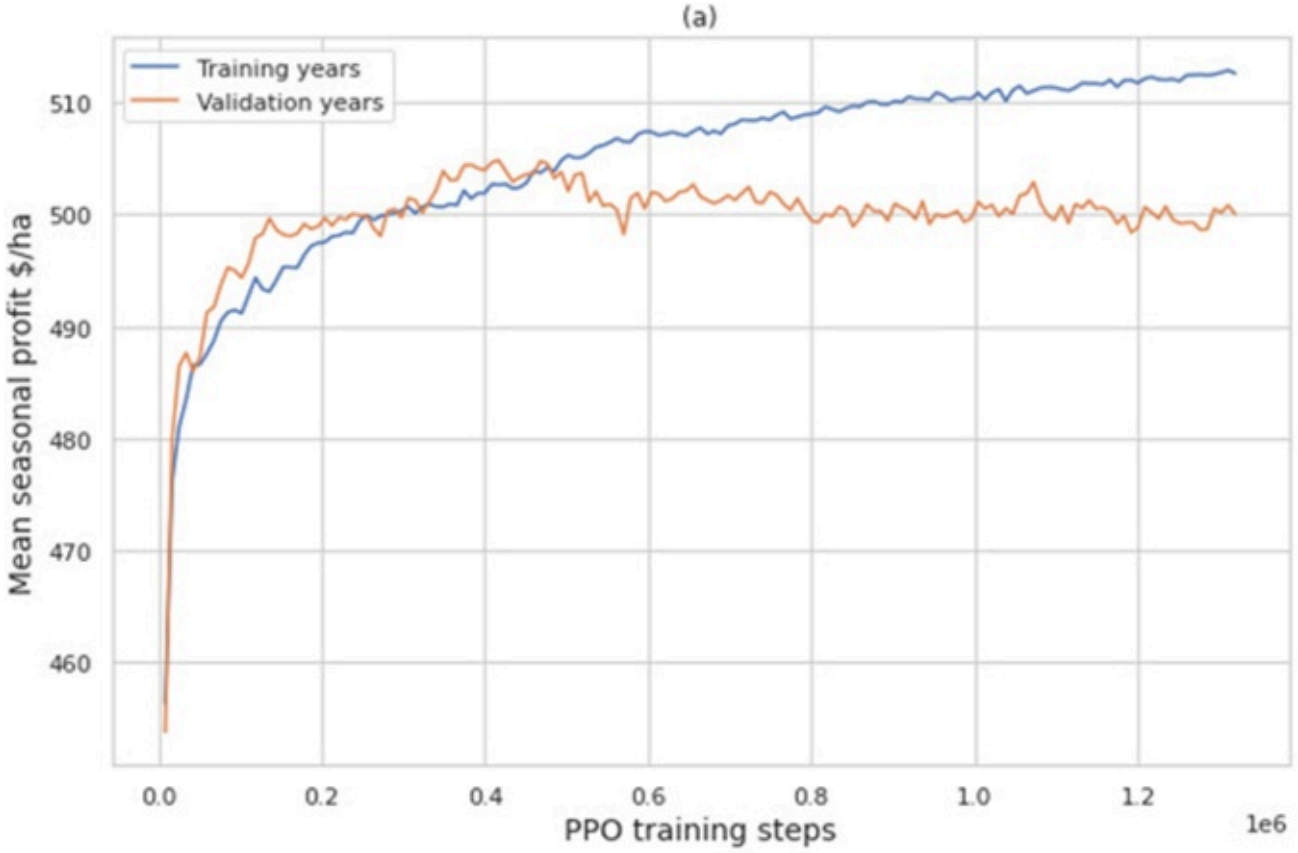
To summarise, four DRL agents will first be trained to evaluate the action sets outlined in **Table 1**. The strongest-performing action set will then be used to train another four DRL agents to compare the four decision intervals of 1, 3, 5, and 7 days. Using this selected DRL setup (i.e., action set and decision interval), two further experiments will be run – one maintaining the original local rainfall and one with rainfall removed entirely. Lastly, using the same DRL setup, three experiments will be performed for each of the seasonal water use limits 75, 100, 125 mm ha−1. This results in 13 experiments in total.

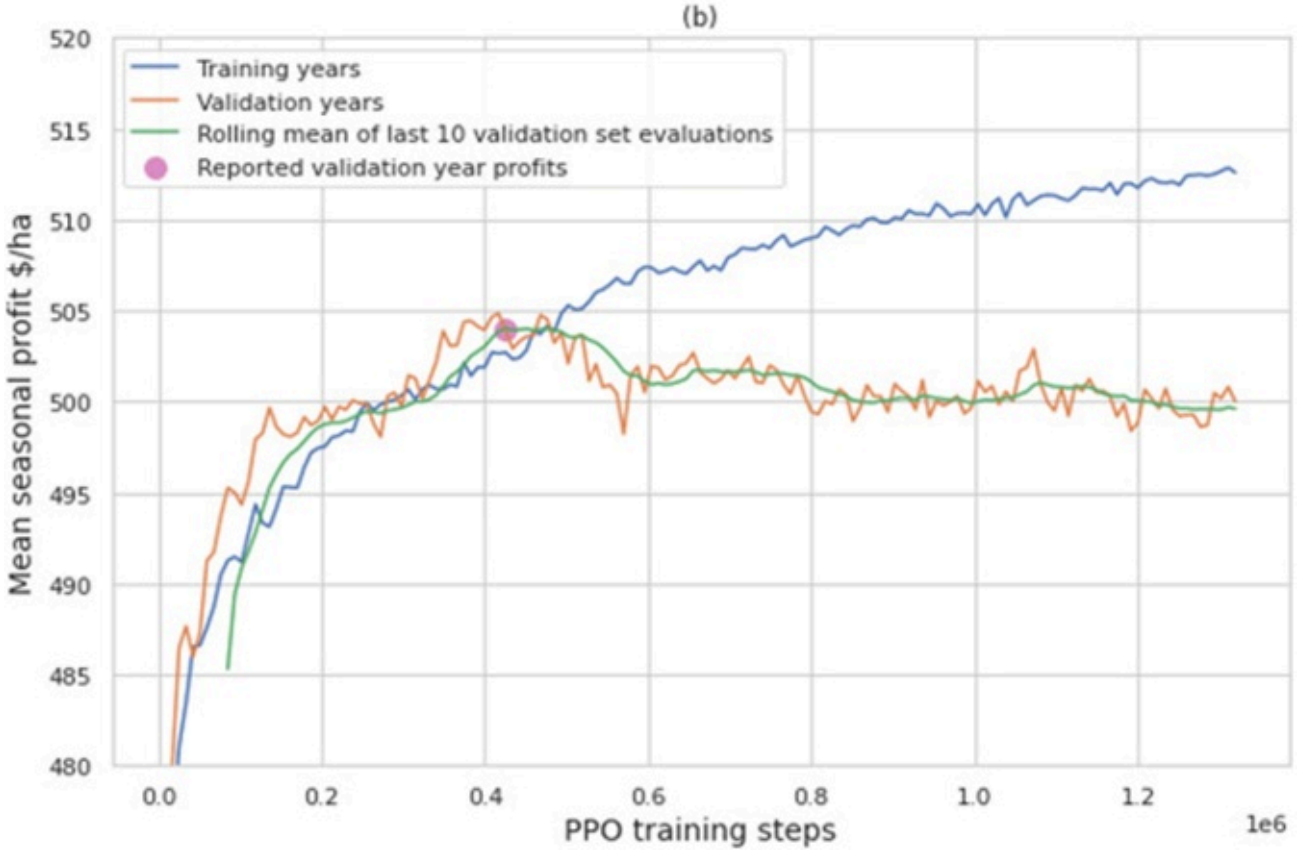
Learning curves, evaluation criteria, and overfitting

A key limitation of artificial neural network-based irrigation strategies is that neural networks frequently overfit to the training data when training continues for too long. Overfitting occurs when training performance continues to improve while validation performance begins to decline. This split arises because the network starts identifying patterns specific to the training set rather than broader patterns that also apply to the validation set. Due to this tendency, it is usually preferable to halt training before the model reaches peak performance on the training set, at which point overfitting has likely already occurred. To monitor the learning process, the DRL agent is regularly tested on both the training and validation sets throughout the training phase, enabling us to detect when overfitting begins.

Figure 3 presents these evaluation outcomes for a DRL agent that selects discrete irrigation depths every 7 days. **Figure 3a**, commonly known as the learning curve, shows the agent's performance relative to training steps (i.e., the number of irrigation decisions executed and used for learning). Training set performance continues to rise steadily. In contrast, validation set performance begins dropping after roughly 500,000 training steps. This gap between training and validation results raises a critical question about how best to report final DRL performance when comparing it with alternative irrigation methods. One option is to report the peak validation profits reached by the agent, which would illustrate the potential of DRL if training were halted at the optimal time. However, this would overstate DRL's capabilities because it presumes access to validation-year information, which would be unavailable to a farmer, since the validation set represents unknown future weather. An alternative is to report the final validation profit once training ends, thereby avoiding the need to assume knowledge of the validation years. Yet, as illustrated in **Figure 3a**, validation profits can fluctuate considerably between evaluation points, making a single result potentially unrepresentative. For this reason, we compute the rolling mean of the last 10 validation evaluations instead **(Figure 3b)**. When assessing DRL performance on the validation set, we report the highest 10-evaluation average.

© 2024 The Author(s). | Published by Agro Intelligence Research Press





© 2024 The Author(s). | Published by Agro Intelligence Research Press

Figure 3. (a) Average training profits and testing profits as a function of training steps, and (b) Average training profits and testing profits as a function of training steps, as well as a rolling average of validation profits over 10 evaluations. The highest point on this rolling average calculation will be reported as the validation profit.

Comparison with an optimal benchmark irrigation heuristic

Prior evaluations of DRL in irrigation scheduling have typically benchmarked DRL agents only against manually selected schedules or simple non-optimized single-value rules [13-15]. One widely adopted irrigation scheduling heuristic among farmers in this area relies on soil-moisture target thresholds. Under this method, irrigation is activated whenever soil-water content falls below a defined threshold, which can be adjusted to reflect the crop's varying sensitivity to water stress across different growth phases. As a reference point for evaluating our DRL agent, we examine a farmer employing four soil-moisture target thresholds aligned with the four primary growth stages in AquaCrop. Using Aquacrop-Gym, we determine a set of four soil-moisture thresholds that yield the highest average profit across the 70 training years. These thresholds are then tested on the 30 unseen validation years to establish a benchmark for measuring the additional benefit provided by DRL. This benchmark produces $513 ha−1 in the training years and $521 ha−1 in the validation years. It also yields $503 ha−1 when tested on the 37 years of historical data originally used to calibrate LARS-WG. Comparing these values with the profits obtained by directly optimizing the heuristic over those same 37 historical years ($506 ha−1) confirms that training and optimizing with synthetic data – along with the selected weather generator – yield reliable strategies for the case study region.

All optimization of the heuristic benchmark and DRL training runs were performed on one AMD Ryzen 7 3700X, 3.60 GHz, 8-core computer equipped with a GeForce RTX 3070 GPU. On this system, completing 1 million DRL training steps required roughly 4 h, while optimizing a set of four soil-moisture thresholds took about 20 min.

Results and Discussion

This section opens with an evaluation of how DRL outcomes shift according to selections of action sets and decision frequencies (Section 3.1). It then examines the role of climate variability in shaping DRL outcomes (Section 3.2) before exploring the addition of binding limits on irrigation water volumes (Section 3.3).

Comparing DRL configurations

Table 2 displays the average seasonal profits realized by several irrigation scheduling techniques following 1 million training steps. Across all variations in how irrigation events were activated, DRL consistently generated superior profits on the training set compared with the fixed, optimized irrigation scheduling heuristic. That said, the fixed heuristic slightly outperformed the various DRL approaches on the validation set. This pattern reveals that deep learning models can master sophisticated irrigation tactics (as evidenced by stronger training results). Yet, they often struggle to translate those gains into better performance under unfamiliar production conditions (namely, the validation set). Because the profit gaps among the strategies remain very narrow, the data here alone cannot justify discarding any particular scheduling method. Among the DRL options, the discrete depth action set delivered the highest validation profit, serving as the reference point for all subsequent comparisons with fixed soil-moisture threshold (SMT) heuristics.

Table 2. Reported seasonal profits (highest mean profits over 10 evaluations) for different action sets.

Scheduling method	Validation set profits $ ha^{-1}	Training set profits $ ha^{-1}
DRL – *discreet depth*	520.1	518.9
DRL – *continuous depth*	512.6	**520.2**
DRL – *binary*	517.8	515.3

© 2024 The Author(s). | Published by Agro Intelligence Research Press

DRL – *soil-moisture threshold*	515.7	517.5
Fixed Optimized SMT	**521.4**	513.1

The following DRL design factor under review concerned the decision interval (i.e., the frequency with which the agent determines irrigation actions). Using the leading action set (discrete depth), the agent was trained at intervals of 1, 3, 5, and 7 days over 1 million steps, and the validation set results were then contrasted. The 3-day interval stood out by delivering peak profits on both training and validation data **(Table 3)**. Profit spreads across these intervals were wider than those observed across the different action sets **(Table 2)**. Such a pattern implies that decision frequency is more significant than the precise rules for setting irrigation quantities once an event begins, at least within the boundaries of our chosen crop model and regional context.

Table 3. Reported seasonal profits (highest mean profits over 10 evaluations) for different decision intervals using the discrete depth action set defined in Section 2.3.

Decision interval	Validation set profits $ ha^{-1}	Training set profits $ ha^{-1}
1 day	497.6	492.2
3 days	**520.1**	**518.9**
5 days	511.0	515.1
7 days	504.0	510.4

Climate variability

Even with comparable spreads and statistical profiles of weather patterns across training and validation data, the outcomes detailed in Section 3.1 indicate that advanced DRL scheduling methods have so far only surpassed basic optimized soil-moisture thresholds during training. This shortfall may stem from the fact that certain weather sequences in the validation window differ markedly (for instance, in the precise ordering of rain events), despite matching broader statistical traits (such as overall seasonal precipitation). Removing rainfall altogether should reduce the divergence between the training and validation series, thereby improving DRL performance on the validation data. To verify this idea, all rainfall inputs were removed from the weather files fed into AquaCrop-OSPy; the DRL agent was retrained, and the fixed-heuristic benchmark was re-optimized accordingly.

Table 4 reports the profit gaps between the DRL agent and the optimized soil-moisture thresholds under conditions with rainfall retained or eliminated. In the absence of rainfall, the optimized fixed thresholds reached $300.6 ha−1 during validation, a decline from Section 3.1 levels owing to the elevated irrigation volumes required to meet crop needs each year. Meanwhile, the DRL agent, which applied discrete depth actions at 3-day intervals, secured $307 ha−1. Thus, stripping away a core source of fluctuation and unpredictability in irrigation planning changed DRL's relative benefit over fixed SMT heuristics from a shortfall of $1.27 ha−1 to a surplus of $6.38 ha−1 — equivalent to a 2.1 % profit rise above the optimized heuristic.

Table 4. Increase in seasonal profits compared to optimized soil-moisture thresholds with and without rainfall.

Weather data	Percentage increase in validation set profits	Increase in validation set profits $ ha^{-1}	Increase in training set profits $ ha^{-1}
With rainfall	– 0.2 %	−1.27	5.87
Without rainfall	+ 2.1 %	6.38	4.07

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Water use restrictions

Up to now, both DRL and heuristic approaches have run without irrigation limits, allowing the agent to apply whatever volumes it deems necessary. Yet this premise fails to align with conditions in our Nebraska case study area, where annual or multi-year constraints govern irrigation water withdrawals. Because extra water yields the largest marginal returns under shortage conditions, we hypothesise that advanced techniques such as DRL could deliver greater benefits under limited supply. To gauge how water restrictions affect DRL's usefulness, the established framework was rerun with caps set at 75, 100, and 125 mm ha−1 for every simulation year. Once a seasonal cap is hit, all further irrigation ceases. **Table 5** highlights the resulting profit differences between the DRL agent and optimized soil-moisture thresholds at these restriction levels. While gains remained modest, DRL delivered greater value in restricted-water settings than optimized heuristics. Applying the tightest limit examined (75 mm ha−1) lifted DRL's added value from -$1.3 ha−1 to +$12.2 ha−1, translating into a 5.6 % profit improvement over the optimized heuristic.

Table 5. The effect of water-use restrictions on validation-set profits and the added value of DRL relative to optimized heuristics.

Irrigation cap mm ha^{-1}	Percentage increase in validation set profits	Increase in validation set profits $ ha^{-1}	DRL validation set profits $ ha^{-1}
75	5.6 %	12.2	228
100	2.5 %	7.1	286.5
125	1.3 %	4.3	332
No cap	– 0.2 %	−1.3	520.1

While optimized heuristics continue to excel in irrigation scheduling, folding expanding farmer data streams into them would necessitate numerous elaborate, interdependent rules that prove hard to craft and fine-tune. To overcome such hurdles, artificial intelligence methods, including Deep Reinforcement Learning (DRL) [8], have emerged as promising tools for dynamically refining choices across the season using any available inputs. Recent work confirms that deep reinforcement learning (DRL) effectively generates irrigation policies within crop simulation platforms [13-15].

This paper examined the additional benefits DRL offers for irrigation scheduling compared with optimized soil-moisture thresholds. Early testing showed that letting the DRL agent pick from discrete irrigation depth options (i.e., 0, 2.5, … 22.5, 25 mm) paired with 3-day decision cycles maximized profits in validation years. Both shorter (e.g., daily) and longer (e.g., 5- or 7-day) intervals produced modestly lower overall crop returns. We connect this pattern to the increased complexity of daily decisions, which enlarges the space for poor choices (such as watering right before unexpected rain). In contrast, extended gaps between decisions increase the likelihood of water stress building up between agent updates.

Although this setup outperformed other DRL variants, it still fell short of the returns from a standard optimized heuristic across the same validation years. Eliminating rainfall from the datasets reduced differences in weather series between the training and validation sets, enabling the DRL agent to improve validation profits by 2.1% over optimized heuristics. In addition, imposing strict irrigation caps (e.g., 75 mm ha−1) enabled the DRL agent to boost validation profits by 5.6% relative to the optimized heuristics.

Findings indicate that, given ordinary climate fluctuations in the study region and unlimited water, optimized soil-moisture thresholds delivered higher average validation-year profits than the DRL agent. This broadly matches observations by Alibabaei *et al.* [13], who found that a basic, non-optimized heuristic outperformed their DRL agent in one of two test years. The modest validation profit gap in our work (–$1.3 ha−1) between DRL and the optimized heuristic should not be interpreted as a fixed quantitative benchmark, as different validation periods could alter the result. Reporting choices — including training stoppage timing, validation profit averaging methods, and repetition counts — may also cause minor

© 2024 The Author(s). | Published by Agro Intelligence Research Press

fluctuations in the gap. Still, the core observation that DRL agents struggle to beat well-tuned thresholds is expected to persist, an aspect often overlooked in prior DRL research [13-15]. Given how straightforward and sensible fixed soil-moisture heuristics are, their strong performance relative to DRL makes them a practical starting point for productivity gains that growers are more likely to implement.

The rise in DRL benefits under tightening water caps supplies useful direction for deploying the technology in actual farming operations. As water becomes scarcer and demand intensifies [37, 38], regulators and officials may tighten total usage limits, a step already taken in various locales [27, 36]. Results here suggest that DRL can support a more efficient distribution of scarce irrigation supplies as constraints grow stricter and more unpredictable, helping producers mitigate the consequences of mounting water limitations. Furthermore, stronger DRL performance without rainfall suggests its suitability for dry climates with scant precipitation, where annual variations in irrigation demand and timing remain relatively stable.

Even though various action sets, decision frequencies, water-use restrictions, and aspects of climate variability were tested, restrictions on available computing power limited the present study to examining only certain combinations of these elements. As one example, the effects of irrigation decision intervals were analyzed solely for discrete depth actions, but upcoming efforts should probe interactions among intervals and other scheduling choices or limitations. Beyond that, other significant design factors — including alternative DRL algorithms, hyperparameter values, neural network designs, and sets of input variables — received no comparison. To illustrate, adopting larger neural networks or more input variables tends to yield more intricate irrigation tactics, potentially yielding higher profits, but at the cost of longer training periods and a higher likelihood of overfitting. Examination of how these and further DRL design decisions influence overall DRL benefits constitutes an essential topic for subsequent research.

Implementing DRL within irrigation scheduling, as carried out here, creates multiple difficulties for the technique that are less severe in conventional domains where DRL has traditionally operated. Should real-world deployment occur, the agent must generate choices spanning several months or possibly years before encountering any reward feedback (such as profits computed after harvest). The underlying reason lies in the delayed visibility of the impacts of any particular irrigation action (e.g., effects on crop output or profit), since the agent perceives an outcome and its associated reward solely at harvest time, based on the collective influence of many irrigation choices made over the full growing period. Crop-simulation platforms partially resolve this by permitting complete seasons to run in just seconds. Nevertheless, eventual real-world trials — despite their high costs and extended duration — will still be necessary to verify whether strategies refined in models operate successfully in actual conditions. The process of transferring knowledge from simulations to practical systems continues to draw interest in robotics research [39]. Yet the requirement of months to assess even one irrigation approach remains distinctive to this application. An additional core difficulty in employing DRL for irrigation is the virtually unlimited variety of environmental conditions (e.g., differing crops, soils, climates, or management boundaries) that significantly influence optimal decisions. Addressing this may necessitate dedicated DRL agents for each setup — an option that could be computationally intensive or time-consuming, akin to challenges with optimized heuristics — or the creation of a single universal agent for every setup, which risks excessive complexity in acquiring strong policies. Moreover, adoption by growers requires clear evidence of benefits tailored to their specific setup; hence, field-based validation across multiple configurations could become essential before producers embrace this novel scheduling technique. As a result, the pursuit of DRL irrigation strategies that function independently of particular configurations emerges as a vital focus for continued research.

Similarly, the irrigation scheduling scenario detailed in this study, along with the bulk of prior publications [4, 5, 13-15, 40], focuses on irrigation management for a single uniform field spanning multiple seasons. Although straightforward to model, this framing tends to favor optimized heuristic methods (as optimization procedures can readily determine parameter sets that maximize average profits for that single field across N seasons). Yet heuristics exhibit limited scalability when complexity rises. For instance, if seasonal water limits vary partway through the season, fixed heuristics would either remain suboptimal or require reoptimization with each change. Additional scenarios involve multiple fields or crops, necessitating prioritization decisions for allocating scarce water resources. One more illustration drawn from our study

© 2024 The Author(s). | Published by Agro Intelligence Research Press

region concerns a 5-year water allocation framework that compels decisions that account for amounts reserved for subsequent years. Within such more elaborate — and truer-to-life — modeling contexts, DRL has the potential to deliver greater benefits.

Conclusion

Given the escalating global water scarcity, investigators are seeking creative approaches to boost agricultural water productivity. Deep Reinforcement Learning (DRL) is among the techniques under exploration and has recently achieved notable results in complex decision-making environments. This paper examines the benefits DRL offers for irrigation scheduling relative to optimized heuristics, drawing on a case study of irrigated maize cultivation in southwest Nebraska. Development of a structure that enables DRL agents to engage with the AquaCrop-OSPy crop-simulation model enabled training and assessment across distinct climate datasets. In conditions of typical climate variability paired with unrestricted irrigation, DRL did not increase profits during validation years above those achieved with optimized soil-moisture thresholds. Removal of rainfall, which represents the dominant driver of climate variability, enabled DRL to exceed the results of soil-moisture thresholds. Introduction of binding irrigation caps likewise allowed DRL to secure higher validation-year profits than optimized thresholds. Overall, the research emphasizes DRL as a promising approach for improving irrigation scheduling. Yet, its superiority over optimized heuristics remains to be fully confirmed across all situations, particularly given greater complexity and data demands. In contrast, optimized heuristics — thanks to their relative straightforwardness — merit consideration as a worthwhile initial measure to elevate water productivity before substantial commitments to AI-based techniques.

Acknowledgements

The work contained in this article was funded by the National Environmental Research Council's Understanding the Earth, Atmosphere, and Ocean Doctoral Training Programme, Grant NE/L002469/1. To support research into these areas of future work, we created *Aquacrop-Gym*, an open-source Python library for training DRL agents and comparing irrigation strategies within AquaCrop-OSPy. Training materials have been developed to guide users in training their own DRL agents and replicating the analysis presented in this work. This library aims to be an actively developed framework that future researchers can use to test new irrigation strategies and compare them with the current best approaches. This framework will accelerate progress in optimized and intelligent irrigation scheduling research and the rate at which this research can be converted to production. All code used in the analysis is available in a public GitHub repository (https://github.com/aquacropos/aquacrop-gym), allowing anyone to duplicate the analysis.

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 03 Apr 2023 Revised: 08 Jun 2023 Accepted: 23 Jun 2023
Published online: 25 November 2024

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

1. Fereres E, Soriano MA. Deficit irrigation for reducing agricultural water use. J Exp Bot. 2006;58(2):147-59.

2. Howell TA. Enhancing water use efficiency in irrigated agriculture. Agron J. 2001;93(2):281-9. https://doi.org/10.2134/AGRONJ2001.932281X.

3. USDA-NASS. 2018 irrigation and water management survey. 2018;3.

4. Cai X, Hejazi MI, Wang D. Value of probabilistic weather forecasts: Assessment by real-time optimization of irrigation scheduling. J Water Resour Plan Manag. 2011;137(5):391-403. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000126.

5. Jamal A, Linker R, Housh M. Optimal irrigation with perfect weekly forecasts versus imperfect seasonal forecasts. J Water Resour Plan Manag. 2019;145(5):06019003. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001066.

6. Linker R. Stochastic model-based optimization of irrigation scheduling. Agric Water Manag. 2021;243:106480. https://doi.org/10.1016/j.agwat.2020.106480.

7. Wang D, Cai X. Irrigation scheduling: Role of weather forecasting and farmers' behavior. J Water Resour Plan Manag. 2009;135(5):364-72. https://doi.org/10.1061/(ASCE)0733-9496(2009)135:5(364).

8. Sutton RS, Barto AG. Introduction to reinforcement learning. Cambridge (MA): MIT Press; 1998;135.

9. Badia AP, Piot B, Kapturowski S, Sprechmann P, Vitvitskyi A, Guo D, et al. Agent57: Outperforming the Atari human benchmark. Proc Int Conf Mach Learn. 2020:484-94. https://doi.org/10.48550/arXiv.2003.13350.

10. Berner C, Brockman G, Chan B, Cheung V, Dębiak P, Dennison C, et al. Dota 2 with large scale deep reinforcement learning. 2019.

11. Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. Science. 2018;362(6419):1140-4. https://doi.org/10.1126/science.aar6404.

12. Vinyals O, Babuschkin I, Czarnecki WM, Mathieu M, Dudzik A, Chung J, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature. 2019;575(7782):350-4. https://doi.org/10.1038/s41586-019-1724-z.

13. Alibabaei K, Gaspar PD, Assunção E, Alirezazadeh S, Lima TM. Irrigation optimization with a deep reinforcement learning model: Case study on a site in Portugal. Agric Water Manag. 2022;263:107480. https://doi.org/10.1016/j.agwat.2022.107480.

14. Chen M, Cui Y, Wang X, Xie H, Liu F, Luo T, et al. A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. Agric Water Manag. 2021;250:106838.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

https://doi.org/10.1016/j.agwat.2021.106838.

15. Yang Y, Hu J, Porter D, Marek T, Heflin K, Kong H. Deep reinforcement learning-based irrigation scheduling. Trans ASABE. 2020;63(3):549-56. https://doi.org/10.13031/trans.13633.

16. Kelly TD, Foster T. AquaCrop-OSPy: Bridging the gap between research and practice in crop-water modeling. Agric Water Manag. 2021;254:106976. https://doi.org/10.1016/j.agwat.2021.106976.

17. Linker R, Ioslovich I, Sylaios G, Plauborg F, Battilani A. Optimal model-based deficit irrigation scheduling using AquaCrop: A simulation study with cotton, potato and tomato. Agric Water Manag. 2016;163:236-43. https://doi.org/10.1016/j.agwat.2015.09.011.

18. Linker R, Kisekka I. Model-based deficit irrigation of maize in Kansas. Trans ASABE. 2017;60(6):2011-22. https://doi.org/10.13031/trans.12341.

19. Schulman J, Wolski F, Dhariwal P, Radford A, OpenAI OK. Proximal policy optimization algorithms. 2017. https://doi.org/10.48550/arXiv.1707.06347.

20. Andrychowicz OA, Baker B, Chociej M, Józefowicz R, McGrew B, Pachocki J, et al. Learning dexterous in-hand manipulation. 2019;39(1):3-20. https://doi.org/10.1177/0278364919887447.

21. Rabault J, Kuchta M, Jensen A, Réglade U, Cerardi N. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. J Fluid Mech. 2019;865:281-302. https://doi.org/10.1017/JFM.2019.62.

22. Liang E, Liaw R, Moritz P, Nishihara R, Fox R, Goldberg K, et al. RLlib: Abstractions for distributed reinforcement learning. Proc Int Conf Mach Learn. 2018;7:4768-80. https://doi.org/10.48550/arXiv.1712.09381.

23. Steduto P, Hsiao TC, Raes D, Fereres E. AquaCrop—the FAO crop model to simulate yield response to water: I. Concepts and underlying principles. Agron J. 2009;101(3):426-37. https://doi.org/10.2134/agronj2008.0139s.

24. Elias M, Salman M, García-Vila M, Fereres E, Raes D, Steduto P. The AquaCrop model – Enhancing crop water productivity. FAO. 2021. https://doi.org/10.4060/cb7392en.

25. McGuire VL. Water-level and recoverable water in storage changes, High Plains aquifer, predevelopment to 2015 and 2013-15. U.S. Geol Surv Sci Investig Rep. 2017;2017-5040:14. https://doi.org/10.3133/sir20175040.

26. Scanlon BR, Faunt CC, Longuevergne L, Reedy RC, Alley WM, McGuire VL, et al. Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. Proc Natl Acad Sci U S A. 2012;109(24):9320-5. https://doi.org/10.1073/pnas.1200311109.

27. Young R, Foster T, Mieno T, Valocchi A, Brozović N. Hydrologic-economic trade-offs in groundwater allocation policy design. Water Resour Res. 2021;57(1):e2020WR027941. https://doi.org/10.1029/2020WR027941.

28. HPRCC. High Plains Regional Climate Center. 2016. https://hprcc.unl.edu/index.php.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

29. Semenov MA, Barrow EM. Use of a stochastic weather generator in the development of climate change scenarios. Clim Change. 1997;35(4):397-414. https://doi.org/10.1023/A:1005342632279.

30. Agarwal A, Babel MS, Maskey S. Analysis of future precipitation in the Koshi river basin, Nepal. J Hydrol. 2014;513:422-34. https://doi.org/10.1016/J.JHYDROL.2014.03.047.

31. Awal R, Bayabil HK, Fares A. Analysis of potential future climate and climate extremes in the Brazos Headwaters Basin, Texas. Water. 2016;8(12):603. https://doi.org/10.3390/W8120603.

32. Foster T, Brozović N, Butler AP. Analysis of the impacts of well yield and groundwater depth on irrigated agriculture. J Hydrol. 2015;523:86-96. https://doi.org/10.1016/j.jhydrol.2015.01.032.

33. Sandhu R, Irmak S. Performance of AquaCrop model in simulating maize growth, yield, and evapotranspiration under rainfed, limited and full irrigation. Agric Water Manag. 2019;223:105687. https://doi.org/10.1016/j.agwat.2019.105687.

34. USDA. National Agricultural Statistics Service—Charts and maps—Prices received: Corn prices received by month, US. 2019. https://www.nass.usda.gov/Charts_and_Maps/Agricultural_Prices/pricecn.php.

35. CropWatch. Nebraska crop budgets. 2019.

36. Berbel J, Expósito A, Gutiérrez-Martín C, Mateos L. Effects of the irrigation modernization in Spain 2002–2015. Water Resour Manag. 2019;33(5):1835-49. https://doi.org/10.1007/s11269-019-02215-w.

37. D'Odorico P, Chiarelli DD, Rosa L, Bini A, Zilberman D, Rulli MC. The global value of water in agriculture. Proc Natl Acad Sci U S A. 2020;117(36):21985-93. https://doi.org/10.1073/pnas.2005835117.

38. Rosa L, Chiarelli DD, Rulli MC, Dell'Angelo J, D'Odorico P. Global agricultural economic water scarcity. Sci Adv. 2020;6(18). https://doi.org/10.1126/sciadv.aaz6031.

39. Zhao W, Queralta JP, Westerlund T. Sim-to-real transfer in deep reinforcement learning for robotics: A survey. 2020 IEEE Symp Ser Comput Intell. 2020:737-44.

40. Linker R, Sylaios G, Tsakmakis I, Ramos T, Simionesei L, Plauborg F, et al. Sub-optimal model-based deficit irrigation scheduling with realistic weather forecasts. Irrig Sci. 2018;36(6):349-62.

© 2024 The Author(s). | Published by Agro Intelligence Research Press