S.I.: MACHINE LEARNING AND BIG DATA ANALYTICS FOR IOT SECURITY AND PRIVACY (SPIOT 2023)

Intelligent control and optimization of hydraulic systems using reinforcement learning

Wenjing Hu[1] · Yuejing Jiang[1]

Received: 20 December 2024 / Accepted: 4 June 2025

Abstract
Hydraulic systems play an essential role in many industrial applications. However, traditional control methods such as PID (proportional–integral–derivative) control have problems such as low accuracy, slow response, and poor adaptability. This paper proposes an intelligent control and optimization framework based on deep neural networks to solve the control difficulties encountered by hydraulic systems in practical applications. First, the working data of the hydraulic system are collected in real time through multiple sensors and preprocessed. Then, a deep neural network is used to model the hydraulic system nonlinearly to capture its complex input–output relationship and achieve precise prediction. Reinforcement learning is applied to optimize the control strategy, and a deep Q network model is used for adaptive control. Under variable working conditions, reinforcement learning continuously adjusts control parameters (such as pump speed and valve opening) through the interaction between the intelligent agent and the environment to achieve optimal system performance. In addition, a multi-objective optimization strategy is adopted to comprehensively consider control accuracy, response speed, and energy efficiency and balance them through a comprehensive reward function. Experimental results show that the proposed method significantly improves the hydraulic system's control accuracy, response speed, and energy efficiency, which is better than traditional control methods. The response time from the change of the system input signal to the stable output of the system is 2.4 s, and the system energy efficiency ratio is less than 0.7 under high load. The research in this paper provides new ideas for the intelligent control and optimization of hydraulic systems and has broad application prospects.

Keywords Hydraulic system · Intelligent control · Machine learning · Deep neural network · Reinforcement learning

1 Introduction

Traditional hydraulic system control methods, such as PID (proportional–integral–derivative) control and classical linear control methods, often rely on preset parameters and models and are difficult to deal with the system's nonlinearity, time-varying nature, and external disturbances [1, 2]. These methods have low control accuracy and response speed in practical applications such as complex working conditions, load changes, and system aging and are difficult to optimize in real time [3, 4]. As the application scenarios of hydraulic systems become increasingly complex, improving their control accuracy and optimizing system performance, especially their adaptability in dynamic and complex environments, have become urgent issues that need to be addressed [5, 6].

https://doi.org/10.1007/s00521-025-11430-z

Published online: 21 July 2025

The purpose of this study is to propose an intelligent control and optimization framework for hydraulic systems based on machine learning. By combining deep learning and reinforcement learning techniques, this paper designs an adaptive control system that dynamically adjusts the control strategy according to real-time data, thereby improving the hydraulic system's response speed, control accuracy, and energy efficiency. The research results of this paper provide a new idea and method for the intelligent optimization of hydraulic systems and an essential reference for future industrial applications. The innovation of this method lies in constructing a closed-loop system that integrates data-driven modeling and dynamic decision-making, effectively balancing the conflicts among control accuracy, response speed, and energy efficiency. This study optimizes the hydraulic system through intelligent control, significantly improving energy efficiency and reducing energy consumption. The intelligent optimization strategy combining deep learning and reinforcement learning is adopted to promote the thoughtful process of hydraulic system control. In addition, this study combines hydraulic control technology, machine learning, and artificial intelligence technology, fostering technological innovation and integration in these fields. By combining deep learning and reinforcement learning for hydraulic systems, the study promotes the practical application of emerging technologies in traditional industries, which helps to accelerate technological progress and industrial upgrading.

2 Related work

In recent years, many scholars have proposed various improvement schemes to address the shortcomings of hydraulic system control accuracy and response speed [7, 8]. For example, model-based control methods (such as model predictive control and robust control) achieve relatively stable control by establishing precise mathematical models, but these methods have high requirements for modeling hydraulic systems and have strong model dependence [9, 10]. In addition, data-driven control methods (such as machine learning algorithms) have gradually attracted attention [11, 12]. By learning the system's input and output data, machine learning can effectively avoid the difficulty of precise modeling and has greater adaptability and flexibility [13, 14]. However, existing research focuses on optimizing a single control objective and lacks an integrated strategy for multi-objective and system-global optimization [15, 16].

To meet the challenges mentioned above, many studies have been devoted to integrating machine learning algorithms into hydraulic systems' intelligent control and optimization strategies [17, 18]. Deep neural networks have been adopted as an effective tool to model the complex nonlinear characteristics of hydraulic systems. On the other hand, reinforcement learning technology has shown its potential to optimize system performance by dynamically adjusting control strategies [19, 20]. Particularly in the face of changing and unpredictable environmental scenarios, reinforcement learning has achieved remarkable results in hydraulic system control due to its excellent adaptive ability [21, 22]. Nevertheless, these advanced methods still need to overcome a series of obstacles in actual deployment, including the complexity of data acquisition, the limitations of algorithm generalization ability, and the long training cycle [23]. Therefore, the focus and difficulty of current research is on how to effectively integrate cutting-edge algorithms such as deep learning and reinforcement learning, aiming to combine the deep analysis of historical data with the rapid response of real-time feedback, thereby improving the accuracy and response speed of hydraulic system control.

3 Data and methods

3.1 Data acquisition and preprocessing

Data acquisition is the basis of intelligent control of hydraulic systems. A hydraulic system's operating status covers several key physical parameters, captured in real time by a series of specialized sensors, including pressure sensors, flow measurement devices, and temperature monitoring sensors, which are used to capture the system's working status data. Specifically, the pressure sensor is responsible for monitoring the working pressure of each component in the hydraulic system; the flow measurement device is used to quantify the flow rate of the fluid; the temperature monitoring sensor focuses on the dynamic changes of the fluid temperature. These sensors are usually deployed in the hydraulic system's core components, such as hydraulic pumps, control valves, and actuators, to comprehensively and accurately reflect the system's overall operating status. To ensure the representativeness and comprehensiveness of the collected data, the experiment covers the operating status of the hydraulic system under various typical working conditions such as different loads, temperature fluctuations, pump speed adjustment, and valve opening changes. The sensor deployment follows the physical characteristics and functional relevance of the core components of the hydraulic system. In addition to the hydraulic pump, control valve, and actuator, the auxiliary circuit pressure and flow parameters are monitored synchronously to capture the dynamic coupling effect of the system fully. The data collection cycle covers system startup, steady-state operation, load mutation, and shutdown stages, combining long-term continuous sampling with high-density transient recording to ensure the adequacy of the data in terms of time scale and working condition dimensions. In the data acquisition process, to ensure the timeliness and accuracy of the acquired data, the sampling frequency of the data acquisition system is set to at least 10 times per second, to capture the dynamic fluctuations of the hydraulic system in a very short time. For different working conditions under experimental conditions, the system flexibly adjusts the sampling frequency and the layout of sampling points according to actual needs, aiming to ensure the complete acquisition of all key information, as shown in Table 1.

The data in Table 1 reflect the hydraulic system's pressure, flow, and temperature values at a specific time. The data acquisition frequency is 10 times per second, which can track the hydraulic system's dynamic changes in real time.

The first step of data cleaning is to detect and process missing values. For missing data caused by sensor failure, interpolation fills in the missing values. Common interpolation methods include linear interpolation and Lagrange interpolation. Linear interpolation is suitable for situations where data change smoothly, while Lagrange interpolation is suitable for scenarios where data changes are more complex. The missing values in this article account for a low proportion (less than 5%) and are processed by interpolation.

Secondly, outliers need to be identified and eliminated. Outliers may be caused by sensor failure, system failure, or sudden changes in the external environment and are usually manifested as points far away from the normal data distribution. To this end, this paper uses two methods, IQR (interquartile range) and Z-score, to detect outliers. The IQR method calculates the data quartile difference and sets the upper and lower limits, and data points outside the range are considered outliers. The Z-score method calculates the deviation of the standard deviation of the data point from the mean, and points exceeding the set threshold are identified as outliers. These abnormal data are removed or replaced with reasonable values estimated by interpolation.

Table 1 Data acquisition

Time (s)	Pressure sensor (MPa)	Flow meter (L/min)	Temperature sensor (°C)
0.1	12.5	45.2	35.4
0.2	12.6	45.5	35.5
0.3	12.7	45.6	35.7
0.4	12.9	45.7	36
0.5	13	45.8	36.2
0.6	13.1	45.9	36.4
0.7	13.2	46	36.6
0.8	13.3	46.2	36.8
0.9	13.5	46.3	37
1	13.6	46.5	37.2

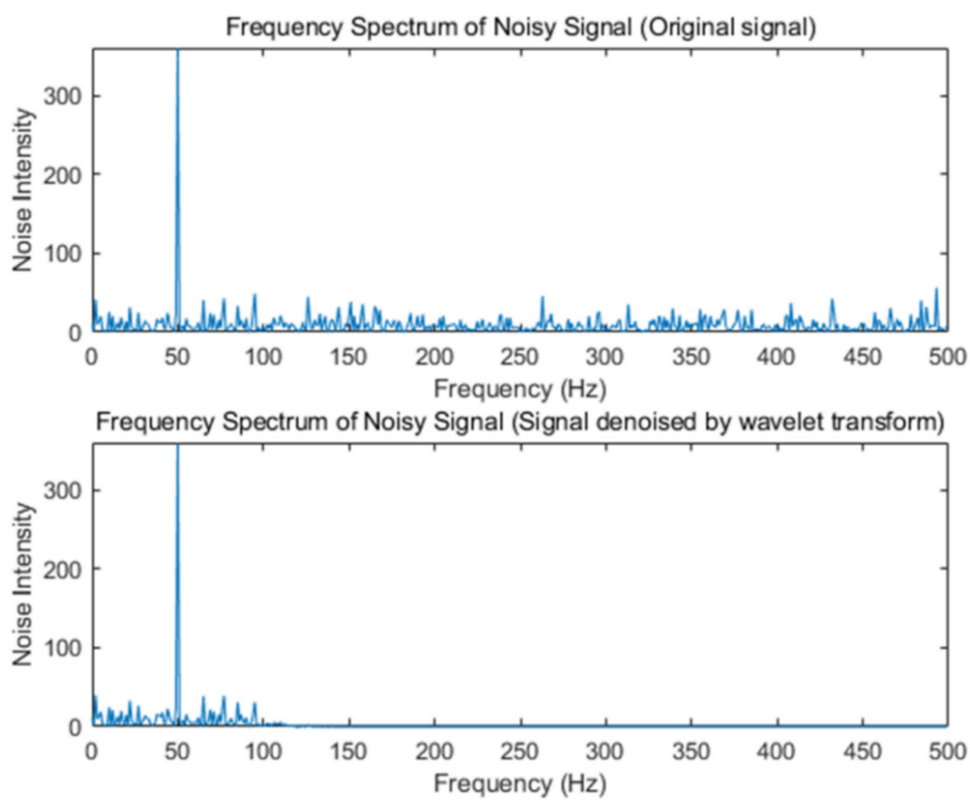
The operating data of hydraulic systems are usually interfered with by noise, which may come from electrical devices, external environmental changes, or errors in the sensor itself. Data noise leads to increased errors during model training, affecting the effect of machine learning models. To reduce the impact of noise, this paper adopts two common denoising methods: wavelet transform and Kalman filtering.

By decomposing the signal into sub-signals of different scales, wavelet transform can effectively remove high-frequency noise and retain the low-frequency characteristics of the signal. In the specific operation, the collected original signal is firstly decomposed by a wavelet, and an appropriate wavelet basis and decomposition layer number are selected. Then, the coefficients of each layer are thresholded to remove unimportant high-frequency noise, and finally, the processed signal is reconstructed by inverse transform. Kalman filtering is a recursive filtering method based on the state space model, which is suitable for processing systems with Gaussian noise. In hydraulic systems, Kalman filtering can filter out noise in measurement data by estimating the optimal value of the current state. Based on the system's dynamic model and measurement model, the Kalman filter continuously iterates and updates the state estimate to obtain a smoother signal. Combining these two denoising methods can effectively retain the key dynamic characteristics of the system while removing noise.

In Fig. 1, the horizontal axis represents the distribution of each frequency component in the signal, and the vertical axis represents the intensity of different frequency components. In the spectrum of the original signal, the amplitude varies greatly, which is mainly caused by noise. These noises affect the signal's stability and accuracy. In the signal spectrum, after wavelet transform denoising, the high-frequency noise is reduced, and the noise is effectively filtered.

Wavelet transform performs well in removing high-frequency noise. It can effectively remove high-frequency interference in the signal and retain the main low-frequency components of the signal. It is suitable for those scenarios dominated by high-frequency noise. Kalman filtering dynamically adjusts the signal through a recursive state estimation method to remove noise while maintaining the stability and dynamic characteristics of the system. It is especially advantageous in systems with complex dynamic changes. It is more stable in removing random noise and is suitable for complex signals containing multiple types of noise.

Fig. 1 Noise spectrum analysis



In the hydraulic system's control process, the measurement values of different sensors have different dimensions and ranges. For example, the dimension of pressure is Pa; the flow unit is liters/minute; and the temperature is Celsius. If data of different scales are not standardized, the training effect of the subsequent machine learning model is affected, resulting in the model being unable to effectively learn the intrinsic relationship in the data. Therefore, in the data processing stage, all sensor data are first normalized to convert data of different dimensions into the same numerical range. The commonly used normalization method is maximum and minimum normalization; that is, by scaling the data according to its minimum and maximum values, it is ensured that the value range of all data is between [0, 1].

In addition, considering that the machine learning model is sensitive to the mean and variance of the data, the standardization method is used to convert the data into a standard normal distribution with a mean of 0 and a variance of 1. In the specific operation, the mean and standard deviation of the data are calculated, and the mean is subtracted from each data point and then divided by the standard deviation. This operation can ensure that the results of different features are not affected by dimensional differences when training the model and can accelerate the training process and improve the convergence speed.

4 Model construction and training

4.1 Nonlinear system modeling

This paper designs a multilayer deep neural network model to model the hydraulic system's nonlinear characteristics precisely. The network's input layer contains multiple neurons, corresponding to the hydraulic system's key input parameters, such as pump speed, valve opening, flow, and pressure. Since the hydraulic system usually involves multiple input variables, the number of neurons in the input layer is dynamically adjusted according to the number of selected features.

The network's hidden layer adopts a multilayer perceptron structure, and the number of neurons in each layer is selected by cross-validation optimization. To deal with the highly nonlinear characteristics of the hydraulic system, this paper uses the ReLU (rectified linear unit) activation function as the activation function of the hidden layer. ReLU has a strong nonlinear expression ability and can avoid the gradient vanishing problem, which is suitable for deep neural networks. The network structure design considers both computing resources and training efficiency and reasonably selects the number of hidden layers and the number of neurons in each layer. Finally, the output layer uses a linear activation function to output the prediction results of the hydraulic system according to the requirements of the regression task, usually physical quantities such as displacement, speed, or pressure of the hydraulic system. For the intelligent control task of hydraulic systems, choosing a suitable network structure depends on the specific application requirements. For example, if there is an obvious temporal sequence and long-term dependency in the system data, LSTM or Transformer may be a more appropriate choice. If the system task is relatively simple and the amount of data is not large, MLP or CNN may be effective enough. Combining CNN or Transformer for multi-sensor data processing and long-time series modeling can improve control accuracy and system response speed for more complex hydraulic systems.

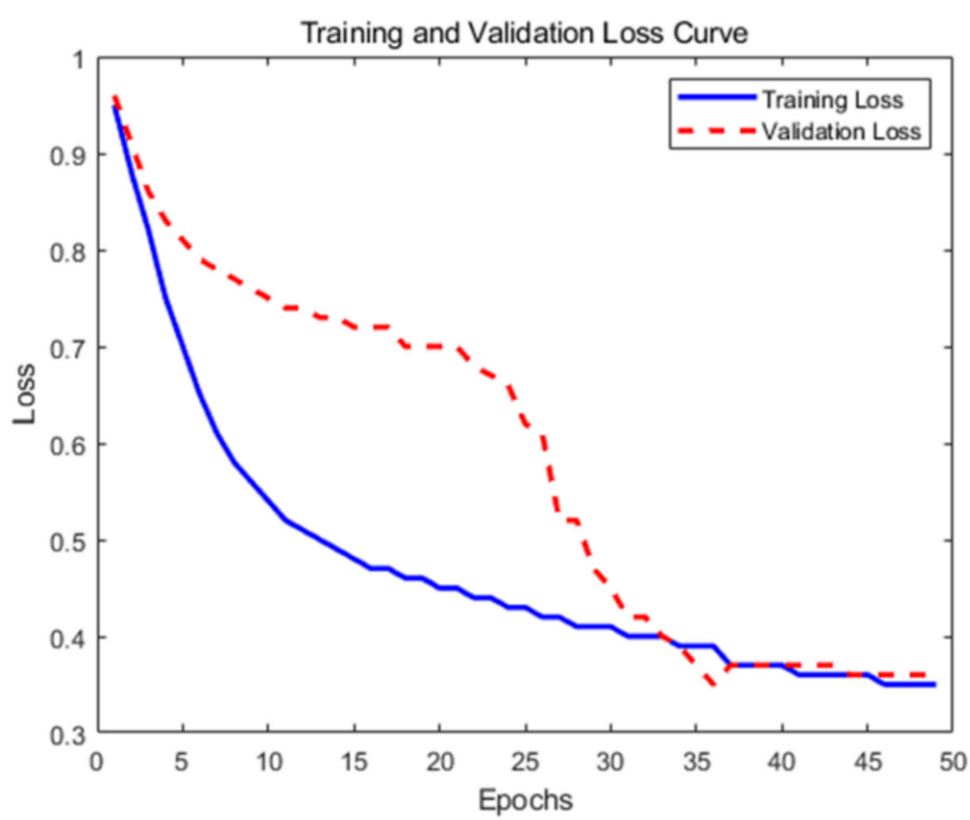
The network's training data come from the hydraulic system's historical operation data, including input–output pairs under different working conditions. To ensure the effectiveness of model training, the dataset is cleaned, denoised, and standardized through the previous data preprocessing steps. The training data include the input parameters of the hydraulic system under various working conditions (such as pump speed, flow, and pressure) and the corresponding outputs (such as the position and speed of the actuator), which are obtained through real-time monitoring and experiments.

During the model training process, the batch gradient descent method is used to optimize the weights and biases of the neural network. The model's weight is updated in each training cycle by calculating the error between the model output and the true value. The optimization goal is to minimize the mean square error; that is, by minimizing the difference between the predicted value and the true value, the model parameters are adjusted so that the network can precisely predict the hydraulic system's behaviors. The learning rate is set to 0.001, and the batch size is set to 32. To prevent overfitting, the dropout technique is used in the training process to randomly discard a part of neurons in each layer to reduce the model's dependence on training data and improve the model's generalization ability. In addition, L2 regularization is used to constrain the size of the weights to reduce the risk of overfitting further.

In Fig. 2, as the training cycle increases, the training set loss decreases, indicating that the model's performance on the training set is getting better and better, gradually reducing the difference between the prediction and the true value. In the initial training cycle, the validation set loss also decreases, indicating that the model's performance on the validation set is also gradually improving. However, as the training progresses, the validation set loss may drop to a certain minimum point and then begin to rise. This indicates that the model has begun to overfit the training data; that is, the model performs better and better on the training set, but begins to perform worse on unseen data (validation set).

In the process of deep neural network modeling, the physical limitations and safety requirements of the hydraulic system are explicitly incorporated into the model design. The parameter settings of the model input layer strictly follow the adjustable range of the actual system, such as the upper and lower limit thresholds of the pump speed, the mechanical constraints of the valve opening, and the safety margin of the actuator displacement. The network output layer ensures that the predicted value is always within the physical range allowed by the system through the activation function boundary constraints. During the training phase, the control actions that exceed the safety threshold are dynamically suppressed by introducing penalty terms to prevent the model from generating infeasible control strategies. At the same time, the system monitors the abnormal changes of key parameters (such as pressure peaks and temperature fluctuations) in real time and performs secondary verification of control instructions in combination with the preset safety protocol to ensure the strong coupling between model decisions and the safe operation of the system.

Fig. 2 Loss change



During the model training process, all key hyperparameter configurations and training results are archived in a structured table format. The table fields include the initial learning rate, batch size, regularization coefficient, number of training cycles, loss function convergence trajectory, and validation set error distribution. The hardware configuration, sensor noise level, and load fluctuation range of each training are synchronously recorded as metadata to ensure the reproducibility of experimental conditions. Training logs are stored in categories through version management tools and support filtering historical records by timestamp or performance indicators. Statistical analysis and visualization tools based on tabular data can quickly locate the optimal parameter combination, identify overfitting or underfitting patterns, and provide a quantitative basis for model iteration.

4.2 Adaptive control and optimization

In reinforcement learning, environment modeling is crucial. The control environment of the hydraulic system is formed by defining the state space, action space, and reward function. In this study, the state space is defined as the real-time state variables of the hydraulic system, including pump speed, flow, pressure, temperature, and the position of the actuator. The system state is composed of quantitative data fed back by the sensor in real time, and each state represents the hydraulic system's working state at a certain moment.

The action space includes the system's controllable parameters, such as the pump's output power, the valve's opening, and the control signal's adjustment amount. According to the current state, the agent can choose an action, that is, to adjust these control parameters to affect the behavior of the hydraulic system. The reward function is designed as a criterion for measuring the quality of the control strategy and is usually associated with the performance indicators of the hydraulic system (such as control accuracy, energy efficiency, and response time). In this study, the reward function comprehensively considers the hydraulic system's control accuracy, response speed, and energy efficiency. For example, the system gives positive rewards when a set goal is achieved (such as a desired execution position) and negative rewards when the response is delayed or overshoots. The agent can learn to adopt the optimal control strategy under different working conditions by continuously optimizing the reward function.

After each training, the validation set is used to evaluate the model performance, and the model parameters are adjusted by calculating the error of the validation set. The early stopping strategy is used during the training process to prevent overfitting; that is, when the error of the validation set no longer decreases in multiple consecutive cycles, the training is stopped early. Q-learning is a basic algorithm in reinforcement learning. A Q value table is constructed to represent the expected reward obtained by performing a certain action in a certain state. The core of Q-learning is that by updating the estimate of each state–action pair, the intelligent agent can gradually learn the optimal strategy. The update of the Q value follows the following formula:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)\right] \tag{1}$$

Among them, s_t is the current state; a_t is the current action; r_t is the reward obtained in the current step; γ is the discount factor; and α is the learning rate. The Q value gradually converges through multiple interactions with the environment, and finally, the optimal state–action strategy can be obtained. In the application of hydraulic systems, the Q-learning algorithm continuously explores different control actions and adjusts the control strategy according to the feedback reward. In the initial stage, the agent acquires more experience through random exploration continuously adjusts control parameters (such as pump speed and valve opening) during the training process and finally achieves optimal control.

A major limitation of Q-learning is that when the state space and action space are too large, the storage and update of the Q value table become unrealistic. To solve this problem, this paper uses a deep Q network to replace the traditional Q-learning method. DQN (Deep Q Network) combines the advantages of deep learning, and Q-learning approximates the Q value function through neural networks, and avoids the construction and storage problems of the Q value table. DQN uses a deep neural network to predict the state–action value function (Q value) in the specific implementation. The network's input is the state of the hydraulic system (such as pump speed, flow, and pressure), and the output is the Q value of all possible actions. By continuously interacting with the environment, the agent stores the historical state–action–reward sequence in the experience pool through the experience replay mechanism and trains by randomly extracting samples to avoid the correlation between data.

During the training of the reinforcement learning strategy, the dynamic adjustment mechanism of the strategy parameters is implemented based on the experimental results of real-time feedback. The learning rate and discount factor are adaptively updated according to the loss convergence trend and Q value stability of the training cycle to avoid premature convergence or oscillation. The exploration rate gradually decays with the progress of training. The initial stage focuses on extensive exploration to cover the state space, and the later stage focuses on fine-tuning the strategy to improve control accuracy. The reward function weight is calibrated in real time according to the needs of multi-objective optimization, and the priority of each goal is dynamically balanced based on the statistical distribution of historical control errors, energy efficiency ratios, and response times. During the training process, the system periodically evaluates the performance of the strategy. If the verification set error or safety index deviates from the preset threshold, the parameter reset and re-optimization process is triggered to ensure that the strategy iteration always evolves toward the global optimal direction.

In addition, DQN adopts the update mechanism of the target network and the target Q value to further stabilize the training process. The target network is a copy of the Q network and is updated at fixed intervals to avoid the overestimation problem that occurs during the Q value update process. In this way, DQN can be trained efficiently in complex environments.

The training process of DQN is as follows: The Q network and the target Q network are initialized, and the initial learning rate, discount factor, experienced pool size, and other hyperparameters are set. The agent interacts with the hydraulic system environment, performs actions, and obtains feedback (state, reward, next state), and this information is stored in the experience pool. A batch of samples is randomly drawn from the experience pool, and the target Q value is calculated through the target Q network to update the Q network's parameters. The parameters of the target Q network are updated at fixed intervals.

The horizontal axis of Fig. 3 represents the pump speed of the hydraulic system, which is usually within a certain range, such as 0 to 100 RPM. The vertical axis represents the pressure of the hydraulic system, which is usually in the range of 0 to 200 MPa. Through continuous training and updating of Q values, the agent gradually learns which state–action pairs have higher Q values, and then forms the optimal control strategy. The strategy corresponding to the action with a higher Q value is preferred by the agent. During the training process, the agent updates the Q value according to the interaction with the environment. After multiple iterations, the Q value gradually converges to a stable value, reflecting the agent's choice of the optimal control strategy for the hydraulic system.

Through the training of Q-learning and DQN algorithms, the control strategy of the hydraulic system is gradually optimized. The optimization goal not only includes minimizing the control error (such as position error and speed error) but also considers multi-objective factors such as response speed and energy efficiency. To comprehensively consider these goals, this paper designs a comprehensive reward function, which weights factors such as control accuracy, system response speed, and energy efficiency, so that the hydraulic system can achieve adaptive control under the multi-objective optimization framework.

Experimental results show that the reinforcement learning method shows significant advantages over traditional control methods (such as PID control) in hydraulic system control. Under conditions with strong nonlinearity and uncertainty, reinforcement learning can automatically adjust control parameters to achieve more precise control and faster response speed. For example, in the positioning control of hydraulic actuators, traditional PID control methods may have large overshoot or oscillation, while the control strategy optimized by reinforcement learning can effectively suppress these phenomena and achieve smooth response and precise positioning.

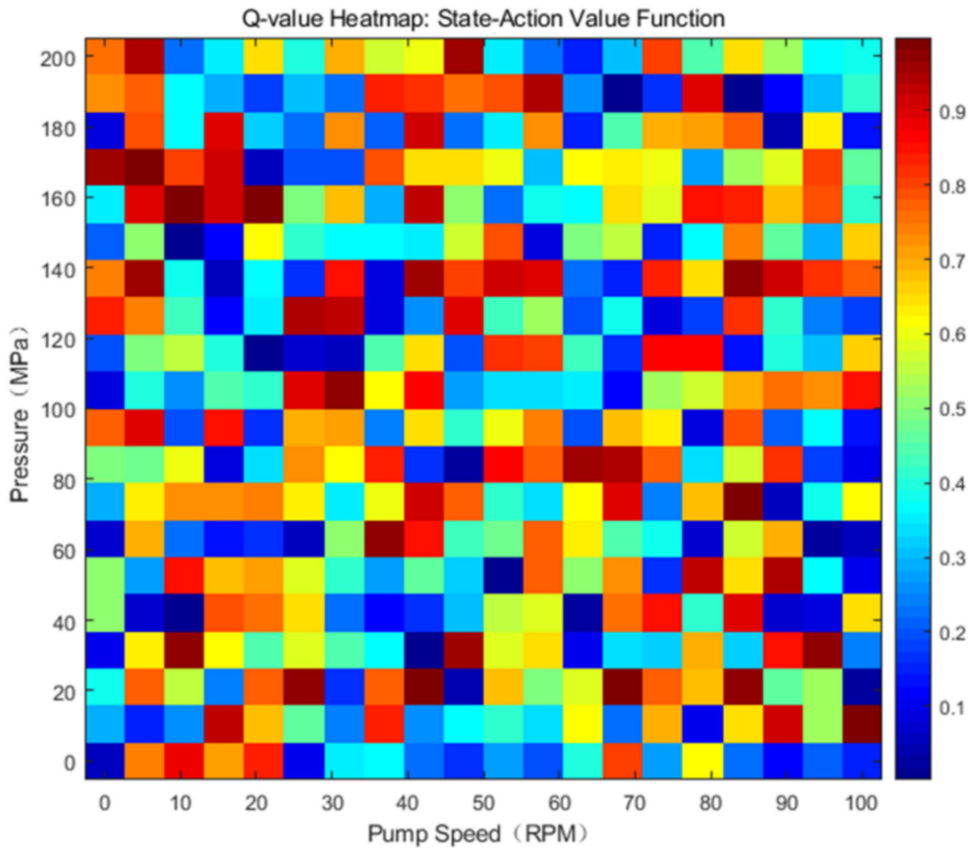


Fig. 3 Q value function heatmap

The reinforcement learning method makes the hydraulic system have a strong adaptive ability by constantly interacting with the environment and adjusting the strategy through feedback. Under complex conditions such as different load changes, temperature fluctuations, and unstable hydraulic flow, reinforcement learning can adaptively adjust the control strategy to maintain the system's stability and efficiency. In addition, by applying deep neural networks, DQN effectively improves the learning ability in high-dimensional and complex state spaces, enabling the hydraulic system to maintain good control performance in more complex operating environments.

4.3 Multi-objective optimization

The optimization goal of the hydraulic system is not only to reduce positioning error or improve response speed but also to minimize energy consumption while ensuring control accuracy. Therefore, in the reinforcement learning framework, multiple goals must be considered comprehensively to avoid the bias caused by single-objective optimization. To this end, the following objective functions are designed. (1) Control accuracy goal: this goal mainly considers the precise positioning of the hydraulic system's actuators (such as pistons and valves). The control accuracy error can be quantified by the error between the target position and the actual position. (2) Response speed goal: this goal focuses on the response time of the hydraulic system. In particular, when the system needs to adjust the working state quickly, the system adjustment speed must be as fast as possible to avoid system performance degradation due to excessive response time. (3) Energy efficiency goal: this goal mainly measures the energy consumed by the hydraulic system in the process of completing the control task. Minimizing energy consumption can not only improve system efficiency but also reduce operating costs, which has important practical value.

For these goals, a weighted comprehensive reward function is constructed to balance the impact of different goals through reasonable weighting coefficients. The reward function is calculated by the following formula:

$$R = w_1 \cdot R_{accuracy} + w_2 \cdot R_{speed} + w_3 \cdot R_{efficiency} \tag{2}$$

Among them, $R_{accuracy}$, R_{speed}, and $R_{efficiency}$ represent the reward values of control accuracy, response speed, and energy efficiency, respectively; w_1, w_2, and w_3 are the weight coefficients of the goals, which are adjusted according to the actual application requirements to ensure the balance of each goal. In this way, the agent can automatically adjust the control strategy when performing the control task to achieve the optimization of multiple goals.

When designing the reward function, it is necessary to ensure that the reward function can precisely reflect the performance of the hydraulic system under different working conditions and can prompt the agent to learn the correct control strategy. Specifically, each sub-goal in the reward function needs to be dynamically adjusted according to the system's real-time state and target state to give the agent a reasonable feedback signal during the training process.

The control accuracy can be quantified by the position error. For example, assuming that the target position is P_{target} and the current position of the actuator is $P_{current}$, the position error is $\epsilon = \left|P_{target} - P_{current}\right|$. The accuracy reward function is designed to be proportional to the inverse of the error. The smaller the error, the greater the reward:

$$R_{accuracy} = -\alpha \cdot \epsilon \tag{3}$$

Among them, α is the weight coefficient of the accuracy reward. The smaller the control accuracy error, the higher the reward. Based on this feedback, the agent adjusts the control strategy to reduce the error. Response speed reward: the response speed is measured by the time delay, that is, the time required from receiving the control command to the actuator reaching the target state. The response speed reward is designed to be proportional to the negative value of the response time:

$$R_{speed} = -\beta \cdot T \tag{4}$$

Among them, T is the response time and β is the weight coefficient of the response speed reward. The shorter the response time, the higher the reward. Energy efficiency reward: energy efficiency is represented by the power consumption of the hydraulic system. To balance energy efficiency and control accuracy, the energy efficiency reward function is designed to be inversely proportional to the power consumption per unit time. The specific form is:

$$R\left(P_{input}\right) = \frac{k}{P_{input} + \epsilon} \tag{5}$$

Among them, k is a constant used to adjust the reward scale to ensure that the reward function is appropriate when optimizing the system. ϵ is a small constant used to avoid division by zero when the power consumption is zero and to ensure that the reward function is still effective when the power consumption is very low.

In the reinforcement learning framework, the agent obtains the reward signal at each time step through interaction with the environment and adjusts the control strategy based on the reward feedback. To deal with the multi-objective optimization problem, this paper uses the weighted sum method; that is, the reward signals of multiple objectives are weighted and summed to obtain a comprehensive reward value for the agent to make decisions. During the training process, the agent continuously optimizes the decision-making strategy to balance various objectives, thereby gaining optimal hydraulic system control.

https://doi.org/10.1007/s00521-025-11430-z

5 Effect evaluation

5.1 Accuracy evaluation

The core of control accuracy evaluation is to calculate the error between the system's output and the expected output. To accurately measure the error, the root-mean-square error (RMSE) and the mean absolute error (MAE) are selected as quantitative indicators of control accuracy.

The root-mean-square error calculation formula is:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{6}$$

Among them, y_i is the expected output; $\hat{y}_i$ is the actual output; and N is the number of data points. RMSE can penalize large errors and is suitable for application scenarios requiring higher accuracy.

The mean absolute error calculation formula is:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{7}$$

MAE measures the average level of error. Unlike RMSE, it does not penalize large errors. Therefore, in some cases where large errors are more tolerant, MAE can provide more intuitive performance feedback.

By calculating RMSE and MAE, the hydraulic system's control accuracy can be comprehensively evaluated. The higher the accuracy, the smaller the RMSE and MAE values, indicating that the hydraulic system can follow the expected trajectory more accurately.

The horizontal axis of the left sub-graph of Fig. 4 represents the control strategy, and the vertical axis represents the accuracy and recall. As can be seen from the figure, the strategy of this paper performs the best, with both high accuracy and very high recall. The horizontal axis of the right sub-graph of Fig. 4 also represents 5 different control strategies, and the vertical axis represents the error value of control accuracy, revealing the significant advantages of the control strategy in the control accuracy of the hydraulic system. The control strategy of this paper performs best in RMSE and MAE, significantly exceeding traditional PID control, reinforcement

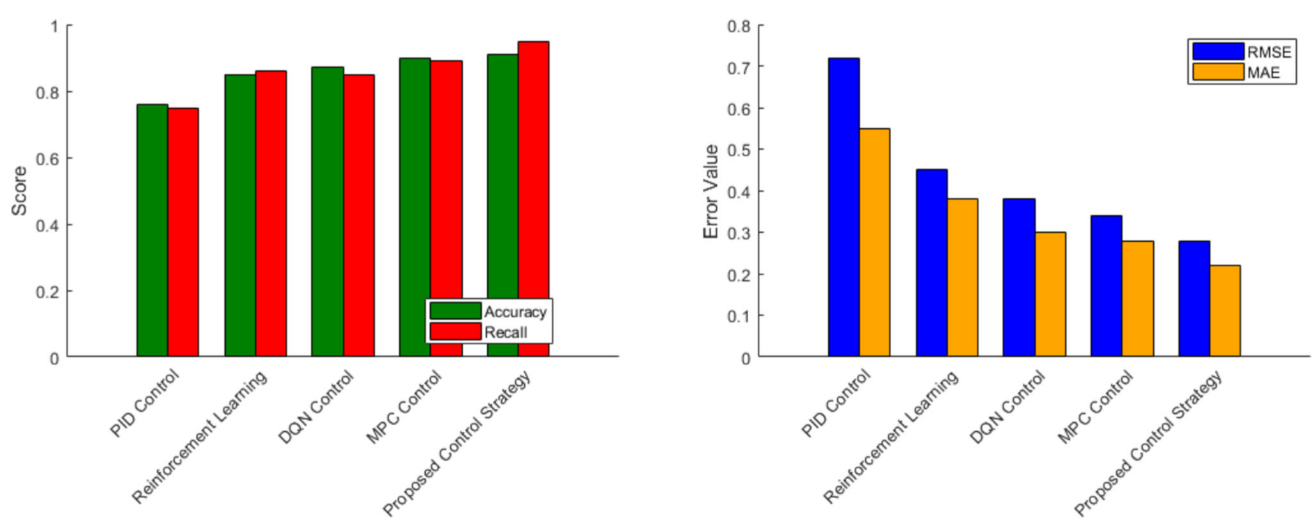


Fig. 4 System control accuracy

learning, DQN control, and MPC (model predictive control) control. The error of PID control is large; reinforcement learning control and DQN control have improved, but there is still a gap in accuracy; and MPC control has higher accuracy, but it is still insufficient compared with the strategy of this paper. The high-precision control requirements of hydraulic systems require more advanced control strategies to meet them. The control strategy of this paper achieves more precise control through optimization algorithms, proving its effectiveness and advantages in hydraulic system control.

5.2 Response speed

The response speed evaluation measures the time required for the system to stabilize the output result from the input signal change. Specifically, the response speed is evaluated by the following steps: recording the time required from the input signal change (such as the change of valve opening or pump speed) to the system output (such as piston position or flow) reaching a stable state. To further quantify the response characteristics, the system's rise time (the time from the initial state to the target value) and stabilization time (when the system output fluctuation range is within the allowable error range) are also calculated. These time domain characteristics can effectively describe the dynamic response capability of the hydraulic system under different operating modes, as shown in Table 2.

Table 2 shows the time required from the input signal change to the system output stabilization. The control strategy in this paper shows the best response capability, with short rise time, stabilization time, and response time. The response time is 2.4 s, which is better than other comparison strategies.

5.3 Energy efficiency evaluation

Energy efficiency evaluation mainly quantifies the system's energy efficiency by calculating the ratio of energy consumption to output power of the hydraulic system. First, the input power (P_{in}) and output power (P_{out}) of the hydraulic system are measured. The input power is usually provided by the motor, and the output power is closely related to the power consumption of the hydraulic actuator. The energy efficiency is calculated by the following formula:

$$\text{Efficiency} = \frac{P_{out}}{P_{in}} \tag{8}$$

Among them, P_{out} is the hydraulic system and P_{in} is the system's input power. The higher the energy efficiency of the system, the more work it can complete while consuming less energy, thus achieving the purpose of energy saving.

Table 2 Response time

Control strategy	Input signal change (valve opening/pump speed)	Rise time (seconds)	Stabilization time (seconds)	Response time (seconds)
PID control	Valve opening change 10%	1.2	3.5	4
Reinforcement learning	Pump speed change 20%	1	2.8	3
DQN control	Valve opening change 15%	0.9	2.5	2.8
Model predictive control (MPC)	Pump speed change 30%	1.1	2.2	3.1
proposed control strategy	Valve opening change 5%	0.8	2	2.4

https://doi.org/10.1007/s00521-025-11430-z

The energy consumption ratio to the hydraulic system's output power (energy/power ratio) is calculated. The lower the ratio, the higher the system's energy efficiency and the lower the energy consumption. The calculation formula for this ratio is:

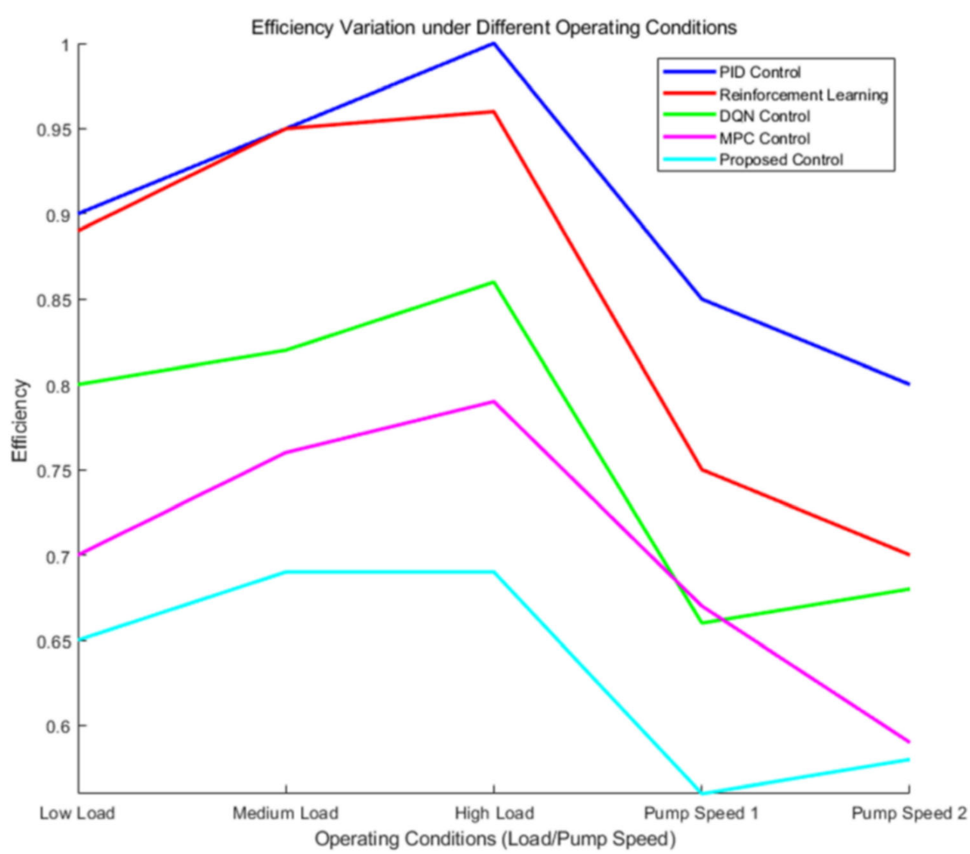
$$\text{Energy/Power Ratio} = \frac{\text{Total Energy Consumption}}{\text{Total Output Power}} \tag{9}$$

The horizontal axis of Fig. 5 shows the response of the hydraulic system under different working conditions, and the vertical axis represents the energy efficiency ratio. The lower the energy efficiency ratio, the more efficient the system is and the lower the energy loss is. As the load increases from low load to high load, the energy efficiency ratios of all control strategies change to a certain extent, but the control strategy in this paper still shows the best performance under various working conditions. Under the condition of pump speed change, the energy efficiency ratio of traditional PID control is the worst, while the energy efficiency ratios of reinforcement learning, DQN, and MPC control strategies are relatively low, indicating the advantages of these methods in dynamic adjustment. The control strategy in this paper also maintains the lowest energy efficiency ratio when the pump speed changes, which is lower than 0.7 under high load, indicating that it has better adaptability and higher energy efficiency under various working conditions.

5.4 Stability evaluation

The stability evaluation aims to verify the response efficiency of the hydraulic system in the face of load changes or external disturbances and ensure that it can maintain a stable operating state in various scenarios. The specific evaluation process is as follows:

Fig. 5 Energy efficiency evaluation



By simulating the behavior of the hydraulic system under different load changes or external disturbances, the output fluctuation of the system is recorded in detail. These fluctuations are analyzed in depth to evaluate the system's stability performance. In addition to evaluating the response speed, stability also involves the ability of the system to quickly return to a stable state after a load change or disturbance. The time required for the system to stabilize the output from the start of the disturbance is recorded, and the system's robustness is evaluated based on the length of the recovery time. At the same time, the output fluctuation amplitude of the system after the disturbance is also an important indicator for measuring its stability. A smaller fluctuation amplitude indicates that the system has a stronger stability in the face of disturbances, as shown in Table 3.

Table 3 shows the output fluctuations. The control strategy of this paper performs best in terms of stability, with the shortest recovery time (2.0 s) and the smallest output fluctuation amplitude (2.5%). This shows that the control method proposed in this paper can provide higher system stability when facing external disturbances or load changes. PID control performs poorly, with the longest recovery time (3.5 s) and the largest output fluctuation amplitude (5.2%), indicating that the traditional control strategy is inefficient in dealing with disturbances. The recovery time based on model predictive control is short (3.0 s), but the output fluctuation amplitude is large, and the performance is not as good as the control strategy proposed in this paper. Reinforcement learning control and DQN control also perform well in terms of recovery time and fluctuation amplitude, but the control strategy of this paper still provides the strongest stability.

The difference in the impact of different control strategies on the performance of hydraulic systems stems from the degree of coupling between their underlying optimization mechanisms and the dynamic characteristics of the system. Traditional PID control relies on linear model assumptions. Under nonlinear time-varying conditions, overshoot and oscillation occur due to fixed parameters, and its steady-state error is positively correlated with load mutations. The reinforcement learning strategy dynamically adjusts the pump speed and valve opening to compensate for the nonlinear hysteresis effect of the system and reduce the interference of pressure fluctuations on the positioning accuracy of the actuator. The DQN algorithm uses the high-dimensional state representation capability of the deep network to accurately map the spatiotemporal correlation of multi-sensor data and effectively suppress parameter perturbations caused by temperature drift. Although model predictive control can improve response speed based on rolling optimization, its reliance on accurate system models can easily lead to decreased stability due to the accumulation of modeling errors under complex coupling conditions.

6 Conclusions

This paper proposes an intelligent control and optimization framework based on deep neural networks, which successfully solves the problems of low precision, slow response, and poor adaptability in traditional hydraulic system control. Through multi-sensor data acquisition and preprocessing, nonlinear modeling by deep neural networks, and reinforcement learning optimization control strategy, this paper significantly improves the hydraulic system's control accuracy, response speed, and energy efficiency. Based on multi-objective

Table 3 System stability

Control strategy	Disturbance type	Disturbance magnitude	Recovery time (seconds)	Output fluctuation amplitude (%)
PID control	Load change	20%	3.5	5.2
Reinforcement learning control	External disturbance	15%	2.8	3.9
DQN control	Load change	25%	2.5	3.2
Model predictive control (MPC)	External disturbance	20%	3	4
Proposed control strategy	Load change	15%	2	2.5

optimization, the system can adaptively adjust control parameters to achieve optimal control under complex working conditions.

Experimental results show that the proposed method is superior to the traditional PID control method and can effectively improve system performance. By combining deep learning and reinforcement learning for hydraulic systems, the study promotes the practical application of emerging technologies in traditional industries, which helps to accelerate technological progress and industrial upgrading. However, the research in this paper still has certain limitations, such as high data requirements during model training, and further adjustments and optimizations may be required under certain extreme working conditions. Future research can further explore algorithms with higher data efficiency, enhanced real-time learning capabilities, and adaptive expansion under various working conditions, to make the intelligent control and optimization of hydraulic systems more complete and practical.

Author's contribution All authors contribute to this study. W. H. and Y. J. contributed to work concept or design and writing; Y. J. was involved in the data collection and methodology; and W. H. was involved in supervision and visualization.

Funding This study did not receive any funding in any form.

Data availability statement Data are available upon reasonable request.

Declarations

Conflict of interests The authors have no relevant financial or non-financial interests to disclose.

Consent to publish All authors in the manuscript are aware of the submission and subsequent publication of the manuscript.

References

1. Yao Z, Liang X, Jiang GP et al (2022) Model-based reinforcement learning control of electrohydraulic position servo systems. IEEE/ASME Trans Mechatron 28(3):1446–1455
2. Yao Z, Xu F, Jiang GP et al (2023) Data-driven control of hydraulic manipulators by reinforcement learning. IEEE/ASME Trans Mechatron 29(4):2673–2684
3. Yuan X, Wang Y, Zhang R et al (2022) Reinforcement learning control of hydraulic servo system based on TD3 algorithm. Machines 10(12):1244
4. Egli P, Hutter M (2022) A general approach for the automation of hydraulic excavator arms using reinforcement learning. IEEE Robot Automation Letters 7(2):5679–5686
5. Li Z, Bai L, Tian W et al (2023) Online control of the raw water system of a high-sediment river based on deep reinforcement learning. Water 15(6):1131
6. Nasrabadi M, Mehri Y, Ghassemi A et al (2021) Predicting submerged hydraulic jump characteristics using machine learning methods. Water Supply 21(8):4180–4194
7. Araya SN, Ghezzehei TA (2019) Using machine learning for prediction of saturated hydraulic conductivity and its sensitivity to soil structural perturbations. Water Resour Res 55(7):5715–5737
8. Kotlar AM, Iversen BV, de Jong Van Lier, Q. (2019) Evaluation of parametric and nonparametric machine-learning techniques for prediction of saturated and near-saturated hydraulic conductivity. Vadose Zone J 18(1):1–13
9. Nevo S, Morin E, Gerzi Rosenthal A et al (2022) Flood forecasting with machine learning models in an operational framework. Hydrol Earth Syst Sci 26(15):4013–4032
10. Goliatt L, Sulaiman SO, Khedher KM et al (2021) Estimation of natural streams longitudinal dispersion coefficient using hybrid evolutionary machine learning model. Eng Appl Comput Fluid Mech 15(1):1298–1320
11. Angelaki A, Singh Nain S, Singh V et al (2021) Estimation of models for cumulative infiltration of soil using machine learning methods. ISH J Hydraul Eng 27(2):162–169
12. Wang Y, Shen H, Wu J et al (2023) Reinforcement-learning-based composite optimal control for looper hydraulic servo systems in hot strip rolling. IEEE/ASME Trans Mechatron 28(5):2495–2504

https://doi.org/10.1007/s00521-025-11430-z

13. Belfadil A, Modesto D, Meseguer J et al (2024) Leveraging deep reinforcement learning for water distribution systems with large action spaces and uncertainties: DRL-EPANET for pressure control. J Water Resour Plan Manag 150(2):04023076
14. Hu S, Gao J, Zhong D et al (2023) Real-time scheduling of pumps in water distribution systems based on exploration-enhanced deep reinforcement learning. Systems 11(2):56
15. Bowes BD, Wang C, Ercan MB et al (2022) Reinforcement learning-based real-time control of coastal urban stormwater systems to mitigate flooding and improve water quality. Environ Sci: Water Res Technol 8(10):2065–2086
16. Tang X, Chen J, Qin Y et al (2024) Reinforcement learning-based energy management for hybrid power systems: state-of-the-art survey, review, and perspectives. Chinese J Mech Eng 37(1):43
17. Backman S, Lindmark D, Bodin K et al (2021) Continuous control of an underground loader using deep reinforcement learning. Machines 9(10):216
18. Samtani P, Leiva F, Ruiz-del-Solar J (2023) Learning to break rocks with deep reinforcement learning. IEEE Robot Automation Letters 8(2):1077–1084
19. Tian W, Xin K, Zhang Z et al (2023) State selection and cost estimation for deep reinforcement learning-based real-time control of urban drainage system. Water 15(8):1528
20. Narayanan VL (2024) Reinforcement learning in wind energy-a review. Int J Green Energy 21(9):1945–1968
21. Hu C, Wang Q, Gong W et al (2022) Multi-objective deep reinforcement learning for emergency scheduling in a water distribution network. Memet Comput 14(2):211–223
22. Blad C, Bøgh S, Kallesøe C (2021) A multi-agent reinforcement learning approach to price and comfort optimization in HVAC-systems. Energies 14(22):7491
23. Gulzar B, Ahmad Sofi S, Sholla S (2024) Cognitive transformation in personal IoT: pioneering intelligent automation. Cyber-Phys Syst 11(2):183–240

Authors and Affiliations

Wenjing Hu[1] · Yuejing Jiang[1]

✉ Yuejing Jiang
jiangyuejing@dlou.edu.cn

Wenjing Hu
huwenjing@dlou.edu.cn

[1] College of Applied Technology, Dalian Ocean University, Dalian 116300, Liaoning, China