MDPI

Article

Automatic Casting Control Method of Continuous Casting Based on Improved Soft Actor–Critic Algorithm

Xiaojun Wu [1,*], Wenze Jiang [1], Sheng Yuan [1,*], Hongjia Kang [1], Qi Gao [2,*] and Jinzhou Mi [2]

1   School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China; jwzddd@163.com (W.J.)
2   China National Heavy Machinery Research Institute Co., Ltd., Xi'an 710016, China
*   Correspondence: xiaojunw@xjtu.edu.cn (X.W.); ysheng@xjtu.edu.cn (S.Y.); xzs_gq@163.com (Q.G.)

Abstract: Continuous casting production is an important stage in smelting high-quality steel, and automatic casting control based on artificial intelligence is a key technology to improve the continuous casting process and the product quality. By controlling the opening degree of the stopper rod reasonably, the mold can be filled with liquid steel stably in the specified time window, and automatic casting can be realized. In this paper, an automatic casting control method of continuous casting based on an improved Soft Actor–Critic (SAC) algorithm is proposed. Firstly, a relational model of the stopper rod opening degree and the liquid steel outflow velocity is established according to historical casting data. Then the Markov Decision Process (MDP) model of the automatic casting problem and the reinforcement learning framework based on the SAC algorithm are established. Finally, a Heterogeneous Experience Pool (HEP) is introduced to improve the SAC algorithm. According to the simulation results, the proposed algorithm can predict the stopper rod opening degree sequence under the constraint of the target liquid level curve. Under different billet specifications and interference conditions, an accuracy of 80% of liquid level in the mold and a stopper rod opening degree stability rate of 75% can be achieved, which is 4.29% and 3.17% higher than those for the baseline algorithms, respectively.

Keywords: continuous casting; automatic casting; liquid level control; stopper rod opening degree control; Soft Actor–Critic; experience pool

1. Introduction

Continuous casting (continuous steel casting) is the process of casting, cooling and slitting of high-temperature liquid steel through a continuous casting machine to obtain billets. As a bridge between steelmaking and rolling, the stability of the continuous casting production not only affects the efficiency of steelmaking tasks but also relates to the quality and cleanliness of the steel produced. The continuous casting machine is mainly composed of devices such as the ladle, tundish, stopper rod, mold, pulling machine, second cooling equipment and cutting equipment.

The tundish is a buffer that feeds the casting line and regulates the supply of liquid steel to the mold. In addition, the tundish ensures that liquid steel is continuously fed to the process, especially while the ladles switch between empty and full [1].

The mold, as the core equipment of the continuous casting machine, is of great importance to the quality of the billet. The liquid steel in the tundish flows into the mold through the stopper rod control, and the liquid steel is initially cooled and solidified in the mold to form the billet shell, which is pulled out from the bottom by the pulling machine. Good control of the liquid level in the mold can effectively avoid the problems of steel leakage, steel overflow and slag rolling, ensuring production safety and avoiding defects on the surface of or inside the billet.

At present, the liquid level in the mold is mainly controlled by controlling the inflow from the tundish through the stopper rod or baffle. The opening degree of the stopper

rod or the position of the baffle is changed according to the liquid level in the mold, so as to change the liquid steel outflow velocity and stabilize the liquid level in the mold. The control of the liquid level in the mold in the complete casting process consists of two main stages: The first is the rapid filling stage of the liquid steel before the start of the pulling machine, and the second is the stable casting stage of the liquid steel after the start of the pulling machine. In the first stage, due to the limitations of the hardware and environment, the sensor cannot accurately detect the liquid level in the mold and, therefore, cannot control the liquid level in the mold in a closed-loop manner, so automatic control in this stage has become a key issue for artificial intelligence in the process of continuous casting.

The liquid steel in the tundish is an incompressible thermal fluid, and its outflow velocity is not only related to the opening degree of the stopper rod or the position of the baffle but also related to the weight of the liquid steel in the tundish, because the weight of the liquid steel can affect the pressure at the outlet, thus affecting the outflow velocity of the liquid steel. The relationship between the velocity and the pressure fields of the liquid steel at the outlet is complex [2], so this paper establishes the relationship model between the opening degree of stopper rod and the liquid steel outflow velocity by using the real casting data for a stopper-rod-type continuous casting machine. This model introduces the weight parameter of the tundish to capture the influence of the liquid steel pressure on the outflow velocity.

The numerical approach is used to simulate the flow of liquid steel under different opening degrees of the stopper rod. The most popular numerical approach is to assume a steady-state, single-phase flow using a Reynold's Averaged Navier–Stokes (RANS) method, together with a turbulence model such as k-ϵ or k-ω [3]. The numerical solutions are regarded as approximate, and they can be accurate in some cases [4]. Therefore, the accuracy of the proposed model can be verified by comparing the results of the numerical simulation and the relational model calculations [5].

Reinforcement learning, as an important branch in the field of machine learning [6], is considered one of the core technologies leading to strong artificial intelligence. It is based on an interactive learning mechanism [7] and does not require precise mathematical equations. It emphasizes that agents learn in interaction with the environment, with good adaptability and robustness [8], and has achieved excellent results in many fields such as quantitative trading, robot control and games [9]. Automatic casting is a complex process related to liquid steel flow, and it is extremely difficult to accurately describe the relationship between the target liquid level, the liquid steel outflow velocity and the stopper rod opening degree. Therefore, in this paper, an automatic casting control method based on reinforcement learning is proposed, which takes advantage of the neural network and reinforcement learning mechanisms to predict the optimal stopper rod opening degree sequence to achieve stable control of the liquid level in the mold.

The contributions of this work are three folds.

(1) A novel relational model between the stopper rod opening degree and the liquid steel outflow velocity is constructed, which can calculate the liquid steel outflow velocity corresponding to different opening degrees of the stopper rod. Compared with previous works, this model can more accurately reflect the relationship between variables by using real casting data.

(2) A control framework for automatic casting based on reinforcement learning is constructed, which can predict the stopper rod opening degree sequence under the constraint of the target liquid level curve during casting. Compared with previous works, this control framework can automatically adapt to different billet specifications and environmental interference without complicated manual calculations.

(3) A novel reinforcement learning algorithm is proposed to improve the performance of the control framework, which can improve the liquid level accuracy in the mold and the stability of the stopper rod opening degree. Compared with the baseline algorithm, the proposed algorithm emphasizes the importance of different samples and optimizes the training process of neural networks.

The remainder of this paper is organized as follows: Section 2 reviews the related work on the control of the liquid level in the mold. Section 3 presents the model of the automatic casting problem. Section 4 presents the automatic casting control framework and the improved reinforcement learning algorithm. Section 5 presents the simulation experiments and analyzes the results. Section 6 gives the conclusions of this paper.

2. Related Work

Automatic casting is a technology used to achieve a stable rise of the liquid level in the mold by controlling the stopper rod or baffle intelligently. It is one of the most important issues in the field of continuous casting and has a direct impact on the quality of the billet. During the development of continuous casting technology, many classical methods for the control of the liquid level in the mold emerged, such as predictive control, fuzzy control, adaptive control, etc.

In [10], a Hammerstein Generalized Predictive Controller (HGPC) was designed, and the performance of the controller under different types of environmental disturbances was verified by simulation. Reference [11] designed a Generalized Predictive Controller (GPC) that modeled white noise as colored noise, which improved the robustness of the controller and took into account the system integrity characteristics in modeling to avoid complex calculations in parameter identification.

Reference [12] designed a fuzzy Proportional–Integral–Derivative (PID) controller with nonlinear compensation, which effectively suppressed fluctuations of the liquid level in the mold under abnormal operating conditions and achieved a better anti-disturbance capability.

Reference [13] designed an adaptive control framework, established an internal mode observer to predict the periodic component of the signal and then established an adaptive control paradigm to effectively reduce the dynamic bulging problem in continuous casting.

However, there are still some challenges in the control of the liquid level in the mold, mainly including inaccurate sensor detection, system delays and limited understanding of its dynamic processes from the control perspective, in addition to disturbances caused by uncertainties such as clogging of the tundish outlet port [14] and argon flow [15,16].

Reinforcement learning has good nonlinear fitting ability and independent decision-making ability [17]. The Soft Actor–Critic (SAC) algorithm [18], as one of the representative algorithms of reinforcement learning in recent years, has achieved the best effect on a series of continuous control benchmark tasks [19]. Therefore, it is applied in this paper for automatic casting control of continuous casting, so that the stopper rod opening degree can respond to different environmental states to achieve stable control of the liquid level in the mold.

3. Problem Modeling

Automatic casting technology means that the stopper rod automatically completes the filling of liquid steel in the mold in the initial stage of continuous casting, making the process of raising the liquid level in the mold fully controllable and reaching the start level of the pulling machine within the specified time window. The setting of the target liquid level curve was in accordance with reference [20], but in order to better understand the process of raising the liquid level in the mold, the specific liquid level value is abstracted in this paper, and the final curve is shown in Figure 1. The period from 0 to t_5 is the automatic casting process. The liquid level in the mold reaches the start level h_2 of the pulling machine at t_5 and then switches to closed-loop control and reaches the stable casting level at t_6.

The sensor of the liquid level in the mold cannot detect a level below h_1 due to hardware limitation and high-temperature steam, so it is necessary to establish a relational model between the stopper rod opening degree and the liquid steel outflow velocity and calculate the stopper rod opening degree at different stages of casting according to the model. Then, the stopper rod opening degree sequence is put into the first-level system (i.e., direct control system of continuous casting machine) of the production line to realize automatic casting. Equation (1) is the model of the liquid level variation in the mold. The

liquid level in the mold is related to the normal inflow Q_{in} of liquid steel and the fluctuation Q_{err} under interference. Due to the existence of various disturbances and noises in the casting process, to simplify the model, the influence of all disturbances on the liquid level in the mold was set as Q_{err}.

$$\begin{cases} Q_{in} = F(p) \\ \frac{dH}{dt} = \frac{Q_{in}(t) - Q_{err}(t)}{S} \end{cases} \tag{1}$$
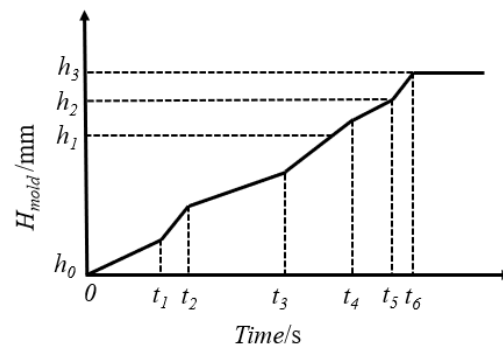
Figure 1. The target liquid level curve of the mold.

As shown in Equation (1), p is the stopper rod opening degree, H is the liquid level in the mold and S is the cross-sectional area of the billet.

In order to establish the relationship F between the stopper rod opening degree p and the liquid steel inflow Q_{in}, this paper used the historical casting data of a steel mill to calculate the weight change relationship between the ladle and the tundish in the casting process, as shown in Equation (2) and indirectly established the relationship F according to the law of conservation of mass. Specifically, the liquid steel in the ladle flows into the tundish at uniform velocity through the fixed opening degree water outlet, and the liquid steel in the tundish flows into the mold at a nonuniform velocity through the stopper rod control. The liquid level in the mold rises gradually. In the casting data used in this paper, the sensor sampling interval for the stopper rod opening degree in association with the weight of the ladle and the tundish was 0.5 s. According to the law of conservation of mass, the difference between the decrease in the weight of the ladle and the increase in the weight of the tundish every 0.5 s is the weight of the liquid steel flowing into the mold from the tundish at that stopper rod opening degree.

$$\begin{cases} \Delta W_{big} = W_{big}^{t} - W_{big}^{t+1} \\ \Delta W_{mid} = W_{mid}^{t+1} - W_{mid}^{t} \\ \text{s.t.} \Delta W_{big} > \Delta W_{mid} > 0 \\ Q_{in} = \frac{\Delta W_{big} - \Delta W_{mid}}{\rho} \\ p \rightarrow Q_{in} \end{cases} \tag{2}$$

As shown in Equation (2), W_{big} is the weight of the ladle, W_{mid} is the weight of the tundish and ρ is the density of the liquid steel.

Considering that the weight of the liquid steel in the tundish also affects the flow velocity at the outlet of the stopper rod, the heavier the liquid steel is, the faster the outflow velocity of the liquid steel can be under the same stopper rod opening degree. Therefore, in order to model the actual casting situation more accurately, the parameter of the tundish weight was introduced based on Equation (1), and the correction coefficient k was added, which was determined by fitting with the historical data. The final equation of liquid level variation in the mold is shown in Equation (3).

$$\frac{dH}{dt} = \frac{Q_{in}(t) \cdot W_{mid}^{t} \cdot k^{-1} - Q_{err}(t)}{S} \tag{3}$$

As shown in Equation (3), assuming the opening degree of stopper rod at time t is p, the volume of liquid steel flowing into the mold from the tundish Q_{in} can be calculated according to Equation (2) under this stopper rod opening. W_{mid}^{t} is the weight of liquid steel in the tundish at time t, which can be directly obtained from the detection data of the weight sensor. Q_{err} refers to the fluctuating volume of liquid steel in the mold caused by environmental interference at time t, which is usually complicated and difficult to predict in real casting. The purpose of setting W_{mid}^{t} in this paper was only to verify the robustness of the proposed method, so its specific value was manually set.

When solving Equation (3), an initial value of k was first given, and then the change of the liquid level in the mold dH was constantly calculated until the end of the casting. The liquid level in the mold was recorded at this time and compared with the value detected by the sensor of liquid level in the mold. If the difference is large, it indicates that the value of k is unreasonable. The value of k was constantly adjusted until the difference was within the allowable range. At this time, Equation (3) can simulate the change process of the liquid level in the mold in the casting process and solve the problem that the sensor of liquid level in the mold cannot work normally in the first stage.

4. Automatic Casting Control Method

4.1. Reinforcement Learning Modeling

Reinforcement learning is the product of combining cognitive science and computational intelligence [21]. By interacting with the environment to learn knowledge, the agent can effectively explore the high-dimensional continuous space [22] and finally make decisions. To clearly characterize the interaction process, reinforcement learning introduces the Markov Decision Process (MDP), which involves three basic elements: state, action and reward. In this paper, an MDP model was established for the automatic casting control problem, as shown in Figure 2.
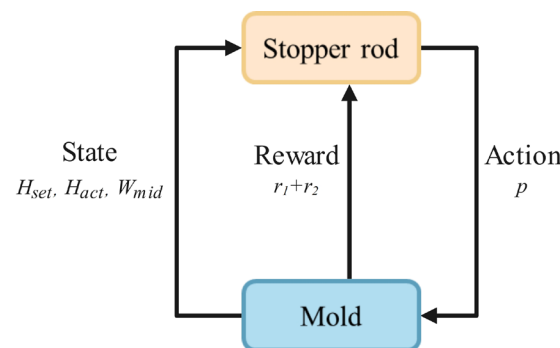
Figure 2. Automatic casting control MDP model.

In order to use the reinforcement learning algorithm, the state space, action space and reward function of automatic casting control were defined in detail based on the MDP model.

State space $S = \{s = [H_{set}, H_{act}, W_{mid}]\}$, where H_{set}, H_{act} and W_{mid} represent the target liquid level in the mold, the actual liquid level in the mold and the weight of the tundish, respectively.

Action space $A = \{a = [p]\}$, where p represents the stopper rod opening degree, which can be any value in the range of 0–30 mm. $p = 0$ indicates that the stopper rod is not opened, and the liquid steel in the tundish cannot flow into the mold. $p = 30$ indicates that the stopper rod is fully opened, and the liquid steel in the tundish can flow into the mold at full speed.

Reward function $R = \{r = [r_1 + r_2]\}$, where r_1 is 0 or 1, indicating the reward obtained by the agent at the end of the control task. $r_1 = 1$ indicates that the agent has completed the task, that is, the stopper rod successfully controls the liquid level in the mold to reach the starting level. In other cases, $r_1 = 0$ will be given. r_2 represents the real-time reward

obtained by the agent during the control process, which is related to the liquid level error $H_{err} = H_{set} - H_{act}$ in the mold. The specific calculation equation is shown in Equation (4).

$$r_2 = \begin{cases} 1, H_{err} \leq 1 \\ 0.1, 1 < H_{err} \leq 5 \\ -0.5, 5 < H_{err} \leq 10 \\ -1, H_{err} \geq 10 \end{cases} \tag{4}$$

The selection of an appropriate reinforcement learning algorithm is crucial to the final control effect. There are complex environmental disturbances in continuous casting production, such as the fluctuation and turbulence of liquid steel in the mold or the vibration of the stopper rod and the mold, so the robustness and generalization ability of the strategy must be high. The SAC algorithm uses a stochastic strategy instead of a deterministic strategy [23] and greatly improves exploration efficiency and training stability by introducing a maximum-entropy objective [24]. Moreover, it is less sensitive to hyperparameters, so the SAC was chosen as the core algorithm of the model.

The state space of the automatic casting control task is a three-dimensional vector, and the dimension is small. In this paper, the policy network and the critic network of the algorithm use a multilayer, fully connected neural network. The input of the policy network is the state, and the output is the action probability distribution function [25]. The input of the critic network is the state, and the output is the value of the state. The optimal number of hidden layers and neurons was selected through experiments. ReLU was used as the activation function.

4.2. Improved SAC Algorithm

The SAC algorithm uses an experience replay mechanism to cache all samples generated by the agent during exploration in the experience pool, including the samples of liquid level fluctuation or stability in the mold and the samples of the stopper rod opening degree vibration or stability. By default, all samples have the same weight [26], and samples are extracted from the experience pool in a completely random way to train the network. However, these samples have different influences on the network training process. The random method ignores the importance and differences of the samples. Therefore, this paper proposes to introduce an additional Heterogeneous Experience Pool (HEP) to cache the samples of the stopper rod opening degree vibration, in order to enhance network training using such samples to reduce unnecessary vibration during the stopper rod movement, thus improving the stability of the continuous casting production. This algorithm is named HEP-SAC in this paper, and its pseudocode is shown in Algorithm 1.

Algorithm 1. Automatic casting control algorithm based on HEP-SAC.

Input: billet width, billet thickness, target liquid level curve.
Output: neural network parameters.
Initialize the experience pool and set the minibatch size.
Initialize the parameters of the neural network.
for each iteration, do
   for each environment step, do
      generate action a_t based on state s_t.
      execute action a_t, generate sample $x_1 = \{s_t, a_t, r_t, s_{t+1}\}$.
      store x_1 in the default experience pool.
      calculate the vibration amplitude of the stopper rod $R_s = a_t - a_{t-1}$.
      if $R_s >$ threshold value,
         store $x_2 = [\{s_{t-1}, a_{t-1}, r_{t-1}, s_t\}\{s_t, a_t, r_t, s_{t+1}\}]$ in the HEP.
      end if
   end for
   $N_1 = $ minibatch$\cdot\eta\cdot2^{-1}$
   $N_2 = $ minibatch$\cdot(1 - \eta)$
   if x_2 quantity $> N_1$ and x_1 quantity $> N_2$

```
    for each gradient step, do
        sampling from HEP and default experience pool.
        update neural network parameters.
    end for
  end if
end for
```

5. Experiment and Analysis

5.1. Experimental Environment and Parameter Setting

In order to verify the availability of the relational model between the stopper rod opening degree and the liquid steel outflow velocity, Ansys Fluent software [27] was used to establish the fluid domain related to the stopper rod in a continuous casting machine. Ansys Fluent is a general-purpose Computational Fluid Dynamics (CFD) software used to model fluid flow, heat and mass transfer, chemical reactions and more. In this paper, Ansys Fluent was used to realize the whole process of numerical simulation from preprocessing to postprocessing, including the establishment of the model and the grid, the setting of the boundary conditions the and solver and the visualization of the results [28].

As shown in Figure 3, the blue side is the inlet boundary, that is, the upper surface of the liquid steel in the tundish. The inlet speed was calculated according to the outflow velocity of the ladle liquid steel. The red side is the outlet boundary, that is, the bottom surface of the outlet pipe.
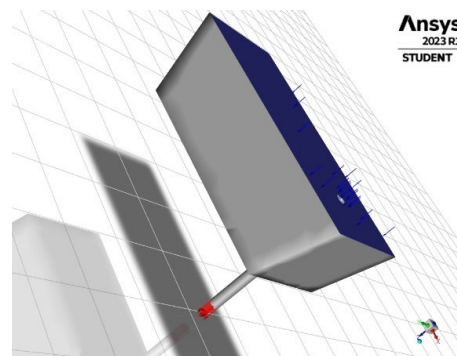
Figure 3. Three-dimensional model diagram of the fluid domain.

In this paper, the Application Programming Interface (API) provided by OpenAI Gym was used to implement the environment function. The TensorFlow framework was used to implement the reinforcement learning algorithm. The algorithm interacts with the environment to complete the training. The software and hardware configurations related to the training platform are shown in Table 1.

Table 1. The detailed configurations of the hardware and software.

Hardware	Configuration	Software	Configuration
CPU	Ryzen 7 4800U	Operating System	Windows 11
RAM	Micron 16G	IDE	PyCharm 2021
Motherboard	LNVNB161216	Python	3.6.2
SSD	SAMSUNG 512G	TensorFlow	2.1.0
GPU	Radeon Graphics	Gym	0.21.0

The environment of the reinforcement learning is important for training and learning [29]. The environment settings in the MDP model included the following: The billet width was 1540 mm, and the billet thickness was 230 mm. The rising process of the target liquid level in the mold included five stages, and the liquid level reached the target height at a uniform speed within the specified time window of each stage. The duration of the

five stages was 7 s, 3 s, 20 s, 20 s and 25 s, respectively. The rise of the liquid steel level in the five stages was 30 mm, 20 mm, 200 mm, 64 mm and 100 mm, respectively.

In order to verify the performance of the HEP-SAC algorithm, this paper compares it with the Twin-Delayed Deep Deterministic Policy Gradient (TD3) and the SAC algorithm, and analyzes its training performance and control performance. The same hyperparameters of all algorithms included the following: The number of training rounds was 350, the number of steps in each round was 150, the batch size was 128, the number of exploration steps was 1000, the number of network updates was 3, the learning rate was 3×10^{-4}, the discount factor was 0.95 and the size of the default experience pool was 5×10^5. The unique hyperparameter settings of each algorithm are shown in Table 2.

Table 2. The hyperparameters of each algorithm.

Parameter	TD3	SAC	HEP-SAC
Exploring noise standard deviation	0.3	—	—
Policy noise standard deviation	0.4	—	—
Delay update frequency	3.0	—	—
Learning rate of alpha	—	3×10^{-4}	3×10^{-4}
Size of HEP	—	—	1×10^6
Sampling ratio	—	—	0.2

The HEP-SAC algorithm uses an additional HEP to cache two ordinary samples at adjacent time steps, so the size of the HEP should be twice that of the original experience pool to prevent sample overflow. The sampling ratio between the HEP and the original experience pool should be determined by experimentation. A large η can lead to a large number of vibration samples collected and reduce the training of the neural network for the accurate liquid level samples, and a small η cannot meet the training of the neural network for the vibration samples. All other parameters of the HEP-SAC algorithm should be the same as in the original SAC algorithm.

In order to quantify the performance of the control framework, two evaluation indexes were set in this paper, namely Automatic Casting Probability (ACP) and Stable Stopper Rod Opening Degree Probability (SSP). The specific calculation is shown in Equation (5). ACP was used to measure the deviation between the liquid level in the mold and the target liquid level curve. The moment when the liquid level error is within the set threshold is called the accurate liquid level moment. The proportion of the accurate liquid level moments in the total casting time is the ACP of the framework under the current casting task. The SSP was used to measure the stability of the stopper rod in the process of movement. The moment when the vibration amplitude of the stopper rod is within the set threshold is called the stable opening degree moment. The proportion of the stable opening degree moment in the total casting time is the SSP of the framework under the current casting task.

$$\begin{cases} ACP = \frac{\sum_{i=0}^{T} H_{err}(t) < 2}{T} \\ SSP = \frac{\sum_{i=0}^{T} R_s(t) < 3}{T} \end{cases} \tag{5}$$

As shown in Equation (5), H_{err} is the liquid level error in the mold, T is the time steps of the automatic casting and R_s is the vibration amplitude of the stopper rod.

The smaller the ACP is, the more the liquid level in the mold deviates from the target liquid level curve, which not only easily causes slag inclusion, and thus affects the quality of billet, but also easily causes steel overflow, and thus poses a safety threat to the workers on site. The smaller the SSP is, the more unnecessary vibration exists in the opening degree sequence of stopper rod, which can easily cause irreversible structural loss to the equipment. Therefore, the larger the two evaluation indexes, the better the performance of the control framework.

5.2. Availability Analysis of the Stopper Rod Flow Control Model

Figure 4 shows the level fitting results of the relational model between the stopper rod opening degree and the liquid steel outflow velocity under different *k*, where the height of the mold was 900 mm and the insertion depth of the pulling machine was 350 mm. IF a position is set indicating that the liquid level in the mold is 0 and this position is 200 mm from the top of the mold, the simulated level changes from −350 mm, and the level enters the effective detection range of the sensor at about 95 s. According to the variation of the level in the figure, the error between the simulated level and the detected level can be controlled to within ±5 mm until 110 s. Compared with the total increase in the level, the ratio is less than 2%, so the relational model proposed in this paper can meet the accuracy requirements. Considering the complexity of the liquid steel flow and the diversity of external disturbances, the value of *k* is suggested to be in the range of 300–310.
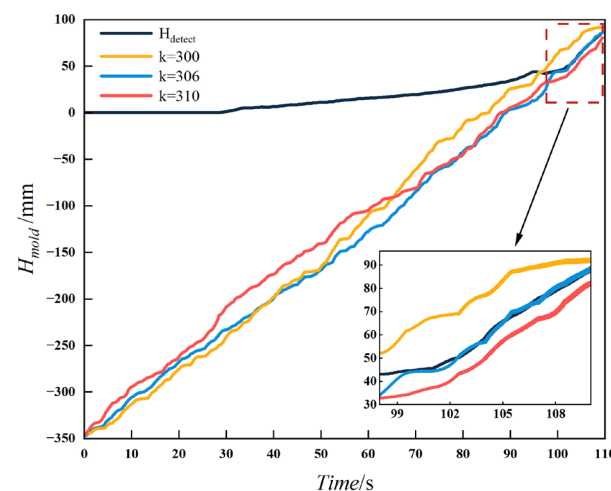
Figure 4. Simulated liquid level in the mold at different *k*.

In order to further verify the availability of the relational model, Ansys Fluent software was used to conduct fluid simulation experiments. Figure 5a is the fluid trace diagram. It can be seen that the liquid steel in the tundish flowed rapidly to the stopper rod and carried out complex flow around it. In this paper, the section 2 mm down from the top of the outlet pipe was taken as the base surface, and the outflow velocity distribution on this surface is shown in Figure 5b. It can be seen that the outflow velocity of the liquid steel gradually decreased from the center point to the pipe wall.
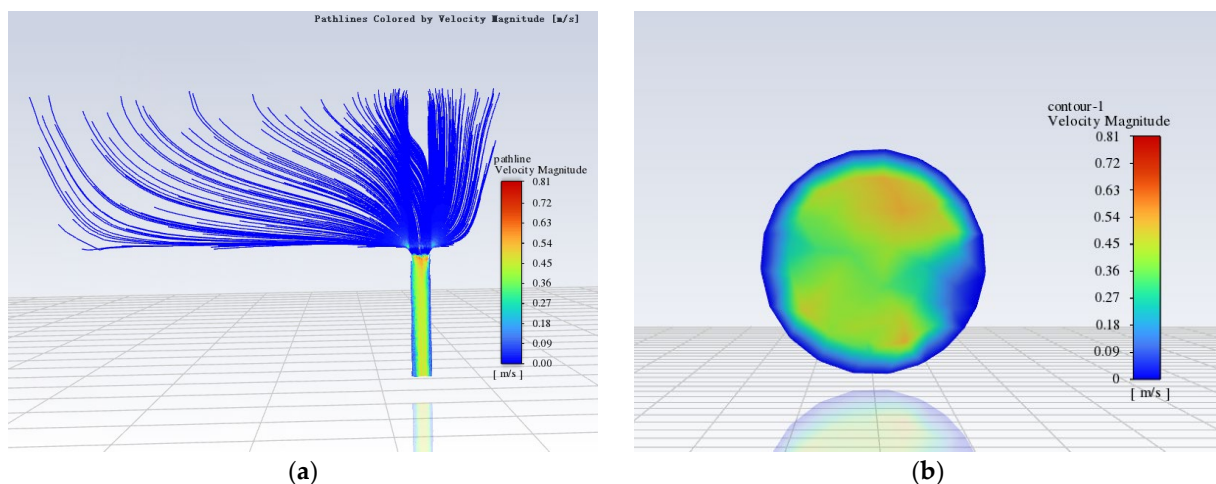
Figure 5. Simulation result diagram. (**a**) Fluid trace diagram; (**b**) fluid outflow velocity nephogram.

In order to quantify the simulation results, V_d was set to represent the difference in outflow velocity between the model and the simulation calculation. It was assumed that the radius of the outlet pipe was 35 mm, the billet width was 1000 mm, and the billet thickness was 230 mm. The liquid level difference H_d in the mold caused by V_d within 1 s under different stopper rod opening degrees was calculated using Equation (6). The results are shown in Table 3. The average absolute level difference was 3.094 mm. Compared with the total level change value of nearly 400 mm during the casting process, it is in the acceptable range. Therefore, the relational model between the stopper rod opening degree and the liquid steel outflow velocity can simulate the liquid level variation in the mold accurately.

$$H_d = \frac{\pi \cdot R^2 \cdot V_d}{S_w \cdot S_t} \tag{6}$$

**Table 3.** Comparison of calculation results between model and simulation.

p/mm	V_d/m·s^{-1}	H_d/mm	p/mm	V_d/m·s^{-1}	H_d/mm
3	0.006	1.003	18	−0.032	−5.351
6	0.007	1.171	21	−0.014	−2.341
9	−0.014	−2.341	24	−0.048	−8.027
12	−0.052	−8.696	27	0.021	3.512
15	0.023	3.846	30	−0.042	−7.024

As shown in Equation (6), R is the radius of the outlet pipe, S_w is the billet width and S_t is the billet thickness.

It is worth noting that, in this model, the outflow velocity of the liquid steel was calculated on the base surface of 2 mm down from the top of the outlet pipe, so the time delay in the process of the liquid steel flowing into the mold was ignored. Moreover, the liquid level fluctuation in the mold was ignored, because the turbulent flow of liquid steel is extremely complicated and it is difficult to model all the characteristics. In addition, the model is not suitable for a baffle-type continuous casting machine.

5.3. Convergence Performance Analysis of HEP-SAC Algorithm

Figure 6 shows the training process of TD3, SAC and HEP-SAC under the same set of random seeds. It can be seen that TD3 algorithm had the fastest learning speed, but the reward value after convergence was the lowest. This is because the deterministic strategy easily causes agents to fall into the overfitting state. The convergence speed and final performance of the HEP-SAC algorithm exceeded that of the baseline algorithms. It shows that the samples stored in the HEP can guide the training of neural networks correctly.
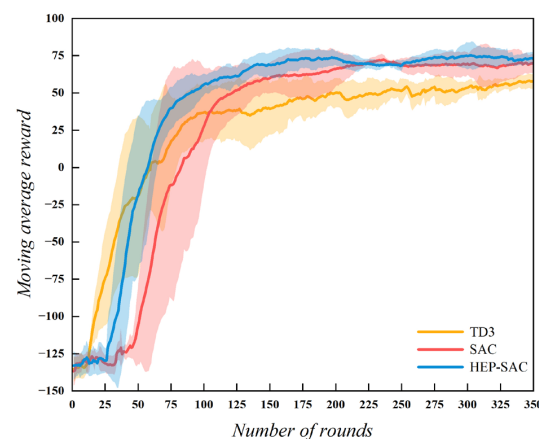

Figure 6. Training process diagram of different algorithms.

5.4. Performance Analysis of Control Framework Based on HEP-SAC

The performance of the control framework was mainly verified by ACP and SSP. The specification of the billet was set as 1540 mm × 230 mm, and two different types of perturbations were added: the first was the sudden rise of the liquid level in the mold at 15 s, and the second was the sudden drop of the liquid level in the mold at 45 s. The level change caused by both disturbances was set at 20 mm to verify the robustness of the control framework. Figure 7a,b shows the time-varying curves of the liquid level error in the mold and the stopper rod opening degree in this task. It can be seen that the proposed method had a smaller liquid level error in the mold and a more stable control sequence under normal and disturbed conditions, and its performance exceeded that of the baseline algorithms.

The performance quantification results of the control framework are shown in Tables 4 and 5. Different billet widths were set, and the billet thickness was uniformly set at 230 mm. The results show that the ACP of the proposed method reached more than 80% under different casting tasks, and the control accuracy improved by 4.29% on average compared with the baseline algorithms. The SSP reached more than 75%, and the control stability improved by 3.17% on average compared with the baseline algorithms.
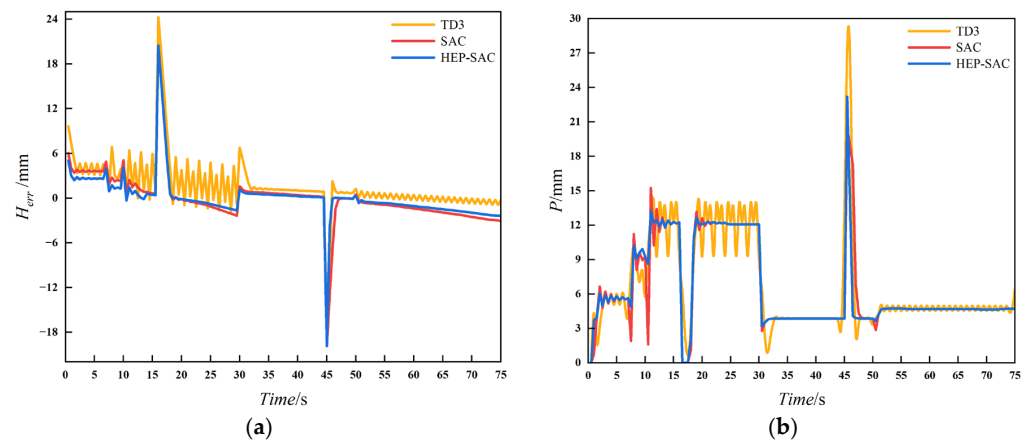
Figure 7. Control effect diagram of different algorithms. (**a**) Curves of the liquid level error in the mold; (**b**) curves of the stopper rod opening degree.

Table 4. ACP of each algorithm in different casting tasks.

Width/mm	No Disturbance			Added Disturbance		
	TD3	SAC	HEP-SAC	TD3	SAC	HEP-SAC
1100	72.0	82.4	85.1	68.0	78.5	81.2
1200	79.2	83.0	86.8	70.1	77.3	82.4
1300	81.2	91.4	96.5	76.3	87.0	91.7
1400	83.5	90.2	94.5	79.2	86.1	89.3
1540	80.4	86.5	92.6	76.4	79.6	86.4
1640	76.3	83.5	89.5	72.3	74.5	84.1

Table 5. SSP of each algorithm in different casting tasks.

Width/mm	No Disturbance			Added Disturbance		
	TD3	SAC	HEP-SAC	TD3	SAC	HEP-SAC
1100	68.7	86.0	89.2	61.8	74.8	78.5
1200	70.2	90.3	93.6	63.2	81.4	85.1
1300	66.9	89.7	93.9	60.2	80.1	84.5
1400	72.3	92.0	94.5	65.1	82.8	86.9
1540	74.4	86.5	90.1	67.0	77.9	79.1
1640	65.3	91.2	95.4	58.8	82.1	85.3

6. Conclusions

In this paper, the relational model between the stopper rod opening degree and the liquid steel outflow velocity was established by using historical casting data. The model used real detection data of the liquid steel weight to convert the outflow of the tundish. Through fitting with the historical casting data, a reasonable correction coefficient k was obtained. Compared with numerical simulation or physical simulation, the model can accurately reconstruct the real continuous casting process. Through the simulation experiment and comparative analysis, the model can accurately simulate the changing process of the liquid level in the mold. By modeling the automatic casting problem as the MDP model, reinforcement learning was used to establish the stopper rod opening degree control framework, and an HEP was introduced to improve the experience replay mechanism of the SAC algorithm. In the control framework, the stopper rod can automatically select the appropriate opening degree according to the difference between the actual liquid level and the target liquid level in the mold at the current moment and adjust the actual liquid level to keep approaching the target liquid level. The control framework not only automatically adapted to different billet specifications, and thus had good generalization, but also corrected the liquid level fluctuation caused by environmental interference in a timely manner, and thus had good robustness. In the HEP-SAC algorithm, an HEP was used to cache the samples of the opening degree vibration of stopper rod, so as to enhance the priority sampling of such samples. With the training of the control framework and the updating of the parameters of the neural network, the policy network not only ensured that the liquid level in the mold accurately tracked the target liquid level curve, but also output a more stable opening degree of the stopper rod during the transition stage of liquid level acceleration, reducing the equipment loss caused by vibration. The experimental results show that the convergence speed and performance of the HEP-SAC algorithm exceeded that of TD3 and SAC algorithm. Compared with the control framework based on baseline algorithms, the framework based on HEP-SAC improved the accuracy of the liquid level in the mold by 4.29% and the stability of the stopper rod opening degree by 3.17%, which shows the effectiveness of the improved method in this paper.

With the development of sensors and other hardware equipment, it will be possible to establish a more accurate relational model between the stopper rod opening degree and the liquid steel outflow velocity, so as to improve the reliability of the reinforcement learning control framework and promote the continuous progress of automatic casting technology.

Author Contributions: Conceptualization, X.W.; data curation, X.W., W.J., Q.G. and J.M.; formal analysis, X.W. and S.Y.; funding acquisition, Q.G.; investigation, H.K.; methodology, W.J.; project administration, S.Y.; resources, Q.G. and J.M.; software, W.J. and H.K.; validation, X.W., W.J. and S.Y.; visualization, X.W.; writing—original draft, W.J.; writing—review and editing, X.W., S.Y., Q.G. and J.M. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

Q_{in}	normal inflow of liquid steel in the mold.
p	stopper rod opening degree.
F	relationship between p and Q_{in}.
H	liquid level in the mold.

Q_{err}	fluctuation of liquid steel in the mold under interference.
S	cross-sectional area of the billet.
W_{big}	weight of liquid steel in the ladle.
W_{mid}	weight of liquid steel in the tundish.
ρ	density of the liquid steel.
K	correction coefficient of W_{mid}.
H_{set}	target liquid level in the mold.
H_{act}	actual liquid level in the mold.
r_1	reward obtained by the agent at the end of the control task.
r_2	reward obtained by the agent during the control task.
H_{err}	liquid level error in the mold.
R_s	vibration amplitude of the stopper rod opening degree.
T	time steps of the automatic casting.
R	radius of the outlet pipe.
V_d	difference of outflow velocity between the model and the simulation calculation.
S_w	billet width.
S_t	billet thickness.
H_d	liquid level difference in the mold caused by V_d within 1 s.

References

1. Birs, I.; Muresan, C.; Copot, D.; Ionescu, C. Model Identification and Control of Electromagnetic Actuation in Continuous Casting Process with Improved Quality. *IEEE/CAA J. Autom. Sin.* **2023**, *10*, 203–215. [CrossRef]
2. Rostamzadeh, A.; Razavi, S.E.; Mirsajedi, S.M. Towards multidimensional artificially characteristic-based scheme for incompressible thermo-fluid problems. *Mechanics* **2017**, *23*, 826–834. [CrossRef]
3. Thomas, B.G. Review on Modeling and Simulation of Continuous Casting. *Steel Res. Int.* **2018**, *89*, 1700312. [CrossRef]
4. Tebeta, R.T.; Fattahi, A.M.; Ahmed, N.A. Experimental and numerical study on HDPE/SWCNT nanocomposite elastic properties considering the processing techniques effect. *Microsyst. Technol.* **2020**, *26*, 2423–2441. [CrossRef]
5. Mramor, K.; Vertnik, R.; Šarler, B. Development of Three-Dimensional LES Based Meshless Model of Continuous Casting of Steel. *Metals* **2022**, *12*, 1750. [CrossRef]
6. Lyu, L.; Shen, Y.; Zhang, S. The Advance of Reinforcement Learning and Deep Reinforcement Learning. In Proceedings of the 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 25–27 February 2022; pp. 644–648. [CrossRef]
7. Shi, Z.; Ma, W.; Yin, S.; Zhang, H.; Zhao, X. Interactive Reinforcement Learning Strategy. In Proceedings of the 2021 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI), Atlanta, GA, USA, 18–21 October 2021; pp. 507–512. [CrossRef]
8. Yu, K.; Jin, K.; Deng, X. Review of Deep Reinforcement Learning. In Proceedings of the 2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 16–18 December 2022; pp. 41–48. [CrossRef]
9. Du, X.; Fuqian, X.; Hu, J.; Wang, Z.; Yang, D. Uprising E-sports Industry: Machine learning/AI improve in-game performance using deep reinforcement learning. In Proceedings of the 2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Chongqing, China, 8–11 November 2021; pp. 547–552. [CrossRef]
10. Sanchotene, F.; de Almeida, G.; Salles, J. Robust predictive controller of the mold level in a steel continuous casting process. In Proceedings of the 2011 9th IEEE International Conference on Control and Automation (ICCA), Santiago, Chile, 19–21 December 2011; pp. 1133–1138. [CrossRef]
11. Guo, G.; Li, W.; Wang, J.; Liu, M. A CGPC controller in continuous casting process. In Proceedings of the 3rd World Congress on Intelligent Control and Automation (Cat. No.00EX393), Hefei, China, 28 June–2 July 2000; Volume 4, pp. 2783–2786. [CrossRef]
12. Feng, Y.; Wu, M.; Chen, X.; Chen, L.; Du, S. A fuzzy PID controller with nonlinear compensation term for mold level of continuous casting process. *Inf. Sci.* **2020**, *539*, 487–503. [CrossRef]
13. Furtmüller, C.; Colaneri, P.; del Re, L. Adaptive robust stabilization of continuous casting. *Automatica* **2012**, *48*, 225–232. [CrossRef]
14. Thomas, B.G.; Bai, H. Tundish Nozzle Clogging-Aplication of Computational Models. In Proceedings of the 18th Process Thecnology Division Conference, Baltimore, MD, USA, 25–28 March 2001.
15. Thomas, B.G.; Huang, X.; Sussman, R.C. Simulation of Argon Gas Flow Effects in a Continuous Slab Caster. *Metal. Mater. Transp.* **1994**, *25*, 527–547. [CrossRef]
16. Liu, F.; Zhou, H.; Zhang, L.; Ren, C.; Zhang, J.; Ren, Y.; Chen, W. Effect of Temperature and Multichannel Stopper Rod on Bubbles in Water Model of a Steel Continuous Caster. *Steel Res. Int.* **2021**, *92*, 2100067. [CrossRef]
17. Zhang, Y.; Li, R.; Li, Y.; Zhang, T.; Zhuang, Y.; Song, Y. Novel TD3 Based AUV Path Tracking Control. In Proceedings of the 2021 China Automation Congress (CAC), Beijing, China, 22–24 October 2021; pp. 4945–4949. [CrossRef]

18. Haarnoja, T.; Zhou, A.; Abbeel, P.; Levine, S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *arXiv* **2018**, arXiv:1801.01290.

19. Banerjee, C.; Chen, Z.; Noman, N. Improved Soft Actor-Critic: Mixing Prioritized Off-Policy Samples with On-Policy Experiences. *IEEE Trans. Neural Networks Learn. Syst.* **2022**, 1–9. [CrossRef] [PubMed]

20. Mi, J.; Wang, X.; Shi, W.; Gao, Q.; Zhong, L. Application of mold automatic start casting control system for slab continuous casting production. *Heavy Mach.* **2021**, *363*, 24–28. [CrossRef]

21. Yan, L.; Liu, W.; Jiang, W.; Li, Y.; Li, R.; Hu, S. Deep Reinforcement Learning based Optimization of Battery Charging and Discharging Management for Data Center. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–22 July 2021; pp. 1–9. [CrossRef]

22. Cao, D.; Zhao, J.; Hu, W.; Ding, F.; Huang, Q.; Chen, Z. Distributed voltage regulation of active distribution system based on enhanced multi-agent deep reinforcement learning. *arXiv* **2020**, arXiv:2006.00546.

23. Li, P.; Wang, Y.; Gao, Z. Path Planning of Mobile Robot Based on Improved TD3 Algorithm. In Proceedings of the 2022 IEEE International Conference on Mechatronics and Automation (ICMA), Guilin, China, 7–10 August 2022; pp. 715–720. [CrossRef]

24. Zhao, F.J.; Zhou, Y. Wind Farm Maintenance Scheduling Using Soft Actor-Critic Deep Reinforcement Learning. In Proceedings of the 2022 Global Reliability and Prognostics and Health Management (PHM-Yantai), Yantai, China, 13–16 October 2022; pp. 1–6. [CrossRef]

25. Zhong, C.; Gursoy, M.C.; Velipasalar, S. Controlled Sensing and Anomaly Detection Via Soft Actor-Critic Reinforcement Learning. In Proceedings of the ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 23–27 May 2022; pp. 4198–4202. [CrossRef]

26. Li, Y.; Aghvami, A.H. Covertness-Aware Trajectory Design for UAV: A Multi-Step TD3-PER Solution. In Proceedings of the ICC 2022-IEEE International Conference on Communications, Seoul, Republic of Korea, 16–20 May 2022; pp. 7–12. [CrossRef]

27. Strativnov, E.; Gunochao, N. Investigation of Transient Boiling Regime of Water and Nanofluids Heated to Saturation Temperature Using CFD Simulation (ANSYS Fluent). In Proceedings of the 2022 IEEE 12th International Conference Nanomaterials: Applications & Properties (NAP), Krakow, Poland, 11–16 September 2022; pp. 1–4. [CrossRef]

28. Matsson, J.E. *An Introduction to ANSYS Fluent 2022*; SDC Publications: Mission, KS, USA, 2022.

29. Zheng, S. The influence of different environments on reinforcement learning. In Proceedings of the 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), Changchun, China, 20–22 May 2022; pp. 401–406. [CrossRef]