Available online at www.CivileJournal.org

Civil Engineering Journal

(E-ISSN: 2476-3055; ISSN: 2676-6957)

Vol. 12, No. 04, April, 2026

Adaptive Real-Time Strain-Rate Control in CRS Consolidation Testing Using SARSA Reinforcement Learning

Muhammad Fadhl 'Abbas [1], Hasbullah Nawir [1], Dimitri Mahayana [2*], Erza Rismantojo [1], Dayu Apoji [1], M. Alifsyah Putra Nasution [2], Targhib Ibrahim [2]

[1] *Faculty of Civil and Environmental Engineering, Bandung Institute of Technology, Bandung 40116, Indonesia.*

[2] *School of Electrical Engineering and Informatics, Bandung Institute of Technology, Bandung 40116, Indonesia.*

Received 30 November 2025; Revised 28 February 2026; Accepted 06 March 2026; Published 01 April 2026

Abstract

This study presents a reinforcement-learning framework for real-time strain-rate control in Constant Rate of Strain (CRS) consolidation testing to hasten the testing process using the SARSA algorithm. The controller adaptively adjusts deformation rate based on evolving pore-pressure ratio, with a reward strategy designed to maintain an average pore-pressure ratio near 30% to ensure partially drained conditions consistent with CRS theory. Two normally consolidated clays with contrasting compressibility were modeled numerically using a 1-D CRS consolidation model to evaluate learning and testing performance. The results show that the SARSA agent autonomously learns soil-specific strain-rate policies and maintains smooth effective stress trajectories and stable pore-pressure ratio responses. Test duration reductions of 60-75% were achieved depending on soil type. The interpreted compression index (Cc) remains consistent with the baseline CRS values, confirming that reinforcement-learning-based strain-rate control can accelerate testing without compromising data integrity. The study demonstrates the feasibility of reinforcement learning for CRS testing and highlights practical potential for soil-responsive, adaptive strain-rate control. Current limitations include simulation-based evaluation, discretized action selection, and the need for multiple runs to achieve optimal convergence.

Keywords: Soil Consolidation; Reinforcement Learning; SARSA Algorithm; Strain Rate Control.

1. Introduction

Consolidation testing remains fundamental for characterizing the compressibility and drainage behavior of saturated clays. The Constant Rate of Strain (CRS) test was first introduced in the seminal work of Smith & Wahls [1], who demonstrated that continuous strain-controlled loading could provide consolidation parameters more efficiently than incremental testing. Subsequent theoretical developments—most notably the finite-strain formulations of Gibson et al. [2]—established the mechanics underlying strain-controlled consolidation. Building upon this theoretical lineage, Wang et al. [3] recently proposed a strain-controlled finite-strain CRD model incorporating nonlinear compression and permeability relationships, representing a significant advancement in describing realistic soil behavior under continuous deformation. Modern CRS devices equipped with servo-actuators and real-time pore-pressure sensing have improved testing precision and repeatability, making the application of such finite-strain models increasingly practical in laboratory environments [4, 5]. In this work, the term "CRS" follows ASTM D4186, although its formulation aligns with strain-controlled CRD mechanics.

* Corresponding author: dimitri@itb.ac.id

https://doi.org/10.28991/CEJ-2026-012-04-023

© 2026 by the authors. Licensee C.E.J, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).

Despite its advantages, CRS testing still faces longstanding challenges. The concern is the strain-rate selection, which strongly affects drainage accuracy and testing time. With a higher strain rate, the testing time can be reduced significantly, while on the other hand, increasing the pore water pressure lowers the test accuracy. Another problem that arises in CRS is over-accelerated deformation, especially when operators rely on fixed-rate schedules [6]. Existing automated CRS systems cannot adapt to evolving drainage conditions and therefore cannot guarantee pore water pressure stability throughout a test [4, 5]. This leads to the well-known trade-off: very slow rates maintain accuracy but lengthen the test considerably.

A wide range of acceptable pore water pressure ratio limits has been reported in the literature. Smith & Wahls [1] originally proposed an upper limit of approximately 50%, while Wissa et al. [7] recommended maintaining ratios below 5% to ensure fully drained conditions. Subsequent studies suggested intermediate limits, with recommended values ranging from 5–15% [8] to 30–50% [7, 9], and in some cases even higher values without significant deviation from conventional consolidation results [10]. More recent CRS studies have shown that pore water pressure ratio does not remain constant during loading but evolves with increasing stress, often exhibiting non-monotonic trends depending on strain rate and flow regime [3, 4, 11]. In particular, transitions between non-Darcy and Darcy flow have been associated with characteristic changes in pore pressure ratio during CRS consolidation, highlighting the limitations of assuming a single hydraulic regime throughout the test. These findings indicate that a single universal pore-pressure-ratio criterion is unlikely to be valid across all soil types and loading conditions.

Recent applied laboratory programs and modified CRS (MCRS) studies further quantify the implications of these mechanisms for test duration. Institutional multi-test programs reported by Deltares [12] and controlled laboratory investigations by Tran & Pham [13] document that increasing the imposed strain rate beyond conservative ASTM guidance—resulting in transient pore-pressure ratios on the order of 20–30%—can reduce total CRS test duration by approximately 30–50%, provided that validation criteria are satisfied. In these studies, *acceptable agreement* is defined by close overlap of e–log p′ compression curves, compression indices C_c differing by less than typical experimental scatter (on the order of ±5–10%), and continuous effective-stress trajectories without evidence of undrained locking.

Back to the problem, recent CRS studies consistently report that pore-pressure ratio evolves non-monotonically with stress and strain rate, and that no single strain-rate criterion is valid across soil types. This behavior fundamentally challenges classical fixed-rate or rule-based control strategies. Reinforcement Learning (RL) offers a promising solution. RL algorithms learn control policies through interaction with an environment rather than relying on predefined loading schedules. Prior engineering applications demonstrate measurable efficiency and stability gains under dynamic system behavior [14–16]. Among RL methods, the on-policy SARSA algorithm is particularly well suited for laboratory applications, as its conservative exploration reduces the risk of unsafe transients—critical when pore-pressure spikes must be avoided.

This study introduces a simulation-based RL framework that couples the SARSA algorithm with a finite-strain CRS consolidation model [3] to develop an adaptive strain-rate controller. The objective is to maintain pore-pressure ratios within a safe, partially drained range—targeted near 30%—while reducing total testing time. The contributions of this work are threefold:

- Development of a physics-consistent learning environment for adaptive CRS control;
- Formulation of a reward structure that encodes hydraulic safety and time efficiency; and
- Demonstration that SARSA can autonomously avoid unsafe strain-rate selections—such as the commonly used 8× baseline rate—while producing soil-specific loading strategies, addressing longstanding limitations of fixed-rate CRS procedures.

The paper is organized as follows. Section 1 presents the introduction, followed by related works in Section 2. Section 3 describes the mathematical model of the CRS. The control system for CRS consolidation is introduced in Section 4. Section 5 presents the results and discussion. Section 6 outlines the limitations and future work, and Section 6 is the conclusion.

2. Literature Reviews

2.1. Consolidation Testing and Automation

The consolidation test methodology has evolved significantly since Terzaghi's initial theoretical framework, with ongoing refinements challenging traditional 24-hour load increment standards. Multiple studies have demonstrated that standard test durations may be unnecessarily long [17].Research indicates that for many soil types, primary consolidation can be effectively assessed in much shorter time frames - potentially as little as 2-12 hours [18, 19]. Furthermore, Crawford [20] emphasizes that consolidation characteristics vary widely depending on soil stress history, void ratio, and structure. The variability in soil permeability means that while some clays may require less than an hour for consolidation, others might need several days, highlighting the critical need for flexible, soil-specific testing approaches.

Automated consolidation testing systems currently face significant challenges in achieving consistent, accurate results through timer-based control, with machine learning showing promise for parameter prediction but not yet solving real-time test control limitations. The research indicates that while commercial manufacturers offer automated loading sequences, comprehensive evaluations reveal critical shortcomings. Timer-based systems often produce either overly conservative test durations or inadequate consolidation, potentially performing less reliably than experienced human operators who can dynamically adjust testing parameters.

Machine learning approaches have demonstrated capability in predicting consolidation parameters showing neural networks and other algorithm can effectively predict soil consolidation characteristics [21–23]. However, these techniques currently focus on parameter estimation rather than direct real-time test control, leaving a significant technological gap in automated geotechnical testing.

2.2. Reinforcement Learning and SARSA

Reinforcement learning (RL) provides a rigorous mathematical framework for learning optimal behavior through Markov Decision Processes (MDPs), which model sequential decision-making under uncertainty [24]. The MDP framework defines critical components: states representing environment information, actions available to agents, transition probabilities between states, and rewards evaluating state-action desirability [25]. Temporal difference learning enables efficient value function estimation by "bootstrapping" - updating value estimates using observed rewards and predicted future state values [24]. The evidence is strong, with multiple seminal works consistently describing this approach as a fundamental method for agents to learn optimal behavior through iterative interaction with complex, uncertain environments [26].

SARSA is an on-policy reinforcement learning algorithm known for its robustness in environments where exploration carries measurable risk, as it incorporates the effects of exploratory actions directly into value updates. Convergence of SARSA under a Lipschitz-continuous policy improvement operator has been mathematically established in prior theoretical work [27]. A subsequent finite-sample analysis under non-i.i.d. conditions demonstrated that the algorithm converges when paired with linear function approximation [28]. Additional theoretical characterization has shown that SARSA converges to a bounded neighborhood even when the Lipschitz constant is arbitrarily large [29]. Empirical investigations have also revealed that SARSA variants exhibit reduced maximization bias relative to off-policy algorithms, improving reliability in uncertain learning domains [30].

Reinforcement Learning (RL) has been demonstrated to deliver notable performance improvements and energy-efficiency benefits across a variety of engineering applications. Significant energy reductions in HVAC systems have been reported, including a 16.7% decrease in cooling demand achieved in a real conference room environment [14]. Additional studies have indicated the potential for 15–30% energy savings under RL-based building control strategies [16]. Increasing adoption of RL has also been identified in construction engineering, encompassing infrastructure management, building energy optimization, and construction machinery control [15]. However, existing literature has not provided direct evidence regarding RL implementations in robotics or laboratory equipment automation, limiting conclusions about deployment in those domains [15]. Although the research landscape suggests broad potential for real-world deployment, RL for building control remains primarily experimental, with only 11% of studies involving actual buildings [16].

3. Mathematical Soil Model for Constant-Rate of Deformation (CRD) Consolidation

The definition of natural strain ε_{Na} being the following [3]:

$$\varepsilon_{Na} = -\ln\frac{H_t}{H_0} = -\ln\left(\frac{1+e}{1+e_0}\right) \quad (1)$$

where H_t is the current specimen height and e is the void ratio. Then, H_0 is initial specimen height and e_0 is the initial void ratio. Figure 1 illustrates the finite strain consolidation model used for the consolidation test, expressed in the Eulerian framework. In this model, H_0 and H_t denote the initial and current heights of the specimen, respectively, while z_E represents the Eulerian coordinate. The coordinate origin is located at the specimen's base, with the upward direction defined as positive. The upper boundary serves as a drainage surface that moves downward at a deformation rate v, whereas the lower boundary is fixed and undrained. Unlike the small-strain consolidation model, this formulation considers variations in specimen height during loading, providing a more realistic representation of practical conditions [3].

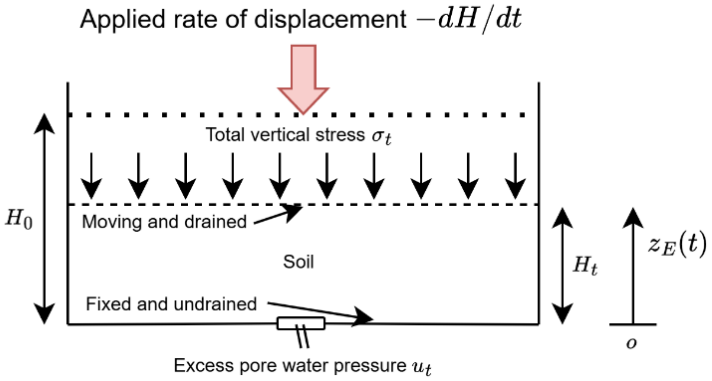


Figure 1. Finite strain model of the consolidation test in the Eulerian description [3]

The governing consolidation equation after simplication by assuming that the unit weight of the solid γ_s is equal to that of the fluid γ_w, is given as follows:

$$\frac{\partial}{\partial z_E}\left[c_v(\varepsilon_{Na})\frac{\partial \varepsilon_{Na}}{\partial z_E}\right] = \frac{D\varepsilon_{Na}}{Dt} \tag{2}$$

where;

$$c_v(\varepsilon_{Na}) = \frac{k(\varepsilon_{Na})}{m_v(\varepsilon_{Na})\gamma_w} = \frac{k(\varepsilon_{Na})}{\gamma_w}\frac{d\sigma'}{d\varepsilon_{Na}} \tag{3}$$

and;

$$c_v(\varepsilon_{Na}) = \frac{k(\varepsilon_{Na})}{m_v(\varepsilon_{Na})\gamma_w} = \frac{k(\varepsilon_{Na})}{\gamma_w}\frac{d\sigma'}{d\varepsilon_{Na}} \tag{4}$$

here, k is the hydraulic conductivity, σ' is the effective stress, v_s is the macroscopic velocity of the solid phase, D/Dt is the total derivative. The coefficient of consolidation can be determined as the following [3]:

$$c_v(\varepsilon_{Na}) = \frac{k_0\sigma_0'(1+e_0)e^{-\varepsilon_{Na}}10^{\frac{(1+e_0)(1-e^{-\varepsilon_{Na}})\left(\frac{1}{C_c}-\frac{1}{C_k}\right)}{}}}{0.434C_c\gamma_w} \tag{5}$$

where C_c is the compression index and σ_0' is the initial effective stress.

In the case of a thin soil layer, we can obtain:

$$\frac{\partial u}{\partial z_E} = -\frac{\partial \sigma'}{\partial z_E} \tag{6}$$

where u is the excess pore water pressure. In the absence of secondary consolidation, the effective stress will be a unique function of natural strain, and we can obtain:

$$\frac{\partial u}{\partial z_E} = -\frac{1}{m_v(\varepsilon_{Na})}\frac{\partial \varepsilon_{Na}}{\partial z_E} \tag{7}$$

where coefficient of volume compressibility m_v is:

$$m_v(\varepsilon_{Na}) = -\frac{1}{1+e}\frac{de}{d\sigma'} = \frac{d(\varepsilon_{Na})}{d\sigma'} \tag{8}$$

As we can see in Figure 1, the upper surface of the consolidation test was moving downward at a deformation rate v, and the settlement at any time t was vt. Then, from [3], the hydraulic boundary condition of the drainage surface is:

$$c_v(\varepsilon_{Na})\frac{\partial \varepsilon_{Na}}{\partial z_E}\Big|_{Z_E=H_t} = v \tag{9}$$

The undrained boundary condition at the lower surface can be written in terms of natural strain:

$$c_v(\varepsilon_{Na})\frac{\partial \varepsilon_{Na}}{\partial z_E}\Big|_{Z_E=0} = 0 \tag{10}$$

The initial condition is given by:

$$\varepsilon_{Na}(z_E, 0) = 0 \tag{11}$$

Let:

$$Z_E = \frac{z_E}{H_0} \tag{12}$$

$$\alpha = \frac{vt}{H_0} \tag{13}$$

where Z_E is the normalized Eulerian coordinate and α is the normalized time.

Following these definitions, the governing Equation 2, boundary condition Equations 9 and 10, and initial condition Equation 11 were rewritten as:

$$\frac{\partial}{\partial z_E}\left[\frac{1}{\beta(\varepsilon_{Na})}\frac{\partial \varepsilon_{Na}}{\partial Z_E}\right] = \frac{D\varepsilon_{Na}}{D\alpha} \tag{14}$$

$$\frac{1}{\beta(\varepsilon_{Na})}\frac{\partial \varepsilon_{Na}}{\partial Z_E}\Big|_{Z_E=\eta} = 1 \tag{15}$$

$$\frac{\partial \varepsilon_{Na}}{\partial Z_E}\Big|_{Z_E=0} = 0 \tag{16}$$

$$\varepsilon_{Na}(Z_E, 0) = 0 \tag{17}$$

where;

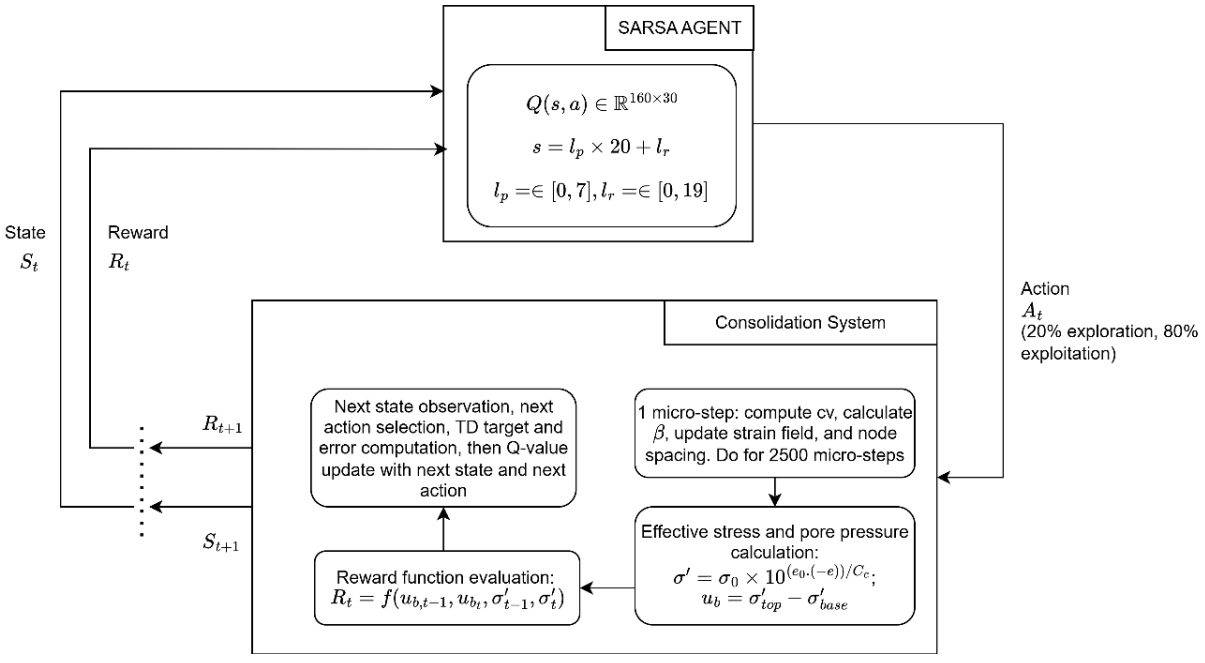
$$\beta(\varepsilon_{Na}) = \frac{vH_0}{c_v(\varepsilon_{Na})} \tag{18}$$

$$\eta(\alpha) = \frac{H_t}{H_0} = \frac{H_0 - vt}{H_0} = 1 - \alpha \tag{19}$$

where, β is the normalized deformation rate parameter and η is the normalized specimen height.

4. Research Methodology

This study adopts a simulation-based methodology to evaluate the feasibility of applying the SARSA reinforcement learning algorithm for adaptive control of the soil consolidation process under a Constant Rate of Deformation (CRD) framework. The methodology consists of four main components: (1) modeling the CRD consolidation process as the learning environment, (2) defining the state and action spaces including operational constraints, (3) designing a multi-objective reward function, and (4) training the SARSA agent and setting its hyperparameters. The relationship among these components and the flow of information within the proposed framework are summarize in Figure 2.



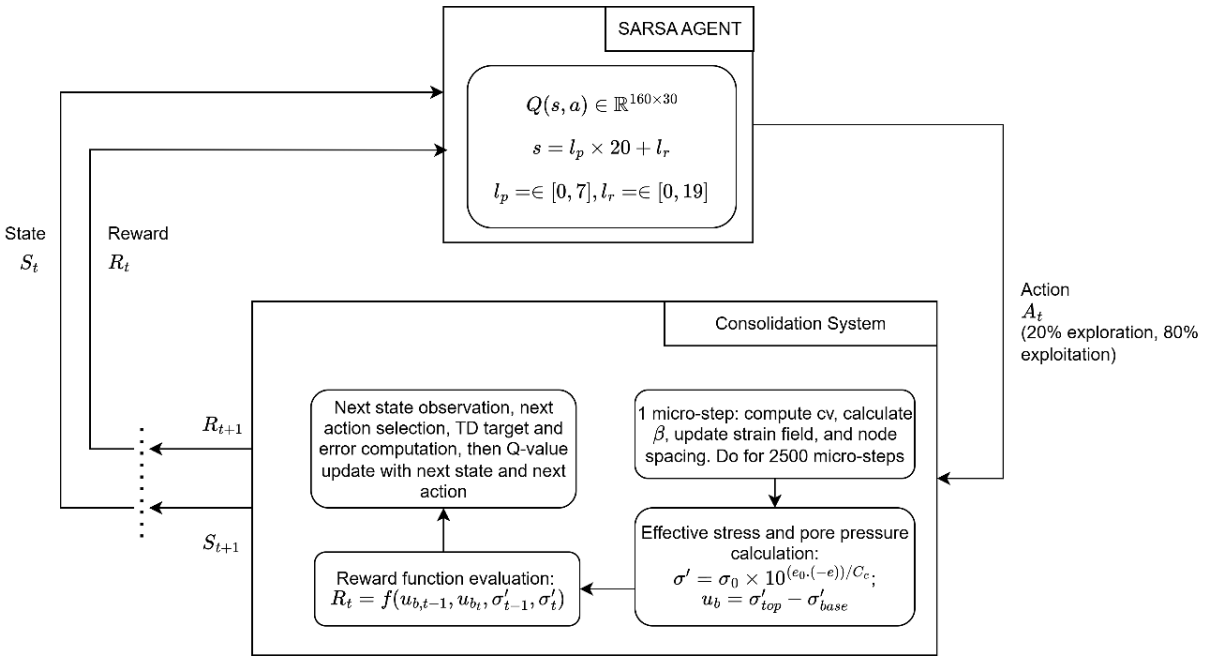


Figure 2. Overview of the SARSA-based reinforcement learning framework integrated with the soil consolidation system

4.1. Consolidation Model Formulation

The simulation environment [26] is built upon the finite-strain consolidation model [3], formulated within the Eulerian description and governed by the nonlinear partial differential equations presented in Section 2. The model captures the coupling between effective stress, pore-water pressure, and deformation under strain-controlled boundary conditions. To obtain a numerically stable solution, the soil specimen is discretized into 101 equally spaced layers along its height. The governing consolidation equations are solved iteratively using a finite-difference scheme in the natural strain domain. At each time step, the soil response is updated considering both the compressibility and permeability evolution with void ratio, ensuring physically realistic dissipation and settlement behavior.

The top boundary condition is a prescribed downward displacement representing the deformation rate imposed by the virtual actuator. The effective stress at the top layer (σ'_{top}) is extracted as the representative measure of external load evolution. Conversely, the pore-water pressure at the bottom layer (u_b) is monitored to characterize drainage response and potential excess pressure buildup. These two outputs serve as key environmental observations that form the feedback inputs to the reinforcement learning agent. These two outputs are what is taken from the real CRD test. This simulation does not consider measurement noise, actuator backlash, and sensor drift. It's assumed that the laboratory testing is perfect.

Two clay samples were selected to model different clay behavior is show in Table 1, taken from [3]. All initial heights are set to be 20 mm, as used in usual consolidation standards. Initially, all soil models were tested using initial speed (or base speed), and their maximum speed (as used for action space explained later) to be used as "low v" and "high v" later.

Table 1. Soil Parameters

Parameters	Soil 1	Soil 2
Initial void ratio, e_0	1.84	1.02
Initial effective stress, σ'_0, (kPa)	10	10
Initial hydraulic conductivity, k_0 (m/s)	3.93e-9	3.48e-9
Compression index, C_c	0.55	0.26
Conductivity–void ratio index, C_k	0.41	0.27
Unit weight of water: γ_w	9810	9810
Initial speed, v_0 (m/s)	1.67e-8	3.33e-8

4.2. State–Action Space Formulation

The soil consolidation environment is modeled as a Markov Decision Process (MDP). The state space (S) represents the dynamic condition of the test and consists of two dimensions: progress level and pore-pressure ratio.

The progress level (p) quantifies the advancement of the consolidation process by considering both mechanical deformation and stress development. It is defined as the maximum between the normalized axial strain ($p_\alpha = \alpha/\alpha_{target}$) and the normalized effective stress ($p_{\sigma'} = \sigma'/\sigma'_{target}$) with respect to their respective target values. The target effective stress is set at 1500 kPa, which is discretized into eight progress intervals, each representing approximately 12.5% of the total consolidation load. Similarly, the axial strain is normalized to a maximum value of 0.40 (40%), also divided into eight discrete levels. The overall progress level at any time step is thus determined as: $p = \max(p_{\sigma'}.p_\alpha)$ ensuring that the metric reflects the dominant consolidation mechanism at each stage, whether governed by stress increase or strain accumulation.

The pore-pressure ratio (r) characterizes the hydraulic condition of the specimen and is defined as the ratio between the excess pore-water pressure at the specimen base and the current effective stress at the top boundary. The ratio ranges from 0% to 100%, corresponding respectively to fully drained and fully undrained conditions. To capture transitional drainage behavior with sufficient granularity, this range is discretized into 20 uniform categories. By combining the eight progress levels and twenty pore-pressure ratio levels, the resulting state space comprises 160 discrete states (8 × 20). These states are mapped linearly into integer indices to facilitate efficient tabular representation in the reinforcement learning algorithm without explicit tuple encoding. The action space (A) correspond to discrete strain-rate control through the speed of the virtual linear actuator. 30 actions are defined linearly spaced between 1 to 8 which is a multiplier to the base v. Higher multiplier correspond to higher speed.

4.3. Reward Function Design

The reward mechanism in this study employs a single-objective, stability-oriented formulation that inherently drives both safety and time efficiency, rather than using an explicit multi-objective weighting scheme. The design principle is to maintain the pore-pressure ratio within a physically safe range while enabling the shortest possible consolidation time. This formulation aligns with reinforcement learning practices where the task objective can be implicitly encoded through scalar rewards that balance progress and constraint satisfaction [16, 28].

At each simulation step, the reward R_t is computed as a function of the previous and current pore-pressure ratios and the corresponding effective stress levels. Let $r_t = \frac{u_t}{\sigma'_t}$ be the pore pressure ratio at time t, $r_{\min}$ and $r_{\max}$ be the lower and upper target bounds, Δr be the difference between r_t and r_{t-1}, then the reward can be defined as follows,

$$R_t = f_{sc}(\sigma'_t)[R_{dir}(r_t) + R_{stab}(r_t, \Delta r) + R_{viol}(r_t)] \tag{20}$$

and Figure 3 shows schematic model for the reward function.

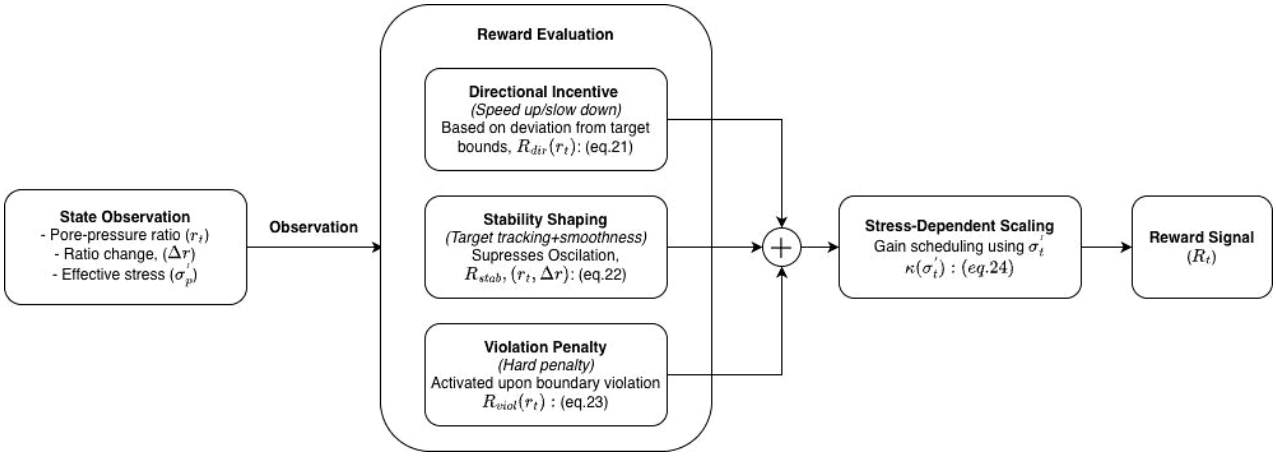


Figure 3. Reward Function Schematic Model

The three components and scaling factor are defined as follows:

- Directional Incentive Term R_{dir}, which encourages speeding up when drainage is *too* effective (ratio too low) and slowing down when pore pressure becomes excessive:

$$R_{\text{dir}}(r_t) = \begin{cases} +\delta_1\,(r_{\min} - r_t)\,. & r_t < r_{\min} \\ -\delta_2\,(r_t - r_{\max})\,. & r_t > r_{\max} \\ 0 \quad . & r_{\min} \le r_t \le r_{\max} \end{cases} \tag{21}$$

with $\delta_1.\,\delta_2 > 0$ and are constant.

- Stability Shaping Term R_{stab}, stabilitzy the ratio to the target and penalizes abrupt changes between steps:

$$R_{\text{stab}}(r_t.\Delta r) = c_1 \exp\left[-\left(\frac{r_t - r_{\text{target}}}{c_2}\right)^2\right] - c_3|\Delta \text{r}|, \quad (22)$$

where c_1 and c_2 controls the width of the "safe band", and c_3 penalizes oscillations.

- Constraint Enforcement Term R_{viol} gives severe penalty for violating safety bounds.

Severe penalty for violating safety bounds:

$$R_{\text{viol}}(r_t) = \begin{cases} -M\,(r_{\min} - r_t), & r_t \le 0 \\ -M\,(r_t - r_{\max}), & r_t \ge 1 \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

with $M \gg 1$ a large penalty coefficient.

- Stress-Dependent Scaling Factor $f_{sc}(\sigma_t')$ multiplies reward which increases with effective stress so that the learning is more active when meaningful CRS loading has begun:

$$f_{sc}(\sigma_t') = 1 + c_{sc}\frac{\sigma_t'}{\sigma_{ref}'} \quad (24)$$

where c_{sc} is a small tuning constant and σ_{ref}' is a normalization stress.

This reward formulation follows reinforcement learning frameworks emphasizing task-oriented scalar reward definitions for control safety and rapid convergence [27, 31]. Similar scalar, constraint-sensitive reward strategies have been successfully employed in engineering control contexts such as adaptive structural control [32] and construction scheduling optimization [33] where stability and efficiency must be balanced implicitly through reward dynamics rather than through weighted multi-objective functions.

4.4. SARSA Training Procedure and Hyperparameters

The State–Action–Reward–State–Action (SARSA) algorithm was employed to train the reinforcement-learning controller within the simulated consolidation environment. SARSA is a temporal-difference (TD) control method that updates the value of the current state–action pair using the reward received and the estimated value of the next state–action pair following the same policy [26]. Unlike off-policy algorithms such as Q-learning, SARSA evaluates and improves the same behavior policy, which makes it inherently more conservative and suitable for safety-critical or real-world physical systems where exploratory actions can cause unstable or irreversible outcomes [29].

The update rule is expressed as:

$$(s.a) \leftarrow Q(s.a) + \eta[r + \lambda Q(s'.a') - Q(s.a)] \quad (25)$$

where α is the learning rate and γ is the discount factor for future rewards.

Each training episode represents a full virtual consolidation test that begins from an initial strain of zero and terminates when the target strain (α = 0.40) or the effective stress limit (1500 kPa) is reached, or when the pore-pressure ratio exceeds the prescribed bounds. The agent interacts with the environment in discrete time steps: at each step it selects an action (motor delay), observes the new state and reward, and updates the Q-table using the SARSA rule.

Exploration and exploitation are balanced through an ε-greedy policy with exponential decay:

$$(\varepsilon_t = \max\,(\varepsilon_{\min}.\varepsilon_0 \times 0{,}995^t) \quad (26)$$

where for example, the initial exploration rate $\varepsilon_0 = 0{,}1$ gradually decreases toward a minimum value of $\varepsilon_{\min} = 0{,}01$. This approach allows the agent to explore the environment extensively during early training while progressively favoring exploitation of learned policies as learning stabilizes.

The Q-table was initialized with zeros for all state–action pairs, i.e., $Q(s.a) = 0\ \forall s.a$. During each iteration, the agent selects an action using the epsilon-greedy strategy, choosing the action with the maximum Q-value for exploitation or a random action for exploration. After executing the action and observing the resulting transition $(s.a.r.s'.a')$, the Q-table is updated according to the SARSA rule. This step-by-step updating mechanism allows the agent to continuously refine its behavior policy based on real-time interaction data, leading to a progressively improved control performance over time.

The hyperparameters used in the final configuration are presented in Table 2.

Table 2. Final Hyperparameter Configuration

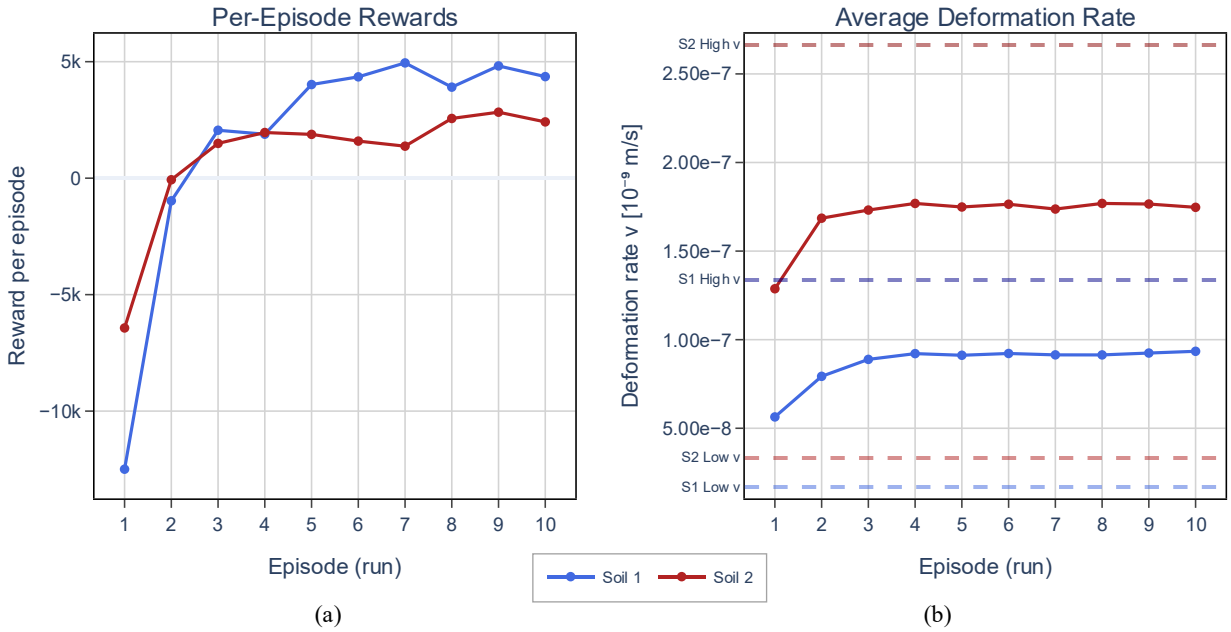
RL Parameter	Symbol	Value	Description
Learning rate	η	0.1	Moderate update speed ensuring smooth convergence
Discount factor	λ	0.95	Balances immediate and long-term reward effects
Min exploration rate	ε_{min}	0.05	Low residual exploration after convergence
Max actions per episode	–	4000	Equivalent to the maximum control adjustments per simulation

Training was executed for ten episodes, after which the cumulative reward and policy performance stabilized. The converged Q-table was then analyzed to verify monotonic improvement in both consolidation efficiency and safety indices.

5. Results and Discussion

5.1. Evolution of Reward and Deformation Rate during Training

Figure 4-a illustrates the evolution of cumulative reward over successive SARSA training episodes, while Figure 4-b shows the corresponding strain-rate selections for Soil 1 and Soil 2.



(a) (b)

Figure 4. Total CRS test time with each episode: (a) absolute time and (b) relative time

The coupling between Figures 4-a and 4-b demonstrates three points: The RL agent autonomously learns "operator-like" behavior, adjusting strain rate in response to pore-pressure feedback; High-plasticity clays require more exploration, consistent with their non-linear consolidation characteristics; SARSA's on-policy nature provides a stable learning trend, avoiding risky strain-rate jumps typical of off-policy methods. The monotonic increase in cumulative reward indicates stable policy convergence, consistent with theoretical expectations for on-policy SARSA learning in constrained environments [26, 28].

5.2. Reduction in CRS Test Duration Through Learning

Figure 5-a presents the absolute CRS test duration for each episode, while Figure 5-b shows the relative duration normalized by the baseline CRS test. After convergence, total test time is reduced by approximately 60–75%, depending on soil type.

These reductions exceed those typically reported in modified CRS (MCRS) and accelerated CRS studies. For example, Tran & Pham [13] and Singh et al. [4] reported time reductions of approximately 30–50% when pore-pressure ratios were allowed to transiently reach 20–30%. Deltares laboratory programs similarly reported reductions on the order of 40% when deviating from conservative ASTM strain-rate limits.

Unlike these previous studies, which rely on predefined strain-rate schedules or empirical thresholds, the present approach autonomously adapts the strain rate based on real-time hydraulic feedback. This distinction explains why larger time reductions can be achieved without compromising drainage stability.

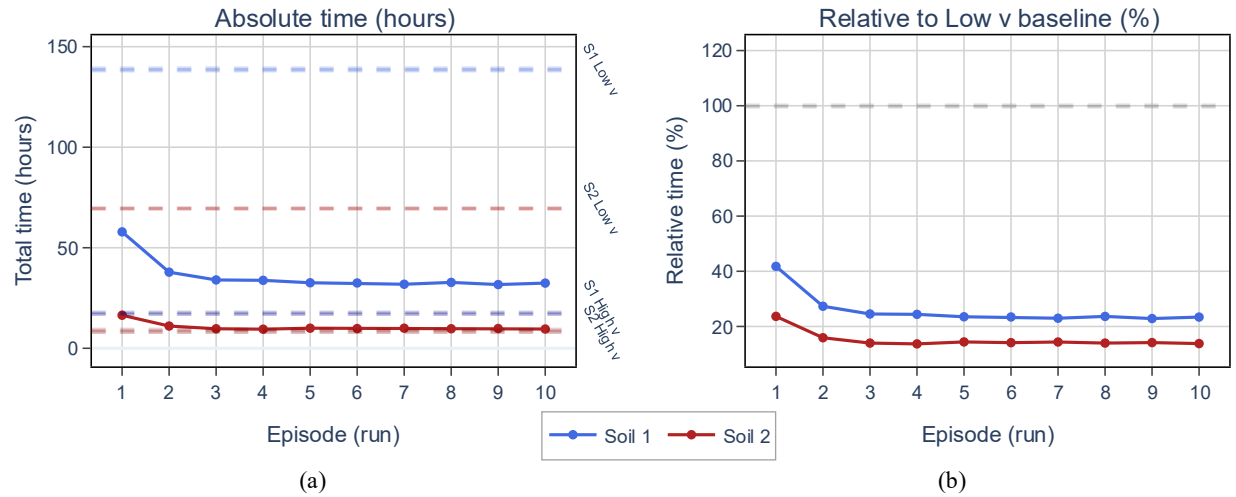


Figure 5. Total CRS test time with each episode: (a) absolute time and (b) relative time

5.3. Consolidation response Under SARSA-Controlled CRS

Figures 6 to 8 collectively illustrate the mechanical and hydraulic response of the soil samples under SARSA-controlled CRS loading. These figures show how the reinforcement-learning controller is able to maintain hydraulic stability while allowing higher strain rates than conventional CRS practice—yet avoiding unsafe pore-pressure behavior. The reward function was intentionally designed to guide the agent toward an average pore-pressure ratio target of approximately 30%, a value chosen to ensure fully drained-to-partially drained conditions without approaching the undrained or unstable region. This target plays a central role in shaping the observed responses.

Figure 6 shows the evolution of effective stress (σ') as a function of normalized strain. Both Soil 1 and Soil 2 exhibit smooth, monotonic increases in σ', indicating that the SARSA-selected strain rates remain within the domain of physically admissible CRS loading. Soil 1, with its high initial void ratio (e_0 = 1.84) and high compressibility (Cc = 0.55), progresses through a more compliant early stress range, resulting in a gentler slope. Soil 2, having a lower initial void ratio (e_0 = 1.02) and lower Cc (0.26), exhibits a steeper initial stress response typical of a denser, stiffer clay matrix. Crucially, the lack of stress oscillations demonstrates that the agent does not over-accelerate deformation, thereby avoiding abrupt pore-pressure spikes. The resulting $\sigma' vs\ \alpha$ trajectories remain suitable for extracting Cc and mv without distortion.

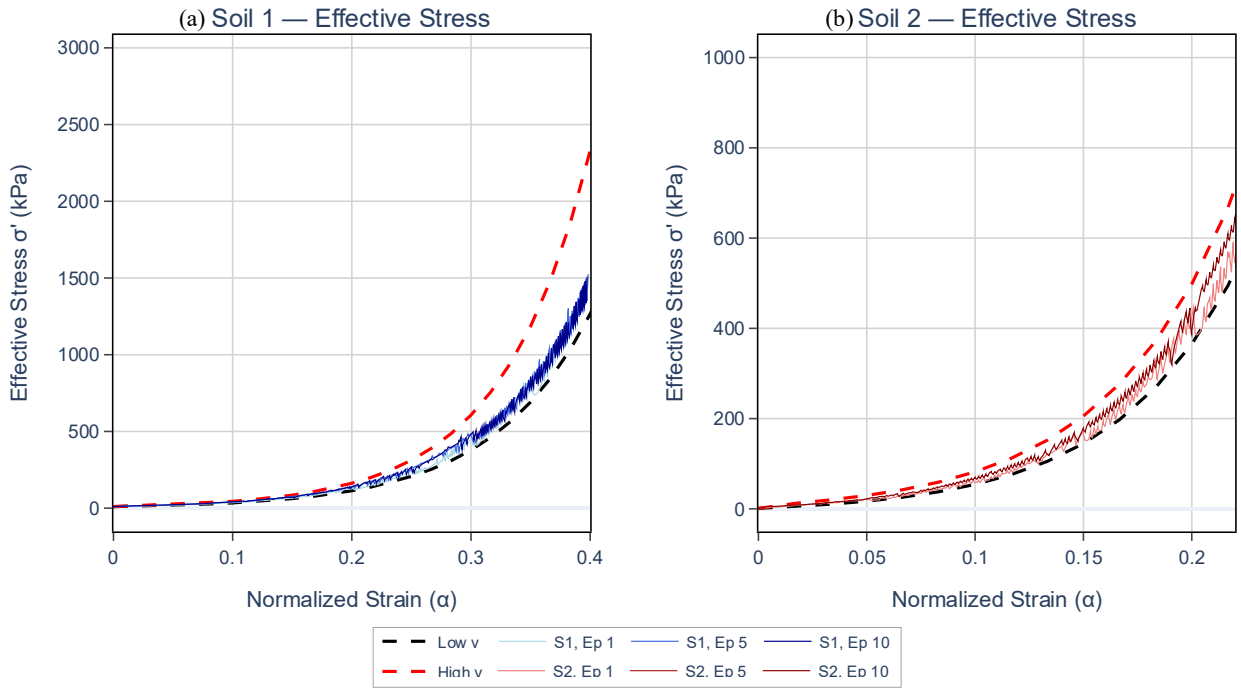


Figure 6. Effective stress buildup in CRS model, (a) Soil 1 and (b) Soil 2

Figure 7 presents the excess pore pressure (u) generated during the CRS loading. The soils respond differently: Soil 1 produces a higher u due to its greater compressibility and void-ratio–dependent hydraulic conductivity, whereas Soil

2 shows significantly lower pore-pressure buildup. The RL controller modulates the strain rate according to the reward system aiming for an average pore-pressure ratio near 30%. As a result, u increases only to levels that keep the pore-pressure ratio within the desired safe zone. This confirms that the reward system's targeting mechanism correctly shapes the strain-rate behavior. A critical observation is that when the strain rate is pushed to the conventional maximum value (8× the base v)—which represents the typical upper-bound CRS strain rate used in practice—both soils exhibit pore-pressure ratios exceeding safe limits. This clearly demonstrates that fixed maximum CRS speeds can violate drainage stability, especially in high-compressibility clays. The SARSA agent avoids selecting the unsafe 8× speed after learning, confirming its awareness of the soil's drainage constraints.

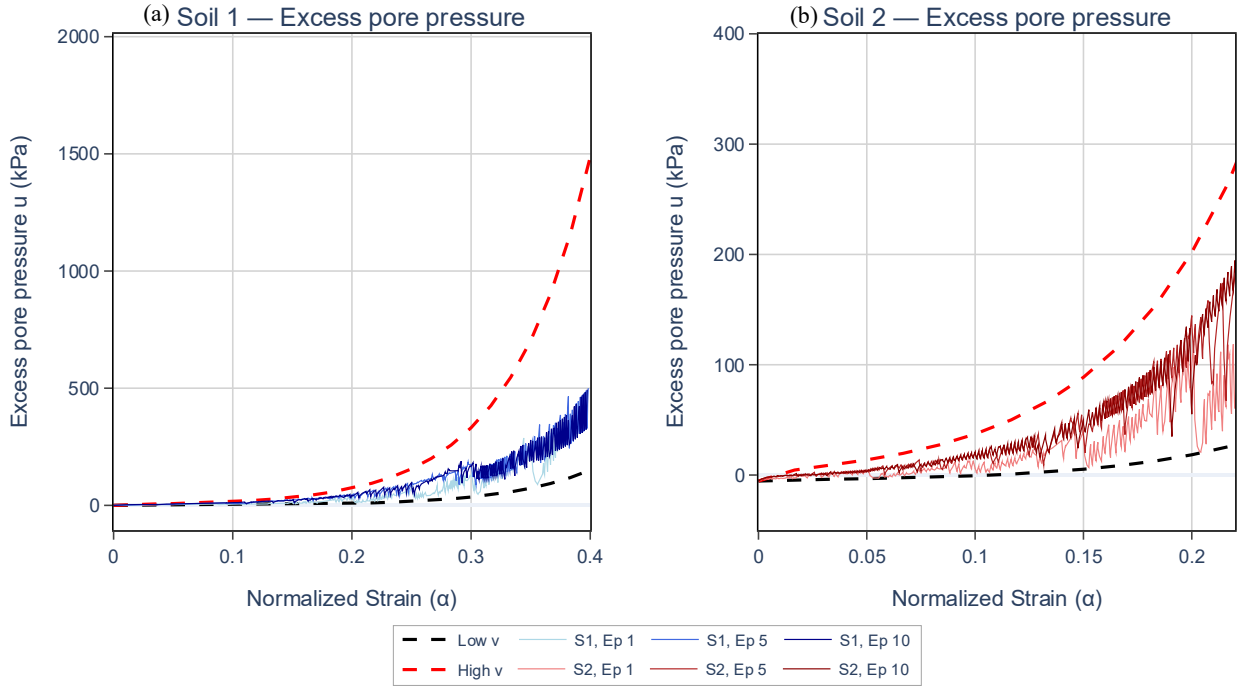


Figure 7. CRS Excess pore pressure u vs normalized strain, (a) Soil 1 (b) Soil 2

The excess pore-pressure ratio, $r = u/\sigma'$, provides the clearest measure of hydraulic safety, as shown in Figure 8. Early episodes contain several exceedances of the acceptable range, particularly during the controller's exploration of the 8x strain rate. However, once the SARSA policy converges, the r-values consistently remain centered around the reward-targeted 30% average. The later episodes show smooth trajectories, free from spikes or divergence.

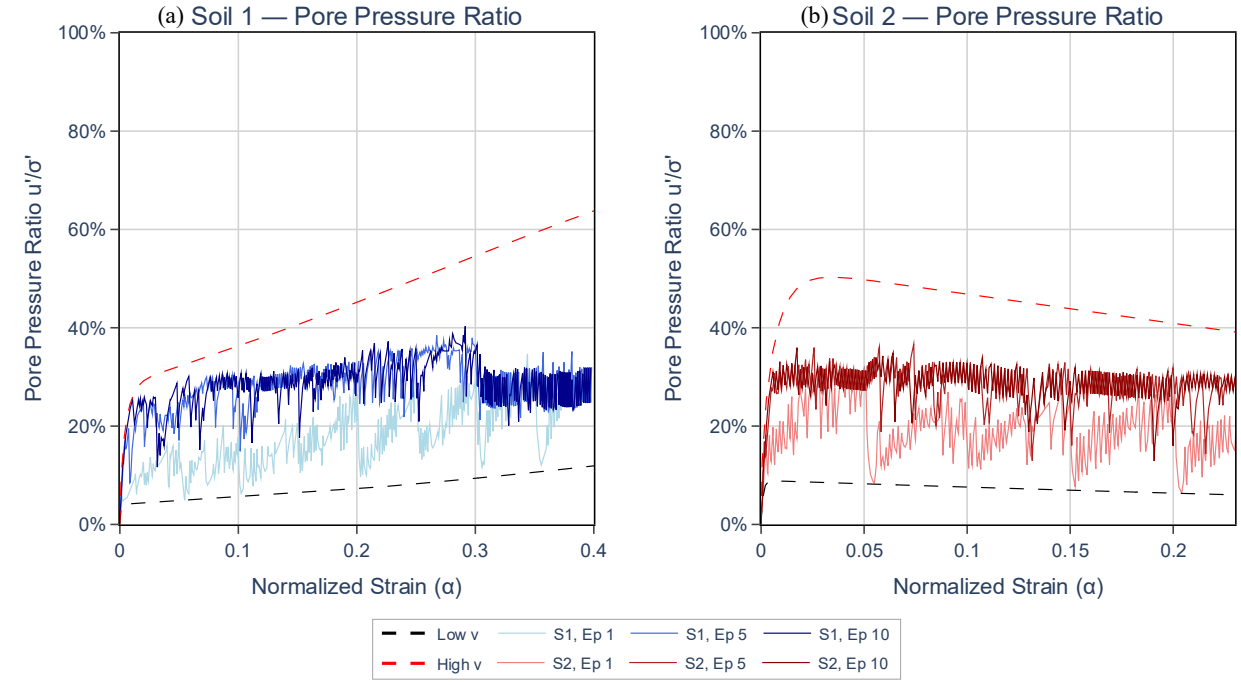
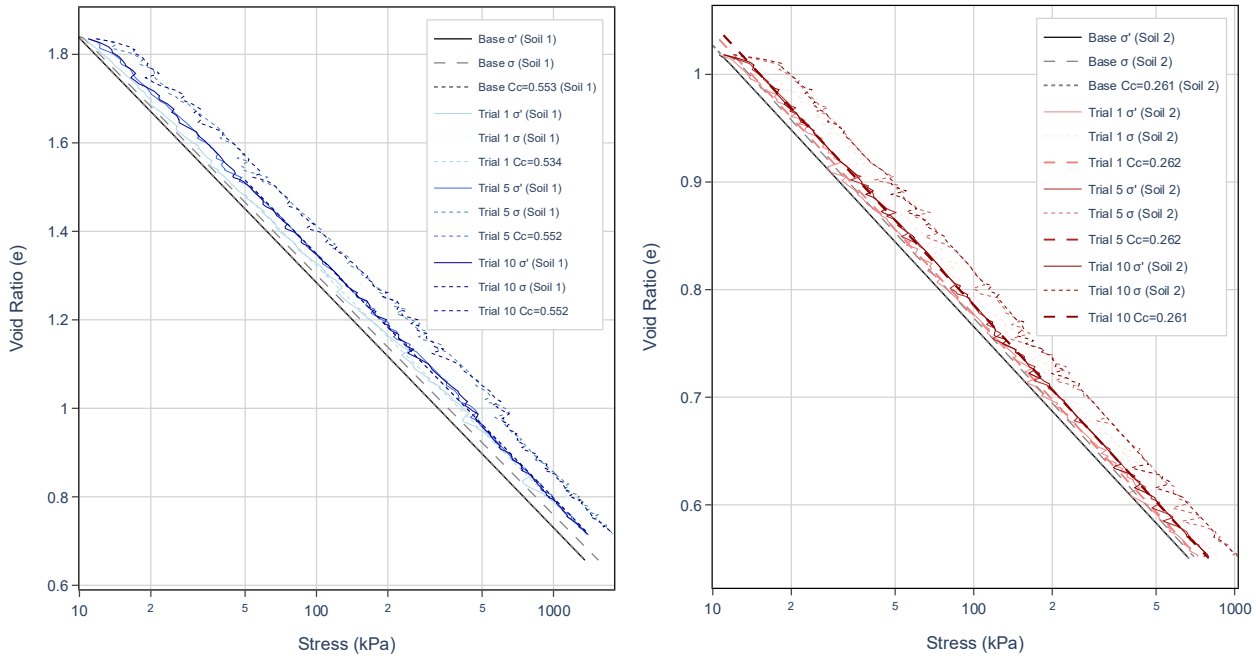


Figure 8. CRS pore pressure to effective stress ratio for selected episodes, (a) Soil 1 and (b) Soil 2

Soil 2 exhibits tighter r-bands due to its faster drainage characteristics, while Soil 1 shows moderately higher variability yet remains within safe limits. These outcomes confirm that the reinforcement-learning controller maintains partially drained conditions essential for accurate CRS interpretation—an improvement over fixed-rate CRS methods, which may unintentionally induce undrained behavior.

5.4. Estimation of Consolidation Parameter Cc

The compression index C_c was interpreted from the linear segment of the $e - \log \sigma'$ curves for the base CRS test and all reinforcement-learning controlled trials which can be seen in Figure 9. The results demonstrate that the adaptive strain-rate strategy does not modify the intrinsic compressibility characteristics of the soils tested. The reference values obtained from the forward CRS test are $C_c = 0.553$ for Soil 1 and $C_c = 0.261$ for Soil 2. In comparison, the mean experimentally reported values from conventional oedometer-based characterization are approximately 0.55 and 0.26, respectively, confirming that the baseline CRS interpretation aligns well with established soil behavior.



(a) (b)

Figure 9. CRS e vs log p' for selected episodes for (a) Soil 1 and (b) Soil 2

The compression index values derived from SARSA-controlled CRS tests closely match baseline CRS results and reported literature values [3]. For Soil 1, Trial 1 yielded $C_c = 0{,}534$, while Trials 5 and 10 produced $C_c = 0{,}552$ and $0{,}552$, respectively. This corresponds to a maximum deviation of 3.4% relative to the base run, and 0.18% between Trial 10 and the baseline. For Soil 2, Trial 1 gave $C_c = 0{,}262$, and Trials 5 and 10 yielded $C_c = 0{,}262$ and $0{,}261$, representing a maximum deviation of 0.38%. The maximum deviation observed is less than 4%, which falls well within typical experimental scatter reported for CRS and oedometer testing [11].

Similar conclusions were reached in modified CRS studies [4], but the present approach achieves these results with greater efficiency and without reliance on predefined operator judgment.

5.5. Discussion

The combined results demonstrate that reinforcement learning provides a substantial advancement in controlling CRS consolidation testing. Traditional CRS loading often relies on fixed strain rates, which assume that the chosen rate is suitable for all soil types. Numerous studies have documented that inappropriate CRS rates may induce undrained or partially undrained responses, violating the theoretical assumptions behind the test [34–36]. This limitation is clearly reflected in the early episodes of the present study, where the conventional 8x deformation rate repeatedly produced pore-pressure ratios well above the safe threshold, especially in Soil 1.

The SARSA-controller overcomes this challenge by continuously adjusting the strain rate in response to real-time pore-pressure conditions. Through its reward mechanism, the agent inherently learns the drainage capacity of the soil. The consistent convergence toward an average pore-pressure ratio close to 30%—a value selected to preserve partial drainage while maintaining test efficiency—demonstrates that reinforcement learning can enforce hydraulic safety more reliably than fixed-rate control. This adaptive behavior also maintains the physical interpretability of the results. The

resulting effective stress trajectories are smooth, and suitable for back-calculating Cc. The estimates from RL-based approaches do not differ from the conventional CRS, which has been explored in Section 5.4. as for m_v and C_v are still not explored in this paper.

The time–reward analysis further highlights the practical value of the RL-based approach. By avoiding unsafe acceleration and instead applying dynamically optimized strain rates, the SARSA controller achieves reductions in overall loading time without sacrificing data integrity. Soil 2 benefits more significantly due to its natural drainage efficiency, while Soil 1 still experiences meaningful time reductions once the policy converges.

Overall, these findings suggest that reinforcement learning is capable of transforming CRS testing from a static, rate-controlled procedure into a dynamic, soil-responsive system. It preserves the theoretical assumptions required for accurate interpretation, reduces test duration, and minimizes hydraulic instability—addressing long-standing limitations in CRS methodology highlighted across geotechnical literature. Further research may explore the integration of reinforcement learning with real-time laboratory instrumentation to enable fully automated CRS devices capable of soil-specific optimization.

6. Conclusion

This study introduced a reinforcement-learning framework based on the SARSA algorithm for adaptive strain-rate control in CRS consolidation testing. The findings demonstrate that the SARSA controller can learn soil-specific loading strategies that maintain partially drained conditions by targeting a pore-pressure ratio near 30%, while significantly reducing total testing time by 60-75%. The approach prevents unsafe over-acceleration associated with traditional fixed upper-bound strain-rate selections, particularly the commonly used 8× multiplier, which produced excessive pore pressures in early exploratory episodes.

The results confirm that reinforcement-learning-based control preserves the accuracy of critical consolidation parameters, including the compression index C_c, which showed deviations less than 4%. This indicates that testing efficiency gains are achieved without compromising interpretive validity, representing a meaningful improvement over conventional fixed-rate CRS procedures.

However, several limitations remain. The current framework models only normally consolidated soils, uses discretized speed levels defined relative to an arbitrarily selected base rate, and requires multiple episodes to converge to optimal policies—conditions that limit immediate laboratory applicability where repeated trials are impractical. The model also lacks generalization capability for a broader range of soils and real instrumentation uncertainties.

Future research will focus on developing a physical CRS apparatus using online reinforcement learning capable of first-run learning efficiency, improving strain-rate continuity through continuous-action RL methods, and expanding evaluation to a wider spectrum of soils with diverse compressibility and permeability characteristics.

Overall, the results establish reinforcement learning as a promising foundation for next-generation intelligent CRS systems capable of autonomous, soil-responsive strain-rate optimization and significantly enhanced testing efficiency.

6.1. Limitations and Future Works

While the results demonstrate the potential of reinforcement-learning-based strain-rate control for CRS consolidation testing, several limitations remain. First, the present numerical framework models only normally consolidated soils. The behavior of overconsolidated or structured soils—which commonly exhibit non-linear recompression and destructuration effects—was not considered, and may present significantly different pore-pressure response characteristics. Second, although the RL agent learns an improved strain-rate policy, the action space is still defined through discretized multipliers of a base rate, which is itself arbitrarily selected. A more rigorous basis for determining the initial permissible speed range is needed to support broader applicability.

Third, achieving optimal efficiency currently requires multiple training episodes; convergence occurred only after repeated full CRS simulations rather than a single run. In practical laboratory operation, testing personnel will not be willing to perform many repeated CRS tests solely for controller learning purposes. Therefore, real-world adoption requires learning efficiency equivalent to a first-run or near real-time optimization method. Fourth, the controller learns behavior specific to the two soils evaluated and does not yet generalize to unknown soil conditions. This lack of generalized prior knowledge limits predictive applicability and requires adaptation before testing begins.

Future work will address these limitations through experimental implementation using a physical CRS device integrated with online reinforcement learning and real-time sensing. Additional research will expand the soil database to include a broader range of compressibility, plasticity, and drainage characteristics, enabling model generalization. The next phase will also explore continuous-action RL algorithms capable of selecting strain rate in an adaptive manner rather than relying on discretized speed levels. Together, these developments aim to evolve the reinforcement-learning CRS controller into a fully automated, practical tool for soil-specific optimization in laboratory environments.

7. Declarations

7.1. Author Contributions

Conceptualization, H.N. and D.M.; methodology, H.N. and D.M.; software, M.A.P.N. and T.I.; validation, D.A. and E.R.; formal analysis, M.F.A.; investigation, M.F.A.; resources, M.F.A.; data curation, M.F.A. and D.A.; writing—original draft preparation, M.F.A.; writing—review and editing, H.N., D.A., and E.R.; visualization, M.F.A. and T.I.; supervision, D.M. and H.N.; project administration, D.M.; funding acquisition, D.M. All authors have read and agreed to the published version of the manuscript.

7.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

7.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

7.4. Conflicts of Interest

The authors declare no conflict of interest.

8. References

[1] Smith, R. E., & Wahls, H. E. (1969). Consolidation Under Constant Rates of Strain. Journal of the Soil Mechanics and Foundations Division, 95(2), 519–539. doi:10.1061/jsfeaq.0001263.

[2] Gibson, R. E., England, G. L., & Hussey, M. J. L. (1967). The theory of one-dimensional consolidation of saturated clays. Geotechnique, 17(3), 261–273. doi:10.1680/geot.1967.17.3.261.

[3] Wang, W., Ke, L., & Gu, Y. (2024). A Strain-Controlled Finite Strain Model for CRD Consolidation of Saturated Clays Considering Non-Linear Compression and Permeability Relationships. Water (Switzerland), 16(19), 2858. doi:10.3390/w16192858.

[4] Singh, D. J., Paramkusam, B. R., & Prasad, A. (2022). Determination of Consolidation Parameters of Geomaterials Using Modified CRS Consolidation Testing System. KSCE Journal of Civil Engineering, 26(3), 1066–1079. doi:10.1007/s12205-021-0386-1.

[5] Prativi, A., Mochtar, N. E., & Mochtar, I. B. (2025). Assessment of Primary and Secondary Compression Parameters of Tropical Fibrous Peat Using Improved-CRS Consolidation Test. Civil Engineering Journal (Iran), 11(5), 1756–1771. doi:10.28991/CEJ-2025-011-05-03.

[6] Head, K. H. (1986). Manual of soil laboratory testing: Volume 1: Soil classification and compaction tests. Pentech Press, London, United Kingdom.

[7] Wissa, A. E. Z., Christian, J. T., Davis, E. H., & Heiberg, S. (1971). Consolidation at constant rate of strain. Journal of the Soil Mechanics and Foundations Division, 97(10), 1393–1413. doi:10.1061/JSFEAQ.0001679.

[8] ASTM D4186. (2012). Standard Test Method for One-Dimensional Consolidation Properties of Saturated Cohesive Soils Using Controlled-Strain Loading. ASTM International, Pennsylvania, United States. doi:10.1520/D4186_D4186M-12E01.

[9] Gorman, C., Hopkins, T., Deen, R., & Drnevich, V. (1978). Constant-Rate-of-Strain and Controlled-Gradient Consolidation Testing. Geotechnical Testing Journal, 1(1), 3–15. doi:10.1520/gtj10363j.

[10] Sheahan, T. C., & Watters, P. J. (1997). Experimental Verification of CRS Consolidation Theory. Journal of Geotechnical and Geoenvironmental Engineering, 123(5), 430–437. doi:10.1061/(asce)1090-0241(1997)123:5(430).

[11] Leroueil, S., & Hight, D. (2013). The Behaviour and Properties of Soft Soils. Géotechnique, 63(4), 393–409. doi:10.1680/geot.2013.63.4.393.

[12] Deltares. (2024). Evaluation linear isotach model: Laboratory study and literature review (Contribution to Regio Deal Bodemdaling Groene Hart, Project 42). Deltares, Delft, Netherlands.

[13] Viet, T. T., & Dung, P. H. (2024). Effect of Strain Rate on Consolidation Characteristics of Artificial Silty Clay: Insights from Modified Constant Rate of Strain Tests. International Journal of GEOMATE, 27(124), 87–94. doi:10.21660/2024.124.4705.

[14] Chen, B., Cai, Z., & Bergés, M. (2019). Gnu-RL: A precocial reinforcement learning solution for building HVAC control using a differentiable MPC policy. BuildSys 2019 - Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, 316–325. doi:10.1145/3360322.3360849.

[15] Asghari, V., Wang, Y., Biglari, A. J., Hsu, S.-C., & Tang, P. (2022). Reinforcement Learning in Construction Engineering and Management: A Review. Journal of Construction Engineering and Management, 148(11), 03122009. doi:10.1061/(asce)co.1943-7862.0002386.

[16] Wang, Z., & Hong, T. (2020). Reinforcement learning for building controls: The opportunities and challenges. Applied Energy, 269, 115036. doi:10.1016/j.apenergy.2020.115036.

[17] Mahmud Al-Neami, M. (2013). Effect of Loading Duration on the Parameters Obtained from Consolidation Test. Engineering and Technology Journal, 31(19), 59–69. doi:10.30684/etj.31.19a.5.

[18] Sridharan, A., Sivapullaiah, P., & Stalin, V. (1994). Effect of Short Duration of Load increment on the Compressibility of Soils. Geotechnical Testing Journal, 17(4), 488–496. doi:10.1520/gtj10309j.

[19] Won, J.-Y., Porter, B., & Cotton, B. (2010). End of Primary Consolidation for the Gulf Coast Soils. GeoFlorida, 1108–1115. doi:10.1061/41095(365)110.

[20] Crawford, C. B. (1986). State of the Art: Evaluation and Interpretation of Soil Consolidation Tests. ASTM Special Technical Publication, 1986, 71–103. doi:10.1520/stp34607s.

[21] Liu, H., Lin, P., & Wang, J. (2023). Machine learning approaches to estimation of the compressibility of soft soils. Frontiers in Earth Science, 11, 1147825. doi:10.3389/feart.2023.1147825.

[22] Yousefpour, N., & Fallah, S. (2018). Applications of Machine Learning in Geotechnics. Proceedings of the Civil Engineering Research in Ireland Conference, Dublin, Ireland.

[23] Kim, M., Senturk, M. A., Tan, R. K., Ordu, E., & Ko, J. (2024). Deep Learning Approach on Prediction of Soil Consolidation Characteristics. Buildings, 14(2), 450. doi:10.3390/buildings14020450.

[24] Dayan, P., & Watkins, C. J. (2002). Reinforcement learning. Stevens' Handbook of Experimental Psychology, 103.

[25] Bufacchi, R. J., & Iannetti, G. D. (2019). The Value of Actions, in Time and Space. Trends in Cognitive Sciences, 23(4), 270–271. doi:10.1016/j.tics.2019.01.011.

[26] Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. MIT Press, Cambridge, United States.

[27] Perkins, T. J., & Precup, D. (2002). A convergent form of approximate policy iteration. Proceedings of the 16th International Conference on Neural Information Processing Systems, 1627-1634. doi:10.5555/2968618.2968820.

[28] Zou, S., Xu, T., & Liang, Y. (2019). Finite-sample analysis for SARSA with linear function approximation. Advances in Neural Information Processing Systems, 778, 8668 - 8678. doi:10.5555/3454287.3455065.

[29] Zhang, S., des Combes, R. T., & Laroche, R. (2023). On the Convergence of SARSA with Linear Function Approximation. Proceedings of Machine Learning Research, 202, 41613–41646.

[30] Tata, G., & Austin, E. (2021). Investigation of Maximization Bias in Sarsa Variants. 2021 IEEE Symposium Series on Computational Intelligence, SSCI 2021 - Proceedings, 1–8. doi:10.1109/SSCI50451.2021.9660081.

[31] Achiam, J., Held, D., Tamar, A., & Abbeel, P. (2017). Constrained policy optimization. Proceedings of the 34th International Conference on Machine Learning (ICML 2017), 70, 22–31.

[32] Sadeghi Eshkevari, S., Sadeghi Eshkevari, S., Sen, D., & Pakzad, S. N. (2023). Active structural control framework using policy-gradient reinforcement learning. Engineering Structures, 274, 115122. doi:10.1016/j.engstruct.2022.115122.

[33] Yao, Y., Tam, V. W. Y., Wang, J., Le, K. N., & Butera, A. (2024). Automated construction scheduling using deep reinforcement learning with valid action sampling. Automation in Construction, 166, 105622. doi:10.1016/j.autcon.2024.105622.

[34] Kulhawy, F. H. (1992). Effect of lateral stress on CPT penetration pore pressures. Journal of Geotechnical Engineering, 118(10), 1657–1660. doi:10.1061/(ASCE)0733-9410(1992)118:10(1657.2).

[35] Marinho, F. A. M. (2005). Nature of Soil–Water Characteristic Curve for Plastic Soils. Journal of Geotechnical and Geoenvironmental Engineering, 131(5), 654–661. doi:10.1061/(asce)1090-0241(2005)131:5(654).

[36] Chai, J., & Carter, J. P. (2011). Deformation Analysis in Soft Ground Improvement. Geotechnical, Geological, and Earthquake Engineering. Springer, Dordrecht, Germany. doi:10.1007/978-94-007-1721-3.