Agricultural Water Management 250 (2021) 106838

Contents lists available at ScienceDirect

Agricultural Water Management

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

# A reinforcement learning approach to irrigation decision-making for rice using weather forecasts

Mengting Chen [a], Yuanlai Cui [a], Xiaonan Wang [b], Hengwang Xie [c], Fangping Liu [c], Tongyuan Luo [d], Shizong Zheng [d], Yufeng Luo [a,*]

[a] State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, Hubei 430072, China
[b] School of Computer Science and Engineering, Changshu Institute of Technology, Changshu, Jiangsu 215500, China
[c] Jiangxi Central Irrigation Experiment Station, Nanchang, Jiangxi 330201, China
[d] Department of Rural Water, Zhejiang Institute of Hydraulics and Estuary, Hangzhou, Zhejiang 310020, China

ARTICLE INFO

Handling Editor: J.E. Fernández

*Keywords:*
Reinforcement learning
Irrigation decision-making
Paddy rice irrigation
Weather forecasts

ABSTRACT

Improving efficiency with the use of rainfall is one of the effective ways to conserve water in agriculture. At present, weather forecasting can be used to potentially conserve irrigation water, but the risks of unnecessary irrigation and the yield loss due to the uncertainty of weather forecasts should be avoided. Thus, a deep Q-learning (DQN) irrigation decision-making strategy based on short-term weather forecasts was proposed to determine the optimal irrigation decision. The utility of the method is demonstrated for paddy rice grown in Nanchang, China. The short-term weather forecasts and observed meteorological data of the paddy rice growth period from 2012 to 2019 were collected from stations near Nanchang. Irrigation was decided for two irrigation decision-making strategies, namely, conventional irrigation (i.e., flooded irrigation commonly used by local farmers) and DQN irrigation, and their performance in water conservation was evaluated. The results showed that the daily rainfall forecasting performance was acceptable, with potential space for learning and exploitation. The DQN irrigation strategy had strong generalization ability after training and can be used to make irrigation decisions using weather forecasts. In our case, simulation results indicated that compared with conventional irrigation decisions, DQN irrigation took advantage of water conservation from unnecessary irrigation, resulting in irrigation water savings of 23 mm and reducing drainage by 21 mm and irrigation timing by 1.0 times on average, without significant yield reduction. The DQN irrigation strategy of learning from past irrigation experiences and the uncertainties in weather forecasts avoided the risks of imperfect weather forecasting.

## 1. Introduction

Rice is one of the major grain crops in China, comprising approximately 30 million hectares of the planting area (NBSPRC, 2019). As a water-loving crop, rice needs a large amount of fresh water for irrigation during its growth. Although the water resources in China are abundant, they are distributed unevenly in time and space and do not match perfectly with the rice irrigation demand. In addition, with rapidly increasing industrial and domestic water consumption, the amount of agricultural water resources has decreased, and water scarcity has become a threat to irrigated rice, driving calls for constant development of new methods of irrigation aimed at increasing the efficiency of allocated irrigation water (Mao, 2002).

An important attempt to achieve this goal is linked to the improvement in rainfall use efficiency. Using rainfall information in weather forecasts to make irrigation decisions can conserve irrigation water. In recent years, public weather forecasts (Mishra et al., 2013; Lorite et al., 2015; Luo et al., 2014; Yang et al., 2016; Traore et al., 2016) and numerical weather forecasts (Perera et al., 2014; Medina et al., 2018) have been applied to determine reference evapotranspiration ($ET_0$) in the development of irrigation schedules. However, these researches did not consider future rainfall, and a large difference was observed between the $ET_0$ and actual irrigation demand, making it difficult to put into practice. Therefore, certain studies incorporated the $ET_0$ and precipitation forecasts for irrigation demand predictions (Perera et al., 2016; Longo-Minnolo et al., 2020). The literature is rich in setting decision rules for determining the irrigated time and amount by using short-term weather forecasts (Cao et al., 2019; Cai et al., 2011;

* Corresponding author.
*E-mail address:* yfluo@whu.edu.cn (Y. Luo).

https://doi.org/10.1016/j.agwat.2021.106838
Received 31 May 2020; Received in revised form 17 February 2021; Accepted 26 February 2021
Available online 6 March 2021

Gowing and Ejieji, 2001; Linker and Sylaios, 2016; Bergez and Garcia, 2010). However, one major issue in previous studies is that the risks due to the uncertainties in weather forecasts (i.e., unnecessary irrigation applied to the field because of unexpected rainfall and potential yield reduction if irrigation is delayed) should be considered in making irrigation decisions using weather forecasts (Mondaca-Duarte et al., 2020). When considering the balance between crop yield and water saving, most studies normally focus on short-term risks by applying a strategy that delays irrigation or removes the expected rainfall from the irrigation water and simply simulates the loss of yield using the average historical data or does not give sufficient consideration to yield, thus failing to obtain the long-term effects of current irrigation decisions, which makes it difficult to obtain the global optimal solution for rice irrigation decision optimization.

The irrigation decision must determine the timing and amount of irrigation applied to the field, which can be formulated as a Markov decision process (i.e., the future state is only related to the current state) (Lee et al., 1991). Reinforcement learning (RL), a machine learning method that uses rewards obtained in the process of interaction with the environment to guide behavior and constantly learns a strategy to maximize the ultimate cumulative rewards, is suitable for decision-making process with the Markov property (Sutton and Barto, 2018). As a branch of RL, the deep Q-learning network (DQN) combines the perception ability of deep learning and its own decision-making ability (Mnih et al., 2013) and can supply new solutions to cognitive decision-making problems in complex states. As a new method in computer science, RL has been widely used in prediction (Görges, 2017), intelligent control (Chen and Su, 2018), decision support (Shen et al., 2017; Wu et al., 2018) and many other fields, but less attention has been given to irrigation management. Bergez et al. (2001) compared the dynamic programming (DP) and RL methods for the identification of optimal initial irrigation strategies with limited irrigation water for maize crops coupled with a stochastic random weather generator. The results showed that RL performed better than DP when only a small number of simulation runs were available. Irukula (2015) optimized the water consumption as much as possible without affecting the crop yield by using the RL algorithm on a maize crop simulation model according to daily conditions such as temperature, solar radiation, rainfall and soil water content. A decrease in water consumption of almost 40% was achieved compared with the constant irrigation method and without any significant decrease in yield and leaf area index. Bu and Wang (2019) designed a smart agriculture Internet of Things (IoT) system based on deep reinforcement learning and presented several representative deep reinforcement learning models. The reviewed water resources studies used RL as a heuristic decision optimization method focusing on use of real-time data collected from the farmland or simulator without considering the future rainfall information contained in a weather forecast. With the improvement in weather forecast quality, use of the reinforcement learning method can excavate hidden information and improve rainfall use efficiency.

In this paper, a RL approach based on the DQN was proposed for irrigation decision-making using weather forecasting. This approach is known as the DQN irrigation decision-making strategy (DQN irrigation strategy), and the paddy rice of Nanchang, a typical rice planting area of China, was chosen as a case study for validation. The rainfall forecast is the main variable of the DQN irrigation strategy, and its accuracy has a great influence on the strategy performance. Therefore, we first evaluated the forecasting performance of daily rainfall for lead times of 1–7 days, and a simple water balance model and a crop water production function, used to simulate the real state of the paddy rice and field, were integrated with the DQN algorithm for irrigation decision-making.

The objectives of this study are (1) to evaluate the quality of the short-term public weather forecasts for different types of paddy rice in different periods in Nanchang, (2) to propose a reinforcement learning strategy based on DQN for irrigation decision-making and to analyze its training and generalization results, and (3) to verify the effect of this proposed strategy in conventional flooded irrigation in terms of water conservation and drainage reduction and to preliminarily illustrate how the proposed method saves water by considering the rainfall forecast.

## 2. Materials and methods

### 2.1. Study area and data

Paddy rice is mainly distributed in southern China and includes double season rice (early rice and late rice) and single season rice (middle rice). In this study, data from Nanchang in southern China on double season rice and single season rice were used to verify the proposed method (Fig. 1).

To increase the number of samples in the training set, the daily observed meteorological data of the paddy rice growing period at three stations near Nanchang from 2012 to 2019 (Table 1) were collected from the China Meteorological Data Sharing Service System (http://data.cma.gov.cn), including daily minimum and maximum temperature, average temperature, average wind speed, sunshine duration, mean relative humidity and precipitation. The weather type of the public weather forecast data that contained rainfall information with a 7-day lead time for the same period was collected from Weather China (http://www.weather.com.cn). Noting that the observed and forecast weather data for certain periods were incomplete, the data in 2012 for the early-rice growth period and the data in 2014 for the middle-rice and late-rice growth period were not included in this study. The data from 2012 to 2018 were set as the training set, and the data of 2019 were set as the verification set.

Information on agricultural management from 2012 to 2019 in Nanchang was obtained from field experiments and included phenological data, crop coefficients and irrigation scheduling. The growth period of paddy rice was divided into six stages: the returning green (RG), early tillering (ET), late tillering (LT), jointing-booting (JB), heading-flowering (HF), milk-ripe (MI) and yellow-ripe (YR) stages. All types of rice were assumed to apply the conventional flooded irrigation regime according to the experience in irrigation from local farmers, and the crop coefficients and the criteria of field water depth in each growth stage of paddy rice are shown in Table 2.

### 2.2. Evaluation of weather forecasting

Previous studies reported that uncertainties exist in the application of weather forecasting to irrigation or evapotranspiration forecasting (Luo et al., 2014; Yang et al., 2016; Cao et al., 2019). To evaluate the performance of weather forecasting, the threat score (TS), missing alarm rate (MAR) and false alarm rate (FAR) were used (Donaldson et al., 1975; Wilks, 2006):

$$TS = \frac{NA}{NA + NB + NC} \tag{1}$$

$$MAR = \frac{NC}{NA + NC} \tag{2}$$

$$FAR = \frac{NB}{NA + NB} \tag{3}$$

where $NA$ is the number of hits (i.e., the days that both forecasts and observations fall in the prescribed threshold ranges), $NB$ is the number of false alarms (i.e., the days that the forecast falls in the threshold ranges but the observation does not), and $NC$ is the number of misses (i.e., the days that the forecast falls outside the threshold range and the observation falls inside the threshold range). According to the cumulative rainfall over 24 h, the weather type was divided into 5 categories, and the threshold values for no rain (NR), light rain (LR), moderate rain (MR), heavy rain (HR), storm or above (ST) were set as [0, 0.1), [0.1, 10), [10, 25), [25, 50), and [50,) mm, respectively. It is worth noting

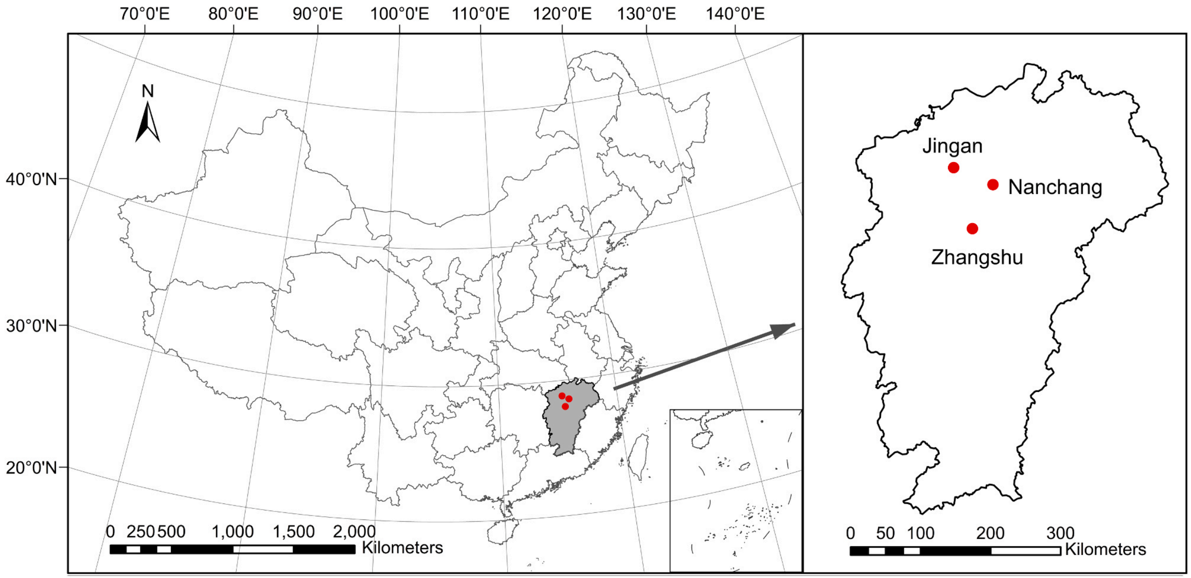


**Fig. 1.** Location of the study area.

**Table 1**
Basic information for stations near Nanchang used in this study.

| Name | Soil type | Latitude | Longitude | Altitude (m) |
|---|---|---|---|---|
| Nanchang | Clay loam | 28°36'N | 115°55'E | 46.9 |
| Zhangshu | Clay loam | 28°04'N | 115°33'E | 30.4 |
| Jingan | Clay loam | 28°52'N | 115°22'E | 78.9 |

**Table 2**
Duration, criteria of field water depth and crop coefficients at each growth stage over rice growth period for early rice, middle rice, and late rice.

| Growth period | RG | ET | LT | JB | HF | MI | YR |
|---|---|---|---|---|---|---|---|
| **A. Early rice** | | | | | | | |
| Start date | 25/04 | 02/05 | 16/05 | 31/05 | 16/06 | 25/06 | 05/07 |
| Ending date | 01/05 | 15/05 | 30/05 | 15/06 | 24/06 | 04/07 | 14/07 |
| $h_{min}$-$h_{max}$-$H_p$ (mm) | 10-30-60 | 20-50-60 | 20-50-60 | 20-50-60 | 20-50-60 | 20-50-60 | Dry |
| Crop coefficients | 1.01 | 1.13 | 1.1 | 1.15 | 1.27 | 1.22 | 0.95 |
| **B. Middle rice** | | | | | | | |
| Start date | 15/06 | 25/06 | 10/07 | 24/07 | 17/08 | 02/09 | 17/09 |
| Ending date | 24/06 | 09/07 | 23/07 | 16/08 | 01/09 | 16/09 | 04/10 |
| $h_{min}$-$h_{max}$-$H_p$ (mm) | 20-40-60 | 20-50-60 | 20-50-60 | 20-50-60 | 20-50-60 | 0-30-50 | Dry |
| Crop coefficients | 0.92 | 1.15 | 1.17 | 1.2 | 1.25 | 1.15 | 1.12 |
| **C. Late rice** | | | | | | | |
| Start date | 22/07 | 29/07 | 12/08 | 26/08 | 16/09 | 25/09 | 09/10 |
| Ending date | 28/07 | 11/08 | 25/08 | 15/09 | 24/09 | 08/10 | 22/10 |
| $h_{min}$-$h_{max}$-$H_p$ (mm) | 10-30-60 | 20-50-60 | 20-50-60 | 20-50-60 | 20-50-60 | 20-50-60 | Dry |
| Crop coefficients | 0.94 | 1.23 | 1.21 | 1.37 | 1.63 | 1.57 | 1.12 |

$h_{min}$ is the lower limit of water depth, $h_{max}$ is the upper limit of water depth and $H_p$ is the maximum allowable water depth after rainfall according to the conventional irrigation strategy.

that the type of storm or above in this case was composed of storms, heavy storms and severe storms due to the low probability of individual occurrences.

### 2.3. Determination of crop evapotranspiration

The actual evapotranspiration is estimated by (Allen et al., 1998):

$$ET_c = K_c \cdot K_s \cdot ET_0 \tag{4}$$

where $K_c$ is the single crop coefficient; $K_s$ is the water stress coefficient, for soil water limiting conditions $K_s < 1$, otherwise, $K_s = 1$; and $ET_0$ is the reference evapotranspiration, mm/day.

Reference evapotranspiration is a key parameter to compute crop water requirement, determine irrigation schedule and conduct reasonable farmland water management. At present, there are many calculation methods. For example, Hargreaves-Samani (Hargreaves and Samani, 1985), Blaney-Criddle (Blaney and Criddle, 1962), McCloud (McCloud, 1955) and other temperature-based methods, Irmak-Allen (Irmak et al., 2003) and other Empirical methods, Priestley-Taylor (Priestley and Taylor, 1972) and other radiation-based method, FAO24 Penman (Doorenbos and Pruitt, 1975), FAO56 Penman-Manteith (Allen et al., 1998) and other synthesis methods. As the FAO56 Penman-Manteith method has a high precision and a solid application foundation, and all the meteorological variables needed can be obtained. Therefore, in this case, FAO56 Penman-Manteith method was selected for reference evapotranspiration and it can be expressed as:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma[900/(T+273)]u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \tag{5}$$

where $R_n$ is the net radiation at the crop surface, MJ m$^{-2}$ day$^{-1}$; $G$ is the soil heat flux density, MJ m$^{-2}$ day$^{-1}$; $T$ is the air temperature at a height of 2 m, °C; $u_2$ is the wind speed at a height of 2 m, m s$^{-1}$; $e_s$ is the vapor pressure of the air at saturation, kPa; $e_a$ is the actual vapor pressure, kPa; $\Delta$ is the slope of the vapor pressure curve, kPa °C$^{-1}$; and $\gamma$ is the psychrometric constant, kPa °C$^{-1}$.

The water stress coefficient is given by (Allen et al., 1998):

$$K_s = \begin{cases} 1 & 0 \le D_r \le RAW \\ \dfrac{TAW - D_r}{TAW - RAW} & D_r > RAW \end{cases} \tag{6}$$

where $D_r$ is the root zone depletion expressing water content in the root zone (i.e., water shortage relative to field capacity); $D_r = 0$ at field capacity, mm; $TAW$ is the total available soil water in the zone, mm; and $RAW$ is the readily available soil water in the root zone, mm.

$TAW$ and RAW are calculated by (Allen et al., 1998):

$$TAW = 1000(\theta_{FC} - \theta_{WP})Z_r \tag{7}$$

$$RAW = p\,TAW \tag{8}$$

where $\theta_{FC}$ is the water content at field capacity, $m^3m^{-3}$; $\theta_{WP}$ is the water content at the wilting point, $m^3m^{-3}$; $Z_r$ is the rooting depth and is set as 0.5, m; p is the average fraction of TAW that is depleted from the root zone before moisture stress and is set as 0.2; and $\theta_{FC}$-$\theta_{WP}$ is estimated as responding to specific soil type (e.g., 0.16 at clay loam).

### 2.4. Crop water production function

When the weather forecast information is considered in irrigation scheduling, irrigation might be delayed to wait for rainfall in the forecast, resulting in water deficit and leading to yield reduction. To evaluate the effect of water deficit on crop yield, the crop water production function was used to establish the functional relationship between water deficit and yield at a specific growth stage. According to previous studies, paddy rice in southern China usually adopts the Jenson model (Jensen, 1968) to quantify the effect. The model is expressed as:

$$\frac{Y_a}{Y_m} = \prod_{i=1}^{n}\left(\frac{ET_a}{ET_m}\right)_i^{\lambda_i} \tag{9}$$

where $Y_a$ is the actual crop yield; $Y_m$ is the maximum crop yield; $ET_a$ is the actual crop evapotranspiration; $ET_m$ is the maximum crop evapotranspiration; $n$ is the number of growth stages; $i$ is the ordinal number of the growth stage, $i = 1,2,\ldots,n$; and $\lambda_i$ is the water stress sensitivity index at growth stage $i$.

In this study, the values of $\lambda_i$ for the three rice types were derived from the literature, as shown in Table 3 (Luo, 2003).

### 2.5. DQN algorithm for irrigation decision

A Markov property is present in the process of irrigation decision-making. The irrigation decision is made based on soil water content and crop condition in the observation stage. It is obvious that the state of the soil system and the crop condition at the current moment is only related to the state of the soil system, climatic conditions, crop conditions and the decision adopted (the amount of irrigation water) at the previous moment. Thus, RL was introduced to solve the irrigation decision problem with the Markov property.

#### 2.5.1. Environment of RL

RL refers to a method of machine learning in which an agent interacts with an environment through a sequence of observations, actions and rewards. The goal of the agent is to find an optimal policy that selects actions in a fashion that maximizes cumulative future rewards. The environment of reinforcement can be expressed as:

$$E = \{S, A, P, R\} \tag{10}$$

**Table 3**
Water stress sensitivity index at each growth stage over rice growth period for early rice, middle rice, and late rice.

| Rice types | TL | JB | HF | MI |
|---|---|---|---|---|
| Early rice | 0.1130 | 0.3999 | 0.6968 | 0.2393 |
| Middle rice | 0.1818 | 0.4524 | 0.6393 | 0.1213 |
| Late rice | 0.2219 | 0.4919 | 0.2339 | 0.0675 |

where $S$ is the state space; $A$ is the action space; $P$ is the transition probability; and $R$ is the reward function.

State space consists of the environmental parameters in the decision-making cycle of the area to be irrigated (i.e., a certain day in the growing period of crops), and it can be expressed as:

$$s_t = \left(P_t, h_t, h_{\min}, h_{\max}, H_p\right) \tag{11}$$

where $P_t$ is the forecast rainfall sequence for the next 7 days on day $t$, mm; $h_t$ is the water depth on day $t$, mm; $h_{\min}$ is the lower limit of water depth, mm; $h_{\max}$ is the upper limit of water depth, mm; $H_p$ is the maximum allowable water depth after rainfall and the value of $h_{\min}$, $h_{\max}$ and $H_p$ are updated in accordance with the growth period of day $t$ (see Table 2).

Action space is the option of irrigation decision-making. In this case, there are 3 possible actions that represent the proportion of the irrigation quota (irrigate until $h_{\max}$) to be supplied to the field on day t. Action 0 supplies 0%, action 1 supplies 50%, and action 2 supplies 100%.

The transition probability is the probability that makes the environment move from the current state to another state after the irrigation decision is executed, including evapotranspiration update, forecast weather data update, and change in depth of water. The change of water depth is expressed as:

$$m_t = \begin{cases} h_{\max} - h_t & a_t = 0 \\ \frac{1}{2}\left(h_{\max} - h_t\right) & a_t = 1 \\ h_t & a_t = 2 \end{cases} \tag{12}$$

$$h_{t+1} = h_t + m_t \tag{13}$$

where $m_t$ is the amount of irrigation water on day $t$, mm; $h_{t+1}$ is the depth of water on day $t+1$, mm.

After executing action $a_t$ at state $s_t$, actual meteorological and field conditions are observed, and the environmental state parameters, including actual evapotranspiration and water depth, should be modified. The modifier formula of water depth can be expressed as:

$$h_{t+1}^{*} = \begin{cases} h_{t+1} + P_t^{1*} - ET_{ct}{}^{*} - f_t & h_{t+1} + P_t^{1*} - ET_{ct}{}^{*} - f_t \le H_P \\ H_P & h_{t+1} + P_t^{1*} - ET_{ct}{}^{*} - f_t > H_P \end{cases} \tag{14}$$

where $h_{t+1}^{*}$ is the actual depth of water on day $t+1$, mm and $P_t^{1*}$, $ET_{ct}^{*}$ and $f_t$ are the actual rainfall, evapotranspiration and deep percolation on day $t$, respectively, mm. In this study, the daily deep percolation was set as a constant value (the value for clay loam was set as 2 mm) before the field water was depleted; otherwise, the value was set as zero.

The reward function is the reward feedback when the environment moves from the current state to another state after the irrigation decision is executed. In this case, the reward function consists of multiplying three subfunctions as follows:

$$r_t = r_{0,t} \times r_{1,t} \times r_{2,t} \tag{15}$$

where $r_{0,t}$ represents the basic reward, $r_{1,t}$ represents the rainfall utilization reward and $r_{2,t}$ represents the yield reward.

$r_{1,t}$ is expressed as:

$$r_{1,t} = \frac{ET_{c,t\sim t+7} + f_{t\sim t+7}}{P_{t\sim t+7} + m_t + \left(h_t - h_{t+7}\right)} \tag{16}$$

where $ET_{c,t\sim t+7}$, $f_{t\sim t+7}$ and $P_{t\sim t+7}$, are the cumulative amount of actual evapotranspiration, deep percolation and rainfall in the next 7 days from day $t$ to day $t+7$ after executing action $a_t$. For example, a 7-day observation was conducted from day $t$. If taking action 2, the water was filled to the upper limit of water depth (i.e., $m_t = h_{\max}$-$h_t$), and drainage occurred if the depth of the water layer reached the maximum allowable water depth after the rainfall ($H_p$); otherwise, no other operation was performed. The total amount of all indicators within 7 days was recorded, where $h_{t+7}$ is the water depth at the end of day $t+7$.

The value of $r_{1,t}$ represents the impact of irrigation decisions on rainfall utilization. If no drainage event occurs, $r_{1,t}$ is equal to 1 which means all rainfall is used for field consumption and there is no waste of water caused by rainfall after irrigation. Otherwise, $r_{1,t}$ is less than 1 which means that drainage wastes part of the rainfall and reduces rainfall utilization.

$r_{2,t}$ is expressed as:

$$r_{2,t} = \left(\frac{ET_{c,t\sim t+7}}{ET_{m,t\sim t+7}}\right)^{\lambda} \tag{17}$$

where $ET_{m,t\sim t+7}$ represents the cumulative amount of actual evapotranspiration from non-stressed treatment in the next 7 days from day $t$ to day $t+7$ after executing action at. λ is water stress sensitivity index in accordance with growth period of day t (see Table 3).

The value of $r_{2,t}$ represents the influence of irrigation decisions on yields. If $r_{2,t}$ is less than 1 after a specific decision is made (e.g., action $a_2$), it indicates that crops will suffer from water deficit in the next few days, resulting in the yield loss. Otherwise, this irrigation decision has no adverse effect on crop yield.

$r_{0,t}$ is expressed as:

$$r_{0,t} = \begin{cases} 1 & a_t = 0 \quad h_t < h_{\min} \\ & a_t = 1\text{or}2 \quad h_t \geq h_{\min} \\ 9 & a_t = 1 \quad h_t < h_{\min} \\ 10 & a_t = 0 \quad h_t \geq h_{\min} \\ & a_t = 2 \quad h_t < h_{\min} \end{cases} \tag{18}$$

The value of $r_{0,t}$ represents whether the irrigation decision conforms to the irrigation baseline according to the depth of traditional irrigation criteria. Eq. (18) indicates that if the basic principle of traditional irrigation is violated (e.g., irrigating before reaching the lower limit of water depth), there will be a small reward (i.e., 1). Moreover, we believe that when the water depth in the field reaches the lower limit of water depth and irrigation is required, taking action a1 should receive a smaller reward than action a2. This difference of 1 is set to compensate for rising costs since action a1 (i.e., small irrigation quota) will increase the irrigation times to some extent. It is worthwhile noting that $r_{0,t}$ will only be determined according to Eq. (18) if there is no drainage in the next 7 days after irrigation, namely $r_1 = 1$, otherwise, $r_0 = 1$. Because when drainage events occur, we want the irrigation decision should be based on the balance of rainfall utilization reward and yield reward without being constrained by irrigation baseline.

To evaluate the strategy, the action-value function (Q) was used to represent the accumulated value of the long-term expected reward (Sutton and Barto, 2018). Given policy $\pi$ for irrigation decision-making, the $Q$ for selecting action $a$ under input sate $s$ can be defined as:

$$Q^{\pi}\left(s, a\right) = E_{\pi}\left\{\sum_{i=0}^{+\infty}\gamma^{i} r_{t+i} | s_t = s, a_t = a\right\} \tag{19}$$

where $\gamma$ is the discount factor and is set as 0.2 in this case; $i$ is the number of subsequent execution steps; and ($s$, $a$) is the joint vector of state and action.

The optimal $Q$ function $Q^*$ can be defined as:

$$Q^{*}\left(s; a\right) = \max_{\pi} Q^{\pi}\left(s; a\right) \tag{20}$$

The optimal stationary policy is:

$$\pi^{*}(s) = \arg\max_{a\in A} Q^{*}\left(s, a\right) \tag{21}$$

### 2.5.2. Deep Q-learning network algorithm

When reinforcement model is complete known, that is, every part of Eq. (17) is known, reinforcement learning problems can be transformed into optimal control problems (i.e., model-based reinforcement problem). The model-based reinforcement problems (i.e., where the transition probability set is given) can be solved efficiently by dynamic programming techniques. However, the transition probability set of our irrigation model is not completely given (i.e., uncertainties exist related to rainfall in the future). On the other hand, the traditional reinforcement learning approach is usually designed for discrete cases while the water depth and precipitation are not distributed discretely. In other words, certain issues are associated with the continuous state space. To address the continuous state space and enhance the learning speed, we leverage deep neural network (DNN) technology to realize a deep Q-learning network algorithm for irrigation decision-making.

According to basic Q-learning algorithm, Q should be updated by the following equation known as off-policy temporal-difference learning (Watkins and Dayan, 1992):

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left[r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)\right] \tag{22}$$

where α is the learning rate and is set as 0.001, 0.0003 and 0.0002 for early rice, middle rice and late rice, respectively.

To deal with the continuous state space, a neural network function approximator with weight $\theta$ was used as a Q-network (Mnih et al., 2015). This approximator can be trained by minimizing a sequence of loss functions $L_k$ ($\theta_k$) that changes at each iteration $k$, and the loss function is defined as:

$$L_k(\theta_k) = E\left[\left(r + \gamma \max_{a_{t+1}} Q\left(s_{t+1}, a_{t+1}; \theta_k^-\right) - Q(s_t, a_t; \theta_k)\right)^2\right] \tag{23}$$

where $r + \gamma maxQ(st+1, at+1; \theta_k^-)$ is the target of iteration $k$, $\theta_k^-$ is the parameter used to compute the target network, and $\theta_k$ is the parameter of the Q-network of iteration $k$. The target network parameter $\theta_k^-$ is only updated with the Q-network parameter $\theta_k$ every C steps.

Thus, the update of action-value function is adjusted by updating the Q-network parameter as:

$$\theta_{k+1} = \theta_k + \alpha \nabla_{\theta_k} L_k(\theta_k) \tag{24}$$

The partial in the $\theta_k$ direction is defined as:

$$\nabla_{\theta_k} L_k(\theta_k) = E\left[\left(r + \gamma \max_{a_{t+1}} Q\left(s_{t+1}, a_{t+1}; \theta_k^-\right) - Q(s_t, a_t; \theta_k)\right) \nabla_{\theta_k} Q(s_t, a_t; \theta_k)\right] \tag{25}$$

In training the neural network, to avoid a sub-optimal policy in which the agent does not have the opportunity to sample state actions with higher returns, an ε-greedy policy was used that selects the action that maximizes $Q(s_t, a_t; \theta_k)$ with probability 1-$\varepsilon$ or a random action with probability $\varepsilon$. Additionally, to ensure the convergence and stability of neural network training, DQN introduced the so-called "experience replay" (Mnih et al., 2015) to break the connection among data and to make better use of historical data samples, i.e., the agent stored data in a database in the process of reinforcement learning and subsequently used the uniform random sampling method to extract data from the database to a minibatch. And the state of each episode was initialized randomly at the training stage to avoid model uncertainties caused by initial conditions (Li et al., 2009). Finally, the extracted data were used to train the neural network, as shown in Fig. 2 and Table 4.

The exact neural network architecture is as follows. The input layer is a matrix of feature vector of stowage samples, the output layer is the approximate Q value of each action. Thus, the number of nodes in input layer is 11, the number of nodes in output layer is 3. There are 2 hidden layers in this case and the first and second hidden layer are both fully-connected and consists of 7 and 5 units respectively.

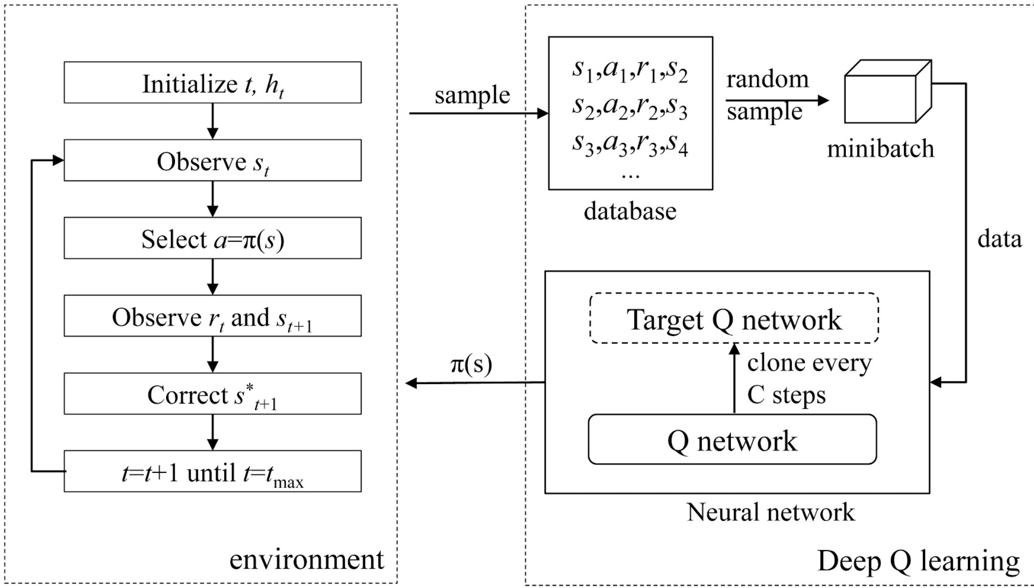


**Fig. 2.** Overall architecture of the system.

**Table 4**
Procedures of the DQN irrigation decision-making algorithm.

| Algorithm 1 Deep Q-learning network for irrigation decision-making algorithm |
|---|
| 1: Initialize replay memory D to capacity $N$ |
| 2: Initialize action-value function Q with random weights $\theta$ |
| 3: Initialize target action-value function Q with weights $\theta^- = \theta$ |
| 4: **For** episode=1,…, M do |
| 5: Collect the environmental condition and initialize $s$ |
| 6: **For** $t$ =1,…, T do |
| 7: With probability ε select a random action $a_t$; otherwise, select $a_t = \max_{a_t} Q(s_t, a_t; \theta^-)$ |
| 8: Execute action at in emulator and observe reward $r_t$ and next state $s_{t+1}$ |
| 9: Store ($s_t$,$a_t$,$r_t$,$s_{t+1}$) in D |
| 10: **If** size of D > size of minibatch |
| 11: Sample random minibatch of ($s_j$,$a_j$,$r_j$,$s_{j+1}$) from D |
| 12: Set $y_j = \begin{cases} r_j & \text{for terminal } s_{j+1} \\ r_j + \gamma \max_{a_{j+1}} Q(s_{j+1}, a_{j+1}; \theta^-) & \text{otherwise} \end{cases}$ |
| 13: Perform a gradient descent step on $(y_j - Q(s_j, a_j; \theta))^2$according to $\theta$ |
| 14: Every C steps clone $\theta^- = \theta$ |
| 15: **If** exploration rate $\varepsilon < \varepsilon_{min}$ |
| 16: $\varepsilon \leftarrow \varepsilon - \Delta\varepsilon$ |
| 17: Correct $s_{t+1}$=$s^*_{t+1}$ |
| 18: **End** for |
| 19: **End** for |

## 3. Results and discussions

### 3.1. Evaluation of rainfall forecast

Statistical indices for the daily rainfall forecasting performance of various rainfall levels with a lead time of 7 days in the growth period for three types of rice are presented in Fig. 3. For all types of rice, Fig. 3a shows that the average threat score (TS) of different precipitation grades ranged between 0.38 and 0.70 and that the TS of all precipitation grades and average TS decreased with increasing lead times. A significant difference was found among the rice types in that the TS of early rice was the lowest, followed by that of middle rice and late rice. For different precipitation grades, the TS of forecasts with no rainfall is significantly higher than that of forecasts with rainfall, and the TS decreased with increasing rainfall level.

Fig. 3b-c also showed that the average missing alarm rate (MAR) and false alarm rate (FAR) ranged from 0.20 to 0.39 and 0.14 to 0.49, respectively. The MAR and FAR of all precipitation grades and the average MAR and FAR decreased with increasing lead times. Similar to TS, the MAR and FAR of early rice were the highest, followed by those of middle rice and late rice. A clear increasing trend is observed for MAR and FAR with increasing rainfall level, except that the trends of HR and SR were almost zero for certain lead times due to low frequency of occurrence. In general, the performance of the rainfall forecasts varied greatly in time and precipitation grades, and the rainfall forecasts performed worse with increasing lead time and rainfall level, which was in agreement with results obtained in previous studies (Luo et al. 2016; Cao et al. 2019). These results suggested that the daily rainfall forecasting performance was acceptable and that potentially uncertain rainfall exists for learning and utilization.

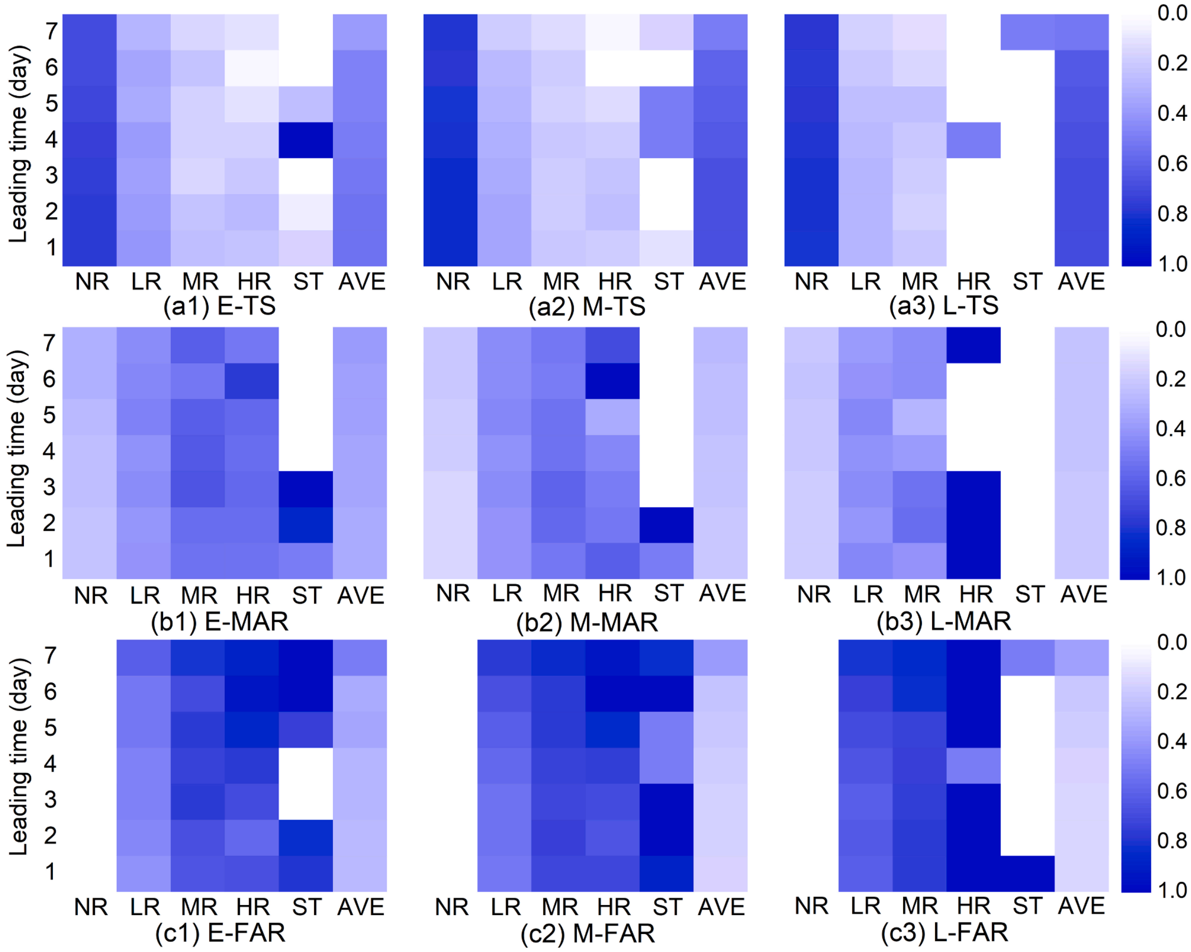


**Fig. 3.** Statistical indices for daily rainfall forecasting performance with a lead time of 7 days in the rice growth period. E, M, and L represents early-rice, middle-rice and late-rice growth periods, respectively and NR, LR, MR, HR, ST and AVE represents no rain, light rain, moderate rain, heavy rain, storm or above and average respectively.

### 3.2. Performance evaluation of DQN irrigation decision-making strategy

Loss is the objective function value during neural network training (Eq. (22)), representing the degree to which the neural network approximates the discrete action value. As shown in Fig. 4a–c, the loss was large at the beginning because the information was insufficient to approximate satisfactorily and subsequently decreased rapidly in 100 iterations. After approximately 500–2000 iterations, the parameters of each iteration differed less, the loss value tended to be stable and fluctuating, and the neural network could better approach the action value under various states.

Slightly different from the loss function, the mean reward first decreased slightly and then increased dramatically because at first the exploration strategy was more frequently used to obtain an adequate sample of reward values on every state (i.e., exploration rate $\varepsilon < \varepsilon_{min}$) which was likely to perform unreasonable action with less reward. And then the exploitation strategy was adopted in deciding on the action with highest reward. After approximately 200–1000 iterations of each rice type, the mean reward increased significantly and stabilized.

As for decomposed rewards, the average values of basic reward($r_0$), rainfall utilization reward($r_1$) and yield reward($r_2$) of samples that $h<h_{min}$ were calculated. Because the proportion of samples that $h>h_{min}$ was significantly large, but the learning rules were simple and the reward values were concentrated (Eqs. (15)–(18)), which made it likely to blur the learning performance of samples that $h < h_{min}$ when calculating the mean reward. As can be seen from Fig. 4d–l, basic reward, rainfall utilization reward and yield reward had some extreme high and low points at the beginning of training since the model was taking the random exploration strategy and adopting unreasonable strategy frequently, which led to over irrigation ($r_1 < 1$, $r_2 = 1$) and potential yield loss ($r_1 = 1$, $r_2 < 1$). Then $r_0$, $r_1$ and $r_2$ rapidly converged and increased, indicating that the model started to optimize the strategy according to the feedback rewards. $r_0$ was in a steady rise in the middle and late stage which proved that $r_0$ had a large learning space and was optimized steadily. While $r_1$ showed a slow and slightly decreasing trend in the middle and late period. Unlike $r_1$, $r_2$ rose slowly in the late period. The reason why the trends of $r_1$ and $r_2$ were different in the middle and late period might be the checks and balances between $r_1$ and $r_2$. To be more specific, the increase of $r_1$ will lead to the decrease of $r_2$. Therefore, the advantages and disadvantages will be automatically weighed in the training, and the improvement of rainfall utilization rate will be marginally sacrificed to ensure no yield reduction. Table 5 shows that the average reward on the verification set was not much different from the average reward on the training set, indicating that the model had a strong generalization ability and can be used to make irrigation decisions with future weather forecasts.

It is worth noting that there are a number of important differences in the trend of reward function among early rice, middle rice and late rice

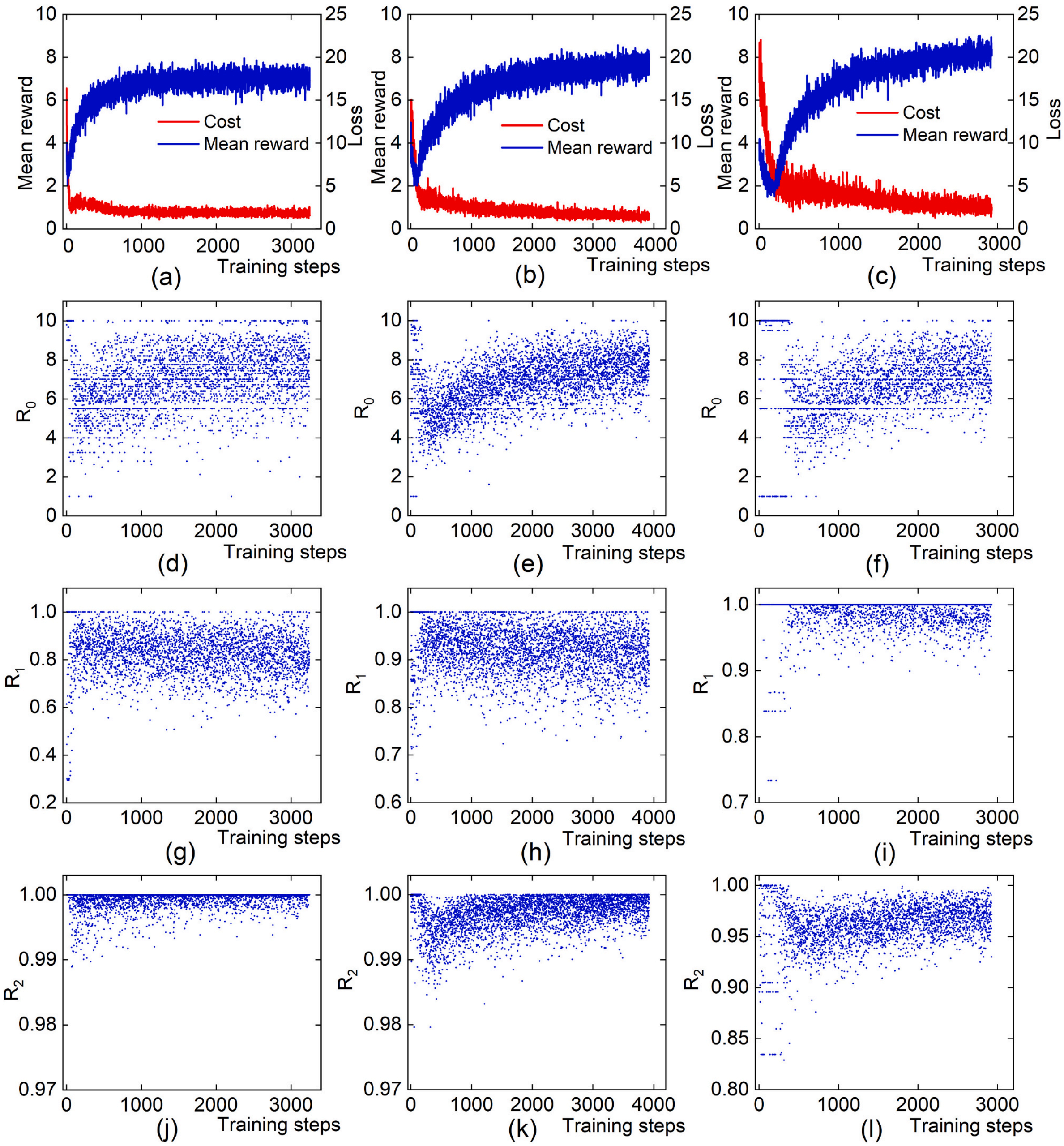


**Fig. 4.** Training curves tracking the loss, mean reward and scatter plot of decomposed reward ($r_0$,$r_1$ and $r_2$) for early rice [(a), (d), (g),(j)], middle rice [(b), (e), (h), (k)] and late rice [(c), (f), (i),(l)].

**Table 5**
Mean reward of training set (2012–2018) and verification set (2019).

| Type | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|---|---|---|
| Early rice | – | 7.88 | 7.08 | 6.85 | 7.50 | 7.43 | 8.47 | 7.57 |
| Middle rice | 8.70 | 9.27 | – | 8.60 | 8.95 | 8.06 | 9.02 | 8.70 |
| Late rice | 9.04 | 8.99 | – | 8.77 | 9.08 | 8.54 | 9.21 | 9.56 |

because the focus of the reward function was to maximize rainfall use efficiency and to avoid the yield loss. With increasing rainfall, the probability of drainage increased and the rainfall utilization reward was less than 1 when drainage occurred, while the probability of drought decreased and the yield reward was less than 1 when water deficit occurred. Thus, the value of the reward function had a certain correlation with rainfall in the rice growth period, which showed that the rainfall utilization reward of early rice with a large amount of rainfall was lower and more scattered than that of middle rice and late rice with less rainfall, whereas yield reward of early rice was more concentrated and larger than that of middle rice and late rice.

### *3.3. Analysis of the water-saving effect of the DQN irrigation decision-making strategy*

The results of the DQN irrigation strategy compared with the results of the conventional irrigation strategy showed obvious decreases in the amount of irrigation water (Fig. 5a1–c1), irrigation timings (Fig. 5a2–c2) and drainage water (Fig. 5a3–c3) without yield losses (i.e., the rate of yield losses in the simulation is equal to 0).

Fig. 5a1–c1 show that the DQN irrigation strategy could further conserve irrigation water with an average decrease of 23 mm compared with the conventional irrigation strategy. The average water-saving rates of early rice, middle rice and late rice were 23%, 6% and 3%, and the standard deviations were 12%, 4% and 4%, respectively, indicating a significant difference among different rice varieties and years, possibly caused by the distinct precipitation distributions and weather forecast patterns during the growth period of the three types of rice. During the early-rice growth period (usually from April to July), with abundant rainfall and lower crop water demand for lower temperature, early rice only needed minimal irrigation to grow properly; thus, delaying irrigation to use future rainfall can conserve a large proportion of irrigation water consumption. Moreover, unpredicted rainfall (a high MAR was found in 3.1) was likely to occur during this period, and this unexpected rainfall can be used as supplementary irrigation by learning and training the data via the DQN algorithm. During the growth period of middle rice (from May to October), much rainfall occurred, but the crop water demand was also high, requiring substantial irrigation. Therefore, the impact of rainfall on irrigation was weakened, making the water-saving effect inferior to that of early rice. During the growth period of late rice (from July to October), although the water demand of rice decreased due to the decrease in temperature, there was less rainfall that could be used as supplementary irrigation due to the significant decrease of rainfall, and thus the water-saving effect was the least. Fig. 5a2–c2 show that the DQN irrigation strategy reduced irrigation timings with an average decrease of 0.98 compared with the conventional irrigation strategy. The proportion of reduction of irrigation times was different across rice varieties, and irrigation timings of early rice decreased more than that of middle rice and late rice. There were 1.4 (31%), 0.6 (5%) and 1.3 (9%) irrigation timing reductions and 0.6, 0.9 and 0.4 standard deviations in early rice, middle rice and late rice, respectively. Fig. 5a3–c3 show that the DQN irrigation strategy reduced drainage water with an average decrease of 21 mm compared with the conventional irrigation strategy. The amounts of drainage water of early rice, middle rice and late rice were reduced by 7%, 8% and 9% with standard deviations of 7%, 8% and 11%, respectively.

Simulation results showed that applying DQN irrigation strategy did not result in yield losses in any given year, not only because the reward function took into account the influence of yield, but also because the lower limit of local traditional irrigation criteria was relatively high, even if irrigation is delayed occasionally, there would be no water deficit affecting the crop evapotranspiration, resulting in the yield loss.

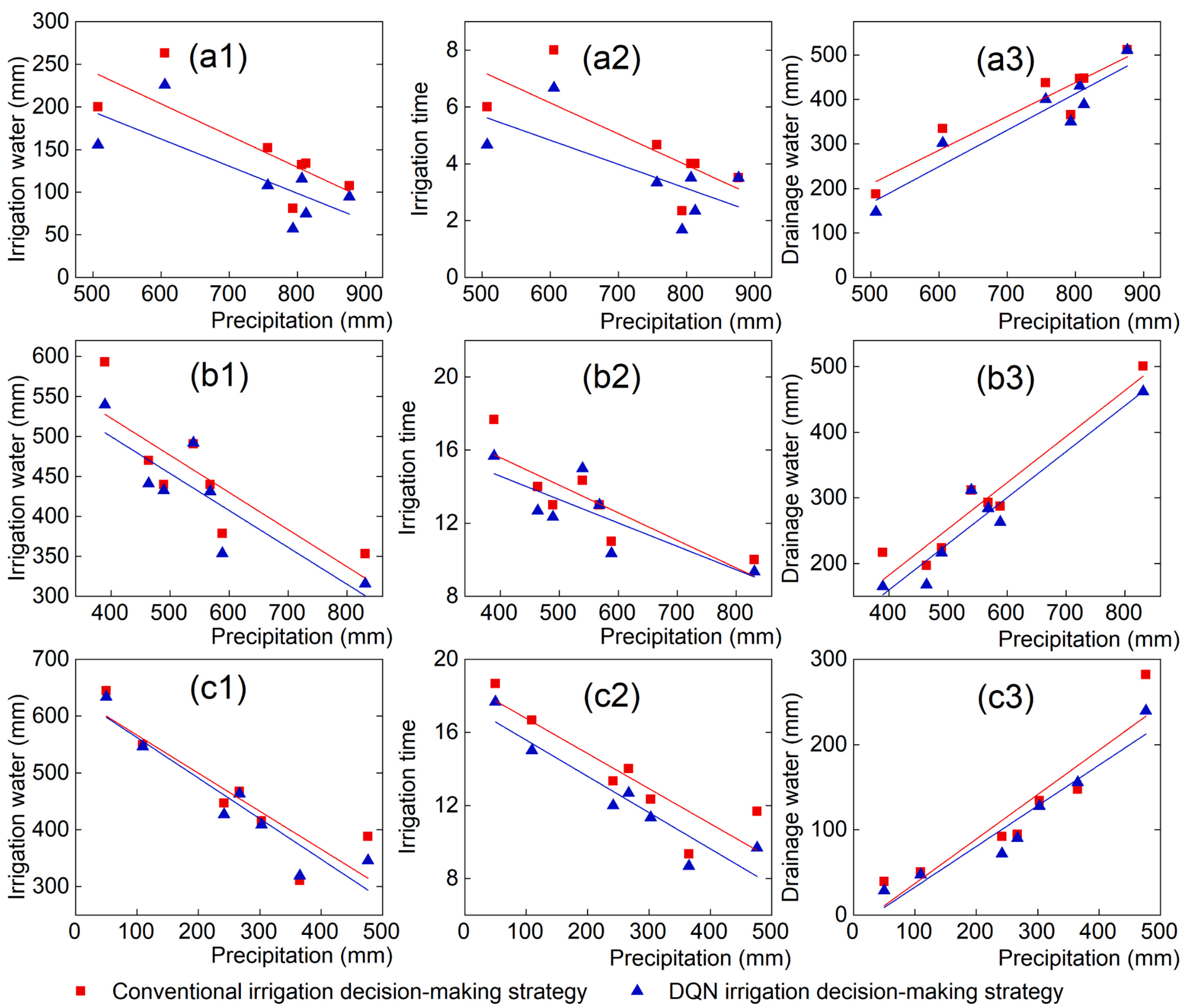


**Fig. 5.** Irrigation water, irrigation timings and drainage water from 2012 to 2019 for early rice [(a1), (a2), (a3)], middle rice [(b1), (b2), (b3)] and late rice [(c1), (c2), (c3)].

Therefore, DQN irrigation decision based on flooded irrigation can be regarded as a relatively safe irrigation strategy and has the potential to further save water.

The variations in the water-saving effect among rice types supported evidence for previous observations (e.g., Cai et al., 2011; Gowing and Ejieji, 2001), which found that using the short-term weather forecasts in wet years could be beneficial due to cost and water savings from unnecessary irrigation. However, in average and dry years, profits could derive from improved efficiency in the use of a limited water supply. Correspondingly, this study showed that the performance of the DQN irrigation strategy for early rice was better than that of middle rice and late rice, the rainfall of which was less than that of early rice. It should be noted that although the costs of irrigation were not included in the DQN model as a limitation of this study, the irrigation timings could indirectly reflect the input of irrigation management.

The value of the action-value function (Q) for each action in the three types of rice under the same criterion of field water depth is shown in Fig. 6. A reasonable increase occurred in the average Q value with the increase in water depth of the field in the case of taking action 0, whereas a decrease in the average Q value was found with an increase in water depth of the field in the case of taking action 1 or 2, and the Q value of action 2 was always higher than that of action 1 because the $r_1$ (rainfall utilization reward) of the reward function for action 2 was higher than that of action 1 when $h < h_{min}$ for reducing the frequency of irrigation. Additionally, the average Q values of action 0 and action 2 were approximately equal at $h = 20$ mm, i.e., $h_{min}$, when predicted precipitation is approximately equal to 0, which indicated that DQN was able to learn the basic irrigation rule that paddy rice should be irrigated when the water depth of the field reached the lower limit of water depth. Considering the future rainfall, when $h > h_{min}$, the Q value was

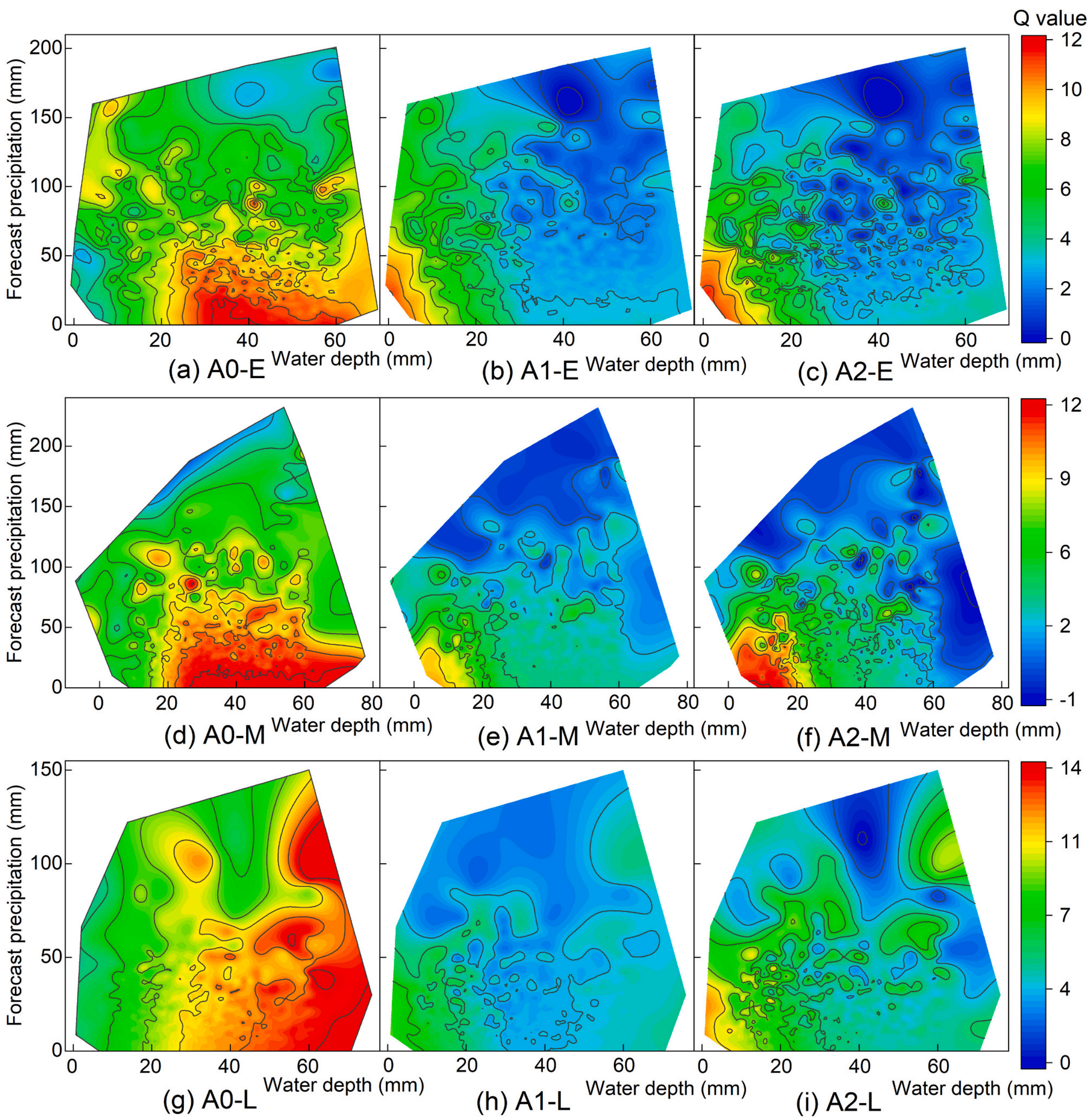


**Fig. 6.** Value of action-value function (Q) of each action in three types of rice. A0, A1, and A2 represent action 0, action 1 and action 2, respectively.

generally at a relatively high stage if taking action 0 but at a low stage if taking action 1 or 2. When $h < h_{min}$, the Q value of action 0 increased with the increase in predicted rainfall, whereas the Q values of action 1 and action 2 decreased with the increase in predicted rainfall. It is worth noting that the above patterns were most obvious in early rice, followed by middle rice, and the least in late rice, which was related to the rainfall during the growth period of paddy rice. It is clear that less rainfall during the growth period of paddy rice gives less information on rainfall in the corresponding weather forecasts. Therefore, there were insufficient samples containing key information (i.e., samples for which rain is predicted in the future) in the model during training; thus the Q value distribution of the late rice model was not as obvious as that of the middle rice and late rice models.

In addition, the forecast frequencies of various rainfall grades within 1–7 days of each action when $h < h_{min}$ were analyzed (Fig. 7). The results showed that the average grade for the rainfall forecast of early rice was the highest (i.e., the average color was deepest), followed by middle rice and late rice. Although the Q value of action 1 was always lower than that of action 2 with the same amount of predicted rainfall and the same criterion of field water depth (see Fig. 6), there were still cases of action 1 (except for late rice) due to random distribution of rainfall. It is apparent from this figure that the average grade for the rainfall forecast of action 0 was higher than that of action 1 and 2. As the forecast lead time increased, a decreasing trend was noted in the average grade for the rainfall forecasts of action 0, an increasing trend occurred in that of action 2, and the average grade for the rainfall forecasts of action 1 first increased and subsequently decreased. These results indicated that the future rainfall was considered in the model. When the field needed irrigation, the DQN model chose to delay irrigation to avoid over-irrigation, leading to wasted water if heavy rainfall was expected in the short-term future (in 1–3 days). Alternatively, the model chose to reduce the amount of irrigation water if rainfall was expected in the middle-term future (in 4–5 days) to relieve the crop from drought and reduce water waste or chose to irrigate immediately according to the conventional rules if there was rainfall in the long-term future (in 6–7 days) or almost no forecast rainfall in the next 7 days. These results are in accordance with the irrigation rules developed by Cao et al. (2019). It is worth noting that the rainfall forecasts were not involved in the reward function of the DQN model. It is possible to hypothesize that even though the model did not directly determine which decision to make in a specific situation, the model obtained feedback through the interaction with the environment and continuously took advantage of implicit domain knowledge, thus summarizing the irrigation experience in the manner of experts.

This paper proposed a reinforcement learning approach to irrigation decision-making for rice using weather forecasts. Certain limitations exist with respect to both sample size and station selection. First, unlike traditional machine learning projects such as autonomous driving and game development, which can produce data artificially, only historical information collected from meteorological data and weather forecast

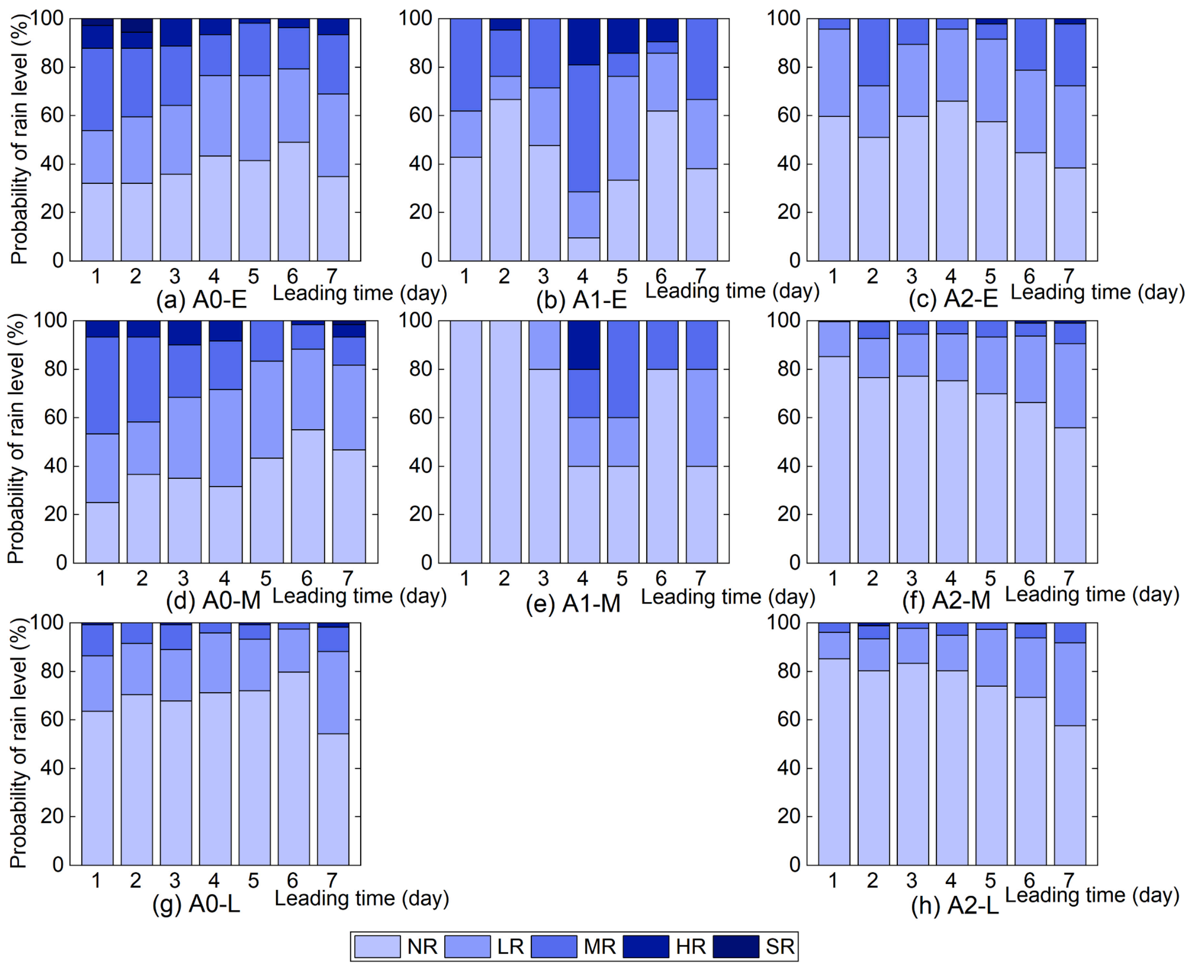


**Fig. 7.** Probability of forecast rainfall of each action in three types of rice.

data were used in training in this study. Although the meteorological data and weather forecast data from three meteorological stations near Nanchang were collected to increase the number of samples in the training set, an obvious gap still existed compared with the large amount of data in traditional machine learning projects, and thus, limitation could reduce the training effect of this new method. Therefore, improving the utilization of samples should be a topic of future research. Second, only Nanchang, a typical rice-growing area in China, was selected for verification in this paper, without consideration of the spatial variability in climate and soil conditions. Thus, more typical stations should be included to assess the spatial variability performance for the proposed method.

In terms of practical application, this research only discussed the water-saving effect of the irrigation decision for a single station. For a small-size irrigated area with quick response to irrigation, this approach can be directly applied to improve the utilization rate of rainfall and increase irrigation profits. However, for large-size irrigated areas, the water-saving effect of irrigation decisions is limited by multi-level constraints such as water distribution efficiency and available water supply. Therefore, in the process of practical application, this proposed method should be embedded into irrigation scheduling optimization of large-size irrigated areas as one a sub-goal.

## 4. Conclusions

In this study, an DQN irrigation decision-making strategy suitable for paddy rice was proposed using weather forecasts, and Nanchang station in southern China was selected to validate the strategy. The forecasting performance of daily rainfall for lead times of 1–7 days was evaluated for three types of rice, and the training process and generalization of the DQN were discussed. The water-saving effects of the DQN irrigation decision and those of the conventional irrigation decision were compared.

The results indicated that the TS of rainfall in the weather forecasts ranged from 0.38 to 0.70 with a certain degree of false alarms and missed alarms, which showed that weather forecasts can be used as reliable inputs for the proposed DQN irrigation decision-making strategy.

The training process of the DQN algorithm converged rapidly, and the nonobvious difference between the average reward of the training set and that of the verification set showed that the strategy had a strong generalization ability and can be used to make irrigation decisions with future weather forecasts.

In our case, simulated results of DQN algorithm indicated that, compared with conventional irrigation decisions, the DQN strategy can conserve irrigation water by 23 mm and reduce drainage water by 21 mm and irrigation timing by 1.0 on average without yield reduction theoretically, considering the risk of irrigation water waste and yield reduction. The results also showed that the DQN strategy was able to summarize the past irrigation experiences after training and select the most appropriate irrigation decisions based on current field water conditions and forecast rainfall patterns, thereby increasing rainfall utilization and conserving irrigation water.

## Declaration of Competing Interest

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

## Acknowledgements

This work was financially supported by the NSFC-MWR-CTGC Joint Yangtze River Water Science Research Project (No. U2040213), Department of Science and Technology of Hubei Province under S&T Cooperation with Foreign Country Project (No. 2019AHB076), Department of Water Resources of Jiangxi Province under Science and Technology Project (No. KT201736) and Department of Science and Technology of Guangxi Zhuang Autonomous Region under Key Research & Development (R&D) Plan (No. AB18126093). The observed meteorological data obtained from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn) and weather forecast data from Weather China (http://www.weather.com.cn) are gratefully acknowledged. We also would like to thank the editor and all anonymous reviewers for their valuable comments.

## References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements, FAO Irrigation and Drainage Paper 56, Rome, Italy.
Bergez, J.E., Garcia, F., 2010. Is it worth using short-term weather forecasts for irrigation management? Eur. J. Agron. 33, 175–181.
Bergez, J.E., Eigenraam, M., Garcia, F., 2001. Comparison between dynamic programming and reinforcement learning: a case study on maize irrigation management, in: Proceedings of the 3rd European Conference on Information Technology in Agriculture (EFITA01), Montpellier, pp. 343–348.
Blaney, H.F., Criddle, W.D., 1962. Determining Consumptive Use and Irrigation Water Requirements. USDA ARS Tech Bull, Washington (DC).
Bu, F., Wang, X., 2019. A smart agriculture IoT system based on deep reinforcement learning. Future Gener. Comput. Syst. 99, 500–507.
Cai, X., Hejazi, M.I., Wang, D., 2011. Value of probabilistic weather forecasts: assessment by real-time optimization of irrigation scheduling. J. Water Resour. Plan. Manag. 137 (5), 391–403.
Cao, J., Tan, J., Cui, Y., Luo, Y., 2019. Irrigation scheduling of paddy rice using short-term weather forecast data. Agric. Water Manag. 213, 714–723.
Chen, T., Su, W., 2018. Local energy trading behavior modeling with deep reinforcement learning. IEEE Access 6, 62806–62814.
Donaldson, R.J., Dyer, R.M., Kraus, M.J., 1975. Objective evaluator of techniques for predicting severe weather events. Bull. Am. Meteorol. Soc. 56 (7), 755–755.
Doorenbos J., Pruitt W.O., 1975. Guidelines for predicting crop water requirements, FAO Irrigation and Drainage paper No. 24, FAO, Rome, Italy.
Görges, D., 2017. Relations between model predictive control and reinforcement learning. IFAC-PapersOnLine 50, 4920–4928.
Gowing, J.W., Ejieji, C.J., 2001. Real-time scheduling of supplemental irrigation for potatoes using a decision model and short-term weather forecasts. Agric. Water Manag. 47, 137–153.
Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Am. Soc. Agric. Eng. 1 (2), 85–97.
Irmak, S., Allen, R.G., Whitty, E.B., 2003. Daily grass and alfalfa-reference evapotranspiration estimates and alfalfa-to-grass evapotranspiration ratios in Florida. J. Irrig. Drain. Eng. 129 (5), 360–370.
Irukula, S., 2015. Reinforcement learning based controller for precision. Acta Univ. Agric. Silvic. Mendel. Brun.
Jensen, M., 1968. Water consumption by agricultural plants. In: Kozlowski, T.T. (Ed.), Water Deficits and Plant Growth. Development, Control, and Measurement, New York, pp. 1–22.
Lee, E.S., Raju, K.S., Biere, A.W., 1991. Dynamic irrigation scheduling with stochastic rainfall. Agric. Water Manag. 19, 253–270.
Li, H., Luo, L., Wood, E.F., Schaake, J., 2009. The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. J. Geophys. Res. 114, D04114.
Linker, R., Sylaios, G., 2016. Efficient model-based sub-optimal irrigation scheduling using imperfect weather forecasts. Comput. Electron. Agric. 130, 118–127.
Longo-Minnolo, G., Vanella, D., Consoli, S., Intrigliolo, D.S., Ramírez-Cuesta, J.M., 2020. Integrating forecast meteorological data into the ArcDualKc model for estimating spatially distributed evapotranspiration rates of a citrus orchard. Agric. Water Manag. 231, 105967.
Lorite, I.J., Ramírez-Cuesta, J.M., Cruz-Blanco, M., Santos, C., 2015. Using weather forecast data for irrigation scheduling under semi-arid conditions. Irrig. Sci. 33, 411–427.
Luo, J., 2003. Water Saving Irrigation Theories and Techniques. Wuhan University Press, Wuhan, China (in Chinese).
Luo, Y., Chang, X., Peng, S., Khan, S., Wang, W., Zheng, Q., Cai, X., 2014. Short-term forecasting of daily reference evapotranspiration using the Hargreaves-Samani model and temperature forecasts. Agric. Water Manag. 136, 42–51.
Luo, Y., Ma, Z., Lyu, X., Guo, L., C, Y, Xiang, Z., 2016. Accuracy assessment of rainfall forecast and analysis on water-saving irrigation strategies during rice growing seasons in South China. J. Drain. Irrig. Mach. Eng. 34, 430–442 (in Chinese).
Mao, Z., 2002. Water-saving irrigation of paddy rice and its effect on environment. Eng. Sci. 4, 8–16 (in Chinese).
McCloud, D.E., 1955. Water requirements of field crops in Florida as influenced by climate. Proc. Soil Sci. Soc. Florida 15, 165–172.
Medina, H., Tian, D., Srivastava, P., Pelosi, A., Chirico, G.B., 2018. Medium-range reference evapotranspiration forecasts for the contiguous United States based on multi-model numerical weather predictions. J. Hydrol. 562, 502–517.
Mishra, A., Siderius, C., Aberson, K., van der Ploeg, M., Froebrich, J., 2013. Short-term rainfall forecasts as a soft adaptation to climate change in irrigation management in North-East India. Agric. Water Manag. 127, 97–106.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M., 2013. Playing atari with deep reinforcement learning. arXiv preprint arXiv: 1312.5602.
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature 518, 529–533.
Mondaca-Duarte, F.D., van Mourik, S., Balendonck, J., Voogt, W., Heinen, M., van Henten, E.J., 2020. Irrigation, crop stress and drainage reduction under uncertainty: a scenario study. Agric. Water Manag. 230, 105990.
NBSPRC, 2019. China Statistical Yearbook. National Bureau of statistics of the People's Republic of China, Beijing China (in Chinese).
Perera, K.C., Western, A.W., Nawarathna, B., George, B., 2014. Forecasting daily reference evapotranspiration for Australia using numerical weather prediction outputs. Agric. Meteorol. 194, 50–63.
Perera, K.C., Western, A.W., Robertson, D.E., George, B., Nawarathna, B., 2016. Ensemble forecasting of short-term system scale irrigation demands using real-time flow data and numerical weather predictions. Water Resour. Res. 52, 4801–4822.
Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92.
Shen, Y., Zhao, N., Xia, M., Du, X., 2017. A deep Q-learning network for ship stowage planning problem. Pol. Marit. Res. 24, 102–109.
Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction. The MIT Press, London, England.
Traore, S., Luo, Y., Fipps, G., 2016. Deployment of artificial neural network for short-term forecasting of evapotranspiration using public weather forecast restricted messages. Agric. Water Manag. 163, 363–379.
Watkins, C., Dayan, P., 1992. Q-learning. Mach. Learn. 8, 279–292.
Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences. Academic Press, San Diego.
Wu, J., He, H., Peng, J., Li, Y., Li, Z., 2018. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. Appl. Energy 222, 799–811.
Yang, Y., Cui, Y., Luo, Y., Lyu, X., Traore, S., Khan, S., Wang, W., 2016. Short-term forecasting of daily reference evapotranspiration using the Penman-Monteith model and public weather forecasts. Agric. Water Manag. 177, 329–339.