vehicles

MDPI

Article

Research on Multi-Objective Optimization of Clutch Engagement Strategy Based on Deep Reinforcement Learning

Ying Liu [1], Chengyou Xie [1], Yongxian Zhang [1], Cheng Zeng [2], Yinmin Huang [3], Tianfu Ai [1,*] and Lie Yang [2,*]

[1] School of Mechatronics and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China; 3543@ecjtu.edu.cn (Y.L.); xcy15279912344@126.com (C.X.); 0429@ecjtu.edu.cn (Y.Z.)
[2] School of Advanced Manufacturing, Nanchang University, Nanchang 330031, China; chengzeng@ncu.edu.cn
[3] School of Naval Architecture, Ocean And Energy Power Engineering, Wuhan University of Technology, Wuhan 430063, China; hyingmin@126.com
* Correspondence: 3527@ecjtu.edu.cn (T.A.); lie.yang@ncu.edu.cn (L.Y.)

Abstract

The optimization of clutch engagement strategies is of great significance for improving vehicle power performance, fuel economy, and driving comfort. Traditional control strategies are difficult to adapt to complex working conditions and lack coordinated optimization of fuel and clutch. This paper proposes a multi-objective optimization method for clutch engagement strategies based on the Deep Deterministic Policy Gradient (DDPG) algorithm. A simulation environment is constructed, which includes a vehicle longitudinal dynamics model, clutch state switching logic, and a reinforcement learning agent. A multi-dimensional state space and action space are designed, and a composite reward function combining power performance, fuel economy, and comfort is developed to achieve multi-objective optimization of the fuel–clutch coordination curve. Experimental results show that the optimized engagement strategy significantly reduces sliding friction power (by 94.07%), power interruption speed (by 8.75%), and jerk (with a maximum reduction of 35.6%), while the average fuel consumption per distance is reduced by 0.39%. Through weight sensitivity analysis, it is found that when the weight of fuel economy is 0.3 and the weight of power performance is 0.5 (Scheme P5E3), the optimal balance among multiple objectives can be achieved. This study provides a new theoretical framework and engineering practice reference for the intelligent control of clutches.

Keywords: clutch optimization; deep reinforcement learning; DDPG algorithm; fuel–clutch coordination; multi-objective optimization

check for updates

Academic Editor: Stephan Rinderknecht

Received: 8 October 2025
Revised: 24 November 2025
Accepted: 27 November 2025
Published: 1 December 2025

Citation: Liu, Y.; Xie, C.; Zhang, Y.; Zeng, C.; Huang, Y.; Ai, T.; Yang, L. Research on Multi-Objective Optimization of Clutch Engagement Strategy Based on Deep Reinforcement Learning. *Vehicles* **2025**, *7*, 147. https://doi.org/10.3390/vehicles7040147

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Dry clutches, characterized by their simple structure, low cost, and convenient maintenance, have emerged as the ideal solution for manual transmissions [1]. Drivers achieve smooth gear-shifting operations by engaging and disengaging the clutch plates from the flywheel through manipulation of the clutch pedal. In these systems, precise control strategies are particularly crucial [2,3]. Common control methods include closed-loop control based on position feedback and adaptive control based on fuzzy logic [4–7]. The core objective of these methods is to optimize the engagement and disengagement processes of the clutch, but these strategies alone cannot ensure smooth power transmission and driving comfort. Therefore, this paper proposes a clutch and throttle coordinated control method based on deep reinforcement learning, which fully considers the coordination between the

clutch and the accelerator pedal in actual driving, providing new insights for solving the collaborative optimization problem between the clutch and the engine.

Research has indicated that dry clutches are prone to significant temperature increases in the friction discs during repeated engagement. This is especially true in applications such as Automated Manual Transmissions (AMT) and Dual Clutch Transmissions (DCT), where the non-linear and time-varying characteristics of friction make gear-shifting quality control difficult [8,9].

Therefore, thermal effects have become an important research topic in clutch control. Skugor et al. [10] proposed an open-loop torque control strategy for an E-type clutch, which introduced a thermal expansion compensation mechanism. Experiments conducted on a custom test bench showed a significant correlation between the displacement changes in the clutch slave cylinder (CSC) due to thermal expansion and the attenuation of the transmitted torque, highlighting the importance of thermal management in clutch control. Mario et al. [11] developed a lumped parameter thermal model for real-time prediction of the temperature of the clutch plates and buffer springs, and proposed a control strategy combining Model Predictive Control (MPC) with thermal compensation. Simulation results demonstrated that this strategy effectively improved vehicle launch quality under different initial temperatures and gradient road conditions.

Accurate real-time estimation of clutch torque transmission is crucial for high-precision control [12], especially for the coordinated control of the engine and dual clutches. Current mainstream methods typically rely on experimentally calibrated torque-displacement characteristics, converting the torque tracking problem into a position control issue for the release bearing [13]. However, due to the strong non-linear and time-varying characteristics of dry clutches, a fixed torque-displacement relationship struggles to accurately describe system dynamics [14].

Based on the dynamics model of a dual-intermediate-shaft DCT, Zhiguo Z et al. [15] developed an improved engine constant-speed control strategy based on the dynamics model of a dual-intermediate-shaft dual-clutch transmission (DCT). The strategy employs a dual-forgetting-factor recursive least squares (DFF-RLS) method to estimate the vehicle's resistance torque, uses a high-order sliding mode observer to reconstruct the angular acceleration signal, and estimates the clutch torque transmission via an unknown input observer. The optimal launch control is achieved by incorporating the minimum principle.

Focusing on an electric vehicle's three-speed mechanical automatic transmission, Peng W et al. [16] designed a torque loop based on adaptive control and a position loop based on Active Disturbance Rejection Control (ADRC), forming a cascaded dual-loop control architecture. Simulation results indicated that this structure effectively coped with non-linearity and disturbances, enhancing torque tracking performance.

To further improve clutch engagement quality, many scholars have conducted multi-objective optimization research. Vu T M et al. [17] designed a fuzzy logic controller that achieved efficient control of the engagement process, ensuring transmission efficiency while suppressing vibration and noise. Park J et al. [18] proposed a sliding friction speed optimization strategy aimed at minimizing the weighted sum of friction energy loss and jerk, significantly improving vehicle launch comfort. Chu et al. [19] proposed a variable slope clamping force loading strategy by analyzing the engagement dynamics of clutches under high-speed conditions. Combined with Support Vector Machine Regression (SVR) and Particle Swarm Optimization (PSO), this strategy optimized loading parameters through multi-objective optimization, effectively reducing engagement time and impact torque, and demonstrated better robustness under various working conditions.

The aforementioned research has primarily focused on the control of the clutch itself, without adequately considering the coordinated operation mechanism between the clutch

and the accelerator pedal in actual driving [20]. Experienced drivers typically maintain continuous power transmission and improve engagement efficiency through the coordination of the accelerator and clutch [21,22]. Therefore, existing control strategies still have shortcomings in terms of humanization and overall performance.

With the focus on the coordinated design of clutch engagement strategies and accelerator pedal control, this study introduces throttle closure control to simulate the manipulation behavior of skilled drivers. Based on this approach, we propose a clutch and throttle coordinated control method using the Deep Deterministic Policy Gradient (DDPG) algorithm. The main innovations of this method are embodied in the following points: the first application of the DDPG algorithm to the coordinated optimization of the clutch engagement process and throttle control by optimizing the high-dimensional continuous control parameters for clutch–throttle coordination, effectively solving the problem of traditional control strategies' insufficient adaptability under complex working conditions; the design of a composite reward function that comprehensively considers power performance, fuel economy, and comfort, achieving multi-objective optimization of the fuel–clutch coordination curve; and through systematic weight sensitivity analysis, determining the optimal balance point among different performance indicators, providing a theoretical basis and design guidance for engineering practice.

2. Modeling of Vehicle Powertrain System

To simulate the impact of different clutch control strategies on vehicles during gear shifting, this paper develops a vehicle forward simulation model as shown in Figure 1. The model mainly consists of seven key component models: vehicle longitudinal dynamics model, driveline model, transmission model, clutch model, engine model, Electronic Control Module (ECM), and driver model.

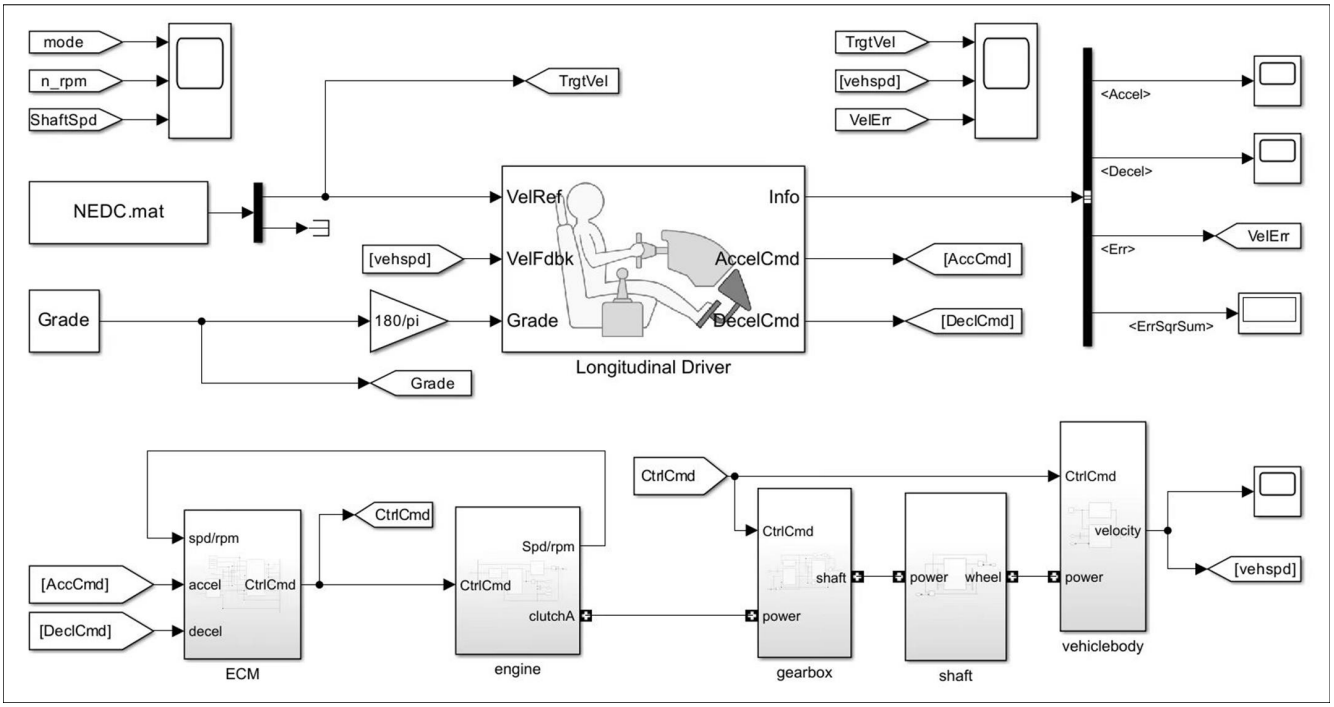


Figure 1. Vehicle Simulation Model.

The driver model uses a PID controller to generate drive and brake signals based on the deviation between the target speed and the actual speed. A clarification on the source of the throttle command is necessary. During normal vehicle operation (non-shifting conditions), the throttle command is generated by the driver model's PID controller to maintain the

target speed. However, during a gearshift event, initiated by a clutch disengagement signal, the throttle control is superseded by a pre-computed, optimized closure curve. The parameters for this curve are output by the DDPG agent, and the instantaneous throttle opening during the shift is determined by referencing this curve based on the elapsed time since the shift began. This creates a clear operational hierarchy where the PID controller handles steady-state tracking, and the DDPG-optimized trajectory handles the transient shift phase. These signals are then input into the ECM and the driving strategy. The ECM calculates the required fuel injection amount based on the drive and brake signals and outputs it to the engine module. A lookup function calculates the engine's output torque at this point. The torque is then controlled by the clutch to connect or disconnect the powertrain. After passing through the transmission, which changes the gear ratio, the torque is transmitted to the vehicle longitudinal dynamics model, causing the vehicle to accelerate until its speed matches the driver's required speed, thus simulating the vehicle's driving state.

To better simulate the working process of the clutch, this study designs the clutch state switching logic as shown in Figure 2 and constructs the corresponding actuator model in Simulink as shown in Figure 3. The model takes the engine output torque (τ_e), engine speed (ω_e), and clutch separation displacement (α_{clutch}) as inputs. It sequentially calculates the maximum static friction torque, the speed difference between the primary and secondary discs, and the system inertia moment, which are then input into the state switching unit. This unit, based on a finite state machine, judges and switches the working state of the clutch and outputs the torque transmitted by the clutch (T_c) in real time.

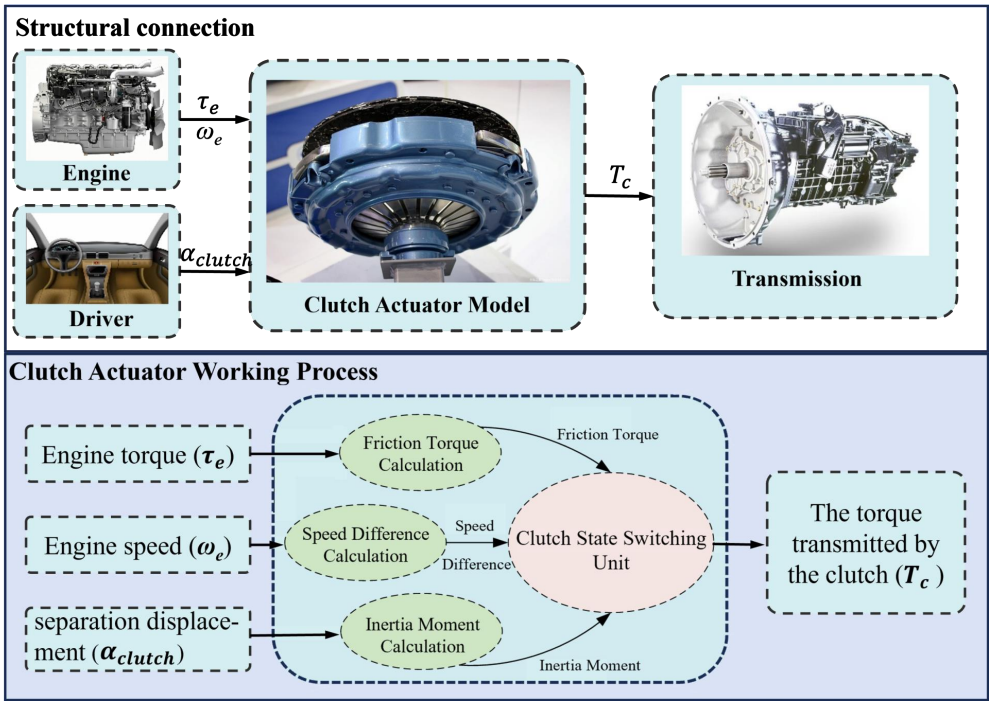


Figure 2. Clutch Actuator Working Process Diagram.

To accurately describe the transition process of the clutch state and establish a reasonable torque transmission mechanism, this paper encodes the clutch states: when the clutch is in the slipping or disengaged state, the state flag is set to 1; when it is in the locked state, the state flag is set to 0. In the early stage of clutch engagement, there is a speed difference between the primary and secondary discs, and the clutch is in the slipping or disengaged

state (state 1). The torque transmitted at this time is the kinetic friction torque, which can be expressed by Equation (1):

$$T_{\text{out}} = T_{\text{fric}} \times C_{\text{clutch}} \tag{1}$$

where T_{fric} represents the kinetic friction force of the friction plates, and C_{clutch} represents the clutch opening.

As the clamping force of the clutch gradually increases, the torque transmitted by the clutch also increases, and the rotational speeds of the primary and secondary friction plates gradually synchronize. When the speed difference between the primary and secondary friction plates is zero, the clutch will switch to the locked state, at which point there will be no relative sliding between the primary and secondary friction plates, and the transmitted torque will be the static friction torque.

The direction of the static friction force is determined by the relative magnitudes of the rotational speeds of the primary and secondary friction plates, while its magnitude can be calculated using Equations (2) and (3).

$$|T_{\text{out}}| = \min \begin{cases} \text{sign}(T_p - T_{\text{inertia}})T_{\text{fric}} \\ T_p - T_{\text{inertia}} \end{cases} \tag{2}$$

$$T_{\text{inertia}} = (J_e + J_{\text{up}})\alpha_{\text{clutch}} \tag{3}$$

where (T_p) represents the engine torque, (T_{inertia}) represents the inertial torque of the engine, (J_e) represents the engine moment of inertia, (J_{up}) represents the moment of inertia of the clutch driving plate, and (a_{clutch}) represents the angular acceleration of the clutch.

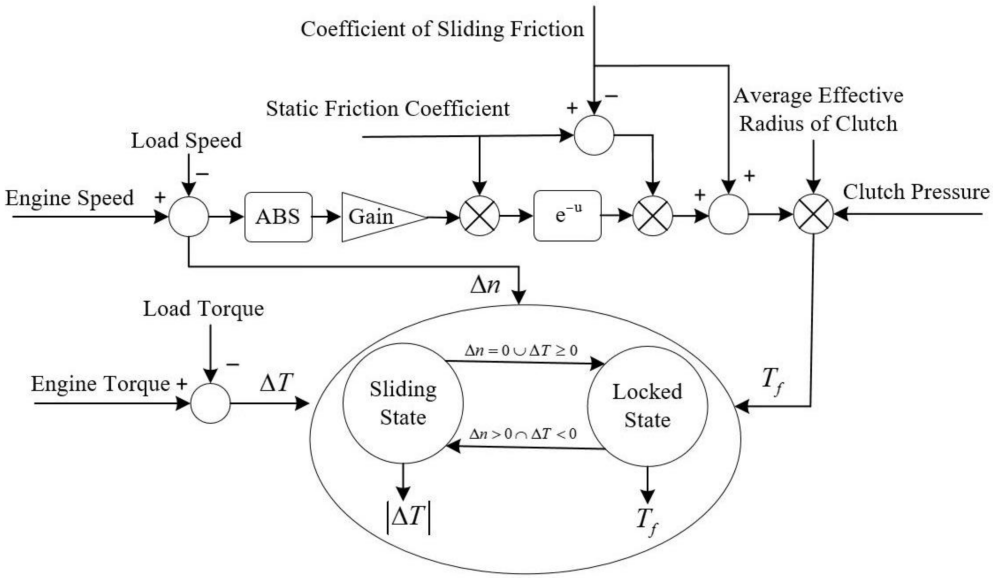


Figure 3. ClutchState Transition Diagram.

When the engine torque exceeds the clutch friction torque, the clutch will experience overload slipping, and the state of the clutch will switch from the locked state to the slipping state. Based on the above analysis, the clutch state transition diagram is designed in Simulink as shown in Figure 3.

3. Deployment and Optimization of Reinforcement Learning Algorithms

To enhance the performance of the fuel–clutch coordination strategy, accurately simulate the driver's operating habits, optimize the vehicle's power transmission efficiency, and improve vehicle performance and driving safety, this paper employs a variable slope

loading curve to simulate the driver's control process of the clutch and accelerator pedal during gear shifting. The vehicle simulation model is used as the training environment to deploy reinforcement learning algorithms to create an agent. The parameters of the variable slope loading curve are set as the action space of the agent, and the simulation results of the vehicle model are used as the state space of the agent. By setting up a reward and punishment mechanism with the goals of reducing torque shock, lowering speed shock, controlling speed fluctuation, and minimizing sliding friction power, the reinforcement learning agent is guided to adjust the key parameters of the clutch and throttle opening during the interactive training process with the training environment. This optimizes the fuel–clutch coordination strategy, improving gear-shifting efficiency while ensuring the smoothness of the gear-shifting process and enhancing the overall driving experience.

3.1. Principle of the DDPG Algorithm

In the field of reinforcement learning, traditional methods such as Q-Learning and DQN have achieved significant success in handling problems with discrete action spaces. However, these methods encounter difficulties when dealing with continuous action space problems. Continuous action spaces require iterative optimization at each step to find the optimal action, which is often impractical in real-world applications. To address this issue, the Deepmind team proposed the Deep Deterministic Policy Gradient (DDPG) algorithm in 2016. As summarized in Algorithm 1, this algorithm is based on the Deterministic Policy Gradient (DPG) algorithm and incorporates deep learning techniques, using deep neural networks as function approximators [23]. The DDPG algorithm improves the Actor-Critic framework and successfully applies it to continuous action spaces while ensuring the stability and efficiency of the algorithm. DDPG draws on the successful experience of DQN, employing experience replay and target network techniques to stabilize the learning process. Additionally, DDPG introduces batch normalization to enhance training stability and uses time-correlated noise generated by the Ornstein-Uhlenbeck process as an effective exploration strategy to promote the algorithm's exploration capabilities in real-world physical environments.

Algorithm 1 Deep Deterministic Policy Gradient (DDPG) Algorithm

Initialize Critic Network $Q(s, a; \theta_Q)$ and Actor Network $\mu(s; \theta_\mu)$
Initialize Target Network $Q'(s, a; \theta_{Q'})$ and $\mu'(s; \theta_{\mu'})$
Initialize Experience replay buffer R
for $episode = 1$ **to** M {Each Episode} **do**
  Initialize random process noise for action exploration
  for $t = 1$ **to** T {Each Time Step} **do**
    Select action according to the current policy and exploration noise:
      $a_t = \mu(s_t; \theta_\mu) + \text{noise}$
    Execute action a_t, observe reward r_t and new state s_{t+1}
    Store the transition (s_t, a_t, r_t, s_{t+1}) to R
    Randomly sample a batch of N transitions (s_i, a_i, r_i, s_{i+1})
    Compute the target value:
      $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}; \theta_{\mu'}); \theta_{Q'})$
    Update critic network by minimizing the loss:
      $L = \frac{1}{N} \sum (y_i - Q(s_i, a_i; \theta_Q))^2$
    Update target networks:
      $\theta_{Q'} \leftarrow \tau\theta_Q + (1 - \tau)\theta_{Q'}$,
      $\theta_{\mu'} \leftarrow \tau\theta_\mu + (1 - \tau)\theta_{\mu'}$
  end for
end for

The optimization of fuel–clutch coordination is a complex task involving numerous variables, requiring a comprehensive consideration of driving conditions, engine performance, driver habits, fuel efficiency, the balance between power and economy, as well as the control system and algorithms [24]. Given the complex interplay of these factors, the DDPG algorithm provides a viable solution suitable for this multi-dimensional optimization problem. The DDPG algorithm was selected for this optimization task due to its capability to handle problems with continuous, high-dimensional action spaces, which aligns with the nature of optimizing the 12 key parameters defining the clutch–throttle coordination curve . While it is recognized that the exploration noise (Ornstein-Uhlenbeck process) can be undirected and that the critic may suffer from overestimation bias, several factors contributed to its effective application here. The physically meaningful constraints imposed on the action space inherently restricted invalid explorations. Furthermore, the use of target networks and soft updates helped stabilize training and mitigate potential value overestimation. Therefore, this paper constructs a throttle-clutch engagement curve optimization model based on the DDPG algorithm, with the system architecture shown in Figure 4.

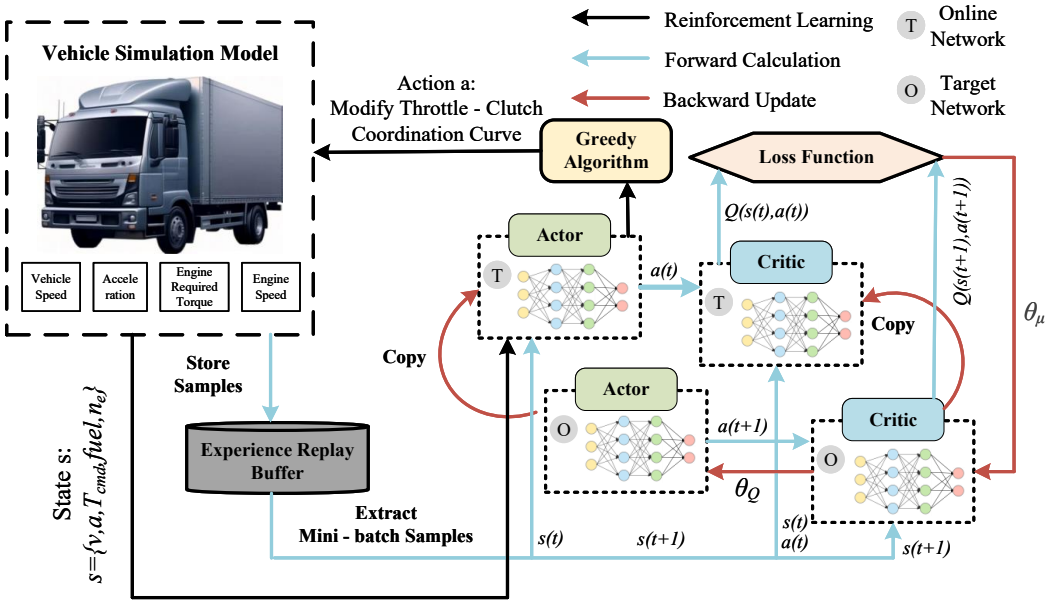


Figure 4. DDPG Throttle-Clutch Coordination Strategy Optimization Framework.

To simulate the control process of the clutch and accelerator pedal by skilled drivers during gear shifting, this paper employs a variable slope loading curve to mimic the driver's actions, as shown in Figure 5. A key consideration for the practical implementation of the optimized clutch–throttle curves is their interaction with the driver. The curves are not designed for direct, precise replication by a human driver. Rather, they represent an optimal mapping that can be embedded within the vehicle's transmission control unit (TCU). In this scenario, the driver's pedal inputs (clutch travel, accelerator position) serve as commands that are then translated by the TCU, using a lookup table derived from the optimized curves, into the precise actuator controls. This allows drivers to achieve near-optimal shift quality through simple, smooth pedal operations, as the system automatically executes the complex coordination strategy. Thus, the proposed method aims to emulate skilled driver behavior through automated control, simplifying the driver's task while enhancing overall vehicle performance. The clutch engagement curve and throttle closure curve are composed of three segments with different slopes. The key points on the clutch curve are determined as $(t_1,0)$, (t_2,a_1), (t_3,a_2), $(t_4,1)$. The key points on the throttle curve are $(t_5,0)$, (t_6,β_1), (t_7,β_2), $(t_8,1)$.

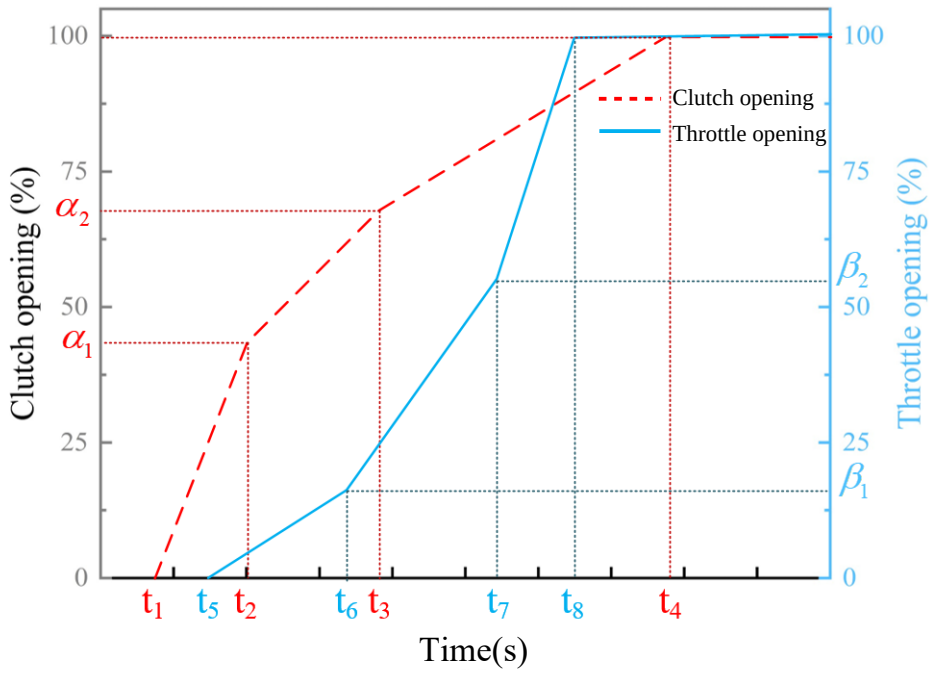


Figure 5. Clutch and Throttle Closure Curves.

In the optimization task of the clutch–throttle coordination curve, the agent is responsible for optimizing the key parameters of the curve. The agent selects actions from the action space and interacts with the vehicle model environment; the interaction experience data obtained from this process is then stored in the experience replay buffer R. As shown in Equation (4), to strike a balance between the exploitation of existing strategies and the exploration of new actions, the agent adopts a noisy action strategy.

The definition of the action space is given by Equation (5). Under the premise of satisfying the aforementioned constraints, the agent outputs continuous actions — selecting four values within the interval [0, 1] to represent the opening degrees of the clutch and the accelerator pedal; and selecting eight values within the interval [0, 2] as the time parameters of the key points. Its specific expression is as follows:

$$act = \begin{cases} \{\alpha_1, \alpha_2, \beta_1, \beta_2, t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8\} \\ \alpha_1, \alpha_2, \beta_1, \beta_2 \in [0, 1] \\ t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8 \in [0, 2] \end{cases} \tag{4}$$

$$a_t = \mu(s_t \mid \theta^{\mu}) + N_t \tag{5}$$

At the start of each training round, a small batch of transition samples is randomly sampled from the experience replay buffer R, denoted as:

$$\{(s_i, a_i, r_i, s_{i+1})\} \in R \tag{6}$$

Subsequently, the critic network is updated using the Temporal Difference (TD) mechanism. First, the target Q-values y_i for each sample are calculated via the target network:

$$y_i = r_i + \gamma Q'\left(s_{i+1}, \mu'\left(s_{i+1} \mid \theta^{\mu'}\right) \mid \theta^{Q'}\right) \tag{7}$$

The loss function of the critic network is given by Equation (8), which is used to measure the mean squared error between the current Q-estimates $Q(s_i, a_i \mid \theta^Q)$ and the target Q-values y_i:

$$L\left(\theta^Q\right) = \frac{1}{N}\sum_i \left(y_i - Q\left(s_i, a_i \mid \theta^Q\right)\right)^2 \tag{8}$$

By minimizing this loss through gradient descent, the parameters of the critic network θ^Q are updated, making the action value evaluation under policy μ more accurate. The actor network is updated based on policy gradients, and its gradient expression is shown in Equation (9):

$$\nabla_{\theta^\mu} J \approx \frac{1}{N}\sum_i \nabla_a Q\left(s, a \mid \theta^Q\right)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu}\mu(s \mid \theta^\mu)_{s_i} \tag{9}$$

Fine-tuning the parameters of the actor network along the gradient direction of the Q-values with respect to actions enables the output actions to achieve higher Q-value scores. Subsequently, the parameters of the target network are slowly aligned using a soft update mechanism, and the update rule is shown in Equations (10) and (11):

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'} \tag{10}$$

$$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'} \tag{11}$$

Among them, the soft update coefficient τ is fixed at 0.001, enabling the target network to gradually follow the online network in the form of an exponential moving average. This effectively suppresses fluctuations in Q-value estimation and significantly improves the stability of the training process.

In each training cycle, the agent outputs a set of parameters as an action and injects them into the vehicle model for simulation execution. The monitor module collects vehicle operating state data in real time and immediately feeds it into the reward function to evaluate the quality of the current action.

3.2. Constructing a Hybrid Reinforcement Learning Model

A skilled driver should complete the clutch operation within 2 s, so the constraints ($t_4 = 2$) and ($t_8 = 2$) are set. Based on the key points, the variable slope engagement equations for the clutch and throttle pedals can be expressed as:

$$\alpha_c = \begin{cases} \dfrac{\alpha_1}{t_2 - t_1}(t - t_1) & t_1 \le t < t_2 \\ \dfrac{\alpha_2 - \alpha_1}{t_3 - t_2}(t - t_2) + \alpha_1 & t_2 \le t < t_3 \\ \dfrac{1 - \alpha_2}{2 - t_3}(t - t_3) + \alpha_2 & t_3 \le t \end{cases} \tag{12}$$

$$\alpha_p = \begin{cases} \dfrac{\beta_1}{t_6 - t_5}(t - t_5) & t_5 \le t < t_6 \\ \dfrac{\beta_2 - \beta_1}{t_7 - t_6}(t - t_6) + \beta_1 & t_6 \le t < t_7 \\ \dfrac{1 - \beta_2}{2 - t_7}(t - t_7) + \beta_2 & t_7 \le t \end{cases} \tag{13}$$

To constrain the basic shape of the closure curves and ensure that the clutch operates within the allowable range, the following constraints are established:

$$0 < \alpha_1 < \alpha_2 < 1 \quad 0 < \beta_1 < \beta_2 < 1 \tag{14}$$

To ensure the coordination between the throttle and clutch, the following constraints are established:

$$t_1 < t_2 < t_3 < t_4 \quad t_5 < t_6 < t_7 < t_8 \tag{15}$$

Meanwhile, considering human reaction time and the response speed of vehicle equipment, the slopes of the closure curves should be within a reasonable range, so the following constraints are established:

$$k_{\min} \leq \frac{\alpha_1}{t_2 - t_1}, \quad \frac{\alpha_2 - \alpha_1}{t_3 - t_2}, \quad \frac{1 - \alpha_2}{t_4 - t_3} \leq k_{\max} \tag{16}$$

$$k_{\min} \leq \frac{\beta_1}{t_6 - t_5}, \quad \frac{\beta_2 - \beta_1}{t_7 - t_6}, \quad \frac{1 - \beta_2}{t_8 - t_7} \leq k_{\max} \tag{17}$$

During gear shifting, the difference in rotational speed between the primary and secondary discs of the clutch can be significant depending on the selected gear. When performing an upshift, the engine speed exceeds the speed of the clutch secondary disc, and the driver should first operate the clutch pedal and then the accelerator pedal. Conversely, during a downshift, the engine speed is lower than the speed of the clutch secondary disc, and the driver should first operate the accelerator pedal and then the clutch pedal. Therefore, a set of additional constraints as shown in Equation (18) is defined to ensure that the driver's operation sequence correctly matches the speed relationship between the engine and the clutch during gear shifting. This reduces vehicle damage or performance degradation caused by improper operation, ensuring smooth gear shifting and vehicle performance.

$$\begin{cases} t_5 > t_1 & \Delta_{\text{gear}} > 0 \\ t_5 > t_1 & \Delta_{\text{gear}} < 0 \end{cases} \tag{18}$$

In the optimization of the fuel–clutch coordination curve, the agent is mainly responsible for optimizing the key parameters of the fuel–clutch coordination curve. Under the above constraints, the agent takes continuous actions, selecting four values within the continuous space of 0 to 1 as the corresponding clutch and throttle pedal openings for each state; and selecting eight values within the continuous space of 0 to 2 as the key point time parameters, as shown in the following equation:

$$act = \begin{cases} \{\alpha_1, \alpha_2, \beta_1, \beta_2, t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8\} \\ \alpha_1, \alpha_2, \beta_1, \beta_2 \in [0, 1] \\ t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8 \in [0, 2] \end{cases} \tag{19}$$

The network structure and training hyperparameters are presented in Table 1. The Actor network, which maps states to actions, was constructed with the following architecture: an input layer corresponding to the state dimension, followed by two fully connected hidden layers with 64 neurons each, using ReLU activation functions. The output layer has 12 neurons (corresponding to the 12-dimensional action vector) with a tanh activation function, whose outputs were subsequently scaled to the appropriate ranges defined for the clutch and throttle parameters.

The Critic network, which estimates the Q-value, was designed as follows: an input layer receiving the concatenated state and action vectors, followed by two fully connected hidden layers with 64 neurons each (ReLU activation), and a single linear output neuron for the Q-value. At the beginning of each training round, the agent selects a set of parameters as an action to be executed and inputs them into the vehicle model for simulation calculation.

The monitor module then supervises and records the dynamic parameters during the vehicle's motion and passes them to the reward function to evaluate the quality of the action.

Table 1. Network Structure and Training Hyperparameters.

Category	Parameter	Value
Network Architecture	Actor Network Structure	Input layer → 64 → 64 → Output layer
	Actor Activation Functions	Hidden: ReLU, Output: tanh
	Critic Network Structure	Input layer (state+action) → 64 → 64 → Output layer
	Critic Activation Functions	Hidden: ReLU, Output: Linear
Training Parameters	Actor Learning Rate	0.001
	Critic Learning Rate	0.01
	Discount Factor (γ)	0.99

3.3. Design of the Reward and Punishment Mechanism

In reinforcement learning, the reward and punishment mechanism refers to guiding the learning process of an agent by providing it with rewards or punishments. This mechanism is a core component of reinforcement learning algorithms, used to evaluate the quality of an agent's actions in a particular state and adjust the agent's behavioral strategy based on the evaluation results. The fundamental principle behind the design of the reward and punishment mechanism is to maximize the agent's long-term cumulative reward [25]. Through the reward mechanism, the agent's actions are assessed and feedback is provided, guiding the agent towards the desired goals.

Rewards are typically represented as scalar values, which can be positive, negative, or zero. Positive rewards indicate that the agent's action is positive or correct, contributing to the achievement of the goal. Negative rewards signify that the action is negative or incorrect, potentially hindering the agent's ability to achieve the goal. Zero rewards imply that the action has no significant impact. The reward and punishment mechanism balances exploration and exploitation. By setting appropriate reward signals, the agent is encouraged to explore unknown areas and discover new effective strategies. At the same time, the mechanism also guides the agent to exploit known effective strategies to obtain higher rewards. The design of the reward function, particularly the use of metrics compared to the previous cycle, requires further justification. While normalized metrics or fixed thresholds could enhance generalization across different vehicles or conditions, they presented a practical challenge during preliminary investigations. Specifically, in the early stages of training with random exploration, the agent could produce policies resulting in exceptionally high values of metrics like Wloss. Using such extreme values for normalization would compress the reward signal for subsequent, better-performing policies, hindering the learning process. Therefore, to provide a stable and consistent learning signal that guides the agent toward progressive improvement, we employed a relative comparison approach. This method rewards the agent for outperforming its previous attempt, effectively driving policy refinement within the context of this specific simulation environment.

During the gear-shifting process in a vehicle, the closing speed of the clutch significantly affects performance. If the clutch closes too quickly, it may generate a large impact torque, which can cause the engine to stall in severe cases. Additionally, rapid closure can lead to significant speed shocks, reducing driving comfort. Conversely, if the clutch closes too slowly, the vehicle will experience prolonged power loss, affecting its power performance and increasing fuel consumption, which is detrimental to fuel economy [26].

Therefore, the reward function is designed from three aspects: power performance, fuel economy, and comfort. The specific reward and punishment mechanism is as follows:

$$R_{\text{all}} = R_p + R_e + R_c \tag{20}$$

where R_{all} represents the cumulative reward, R_p represents the power performance reward, R_e represents the fuel economy reward, and R_c represents the comfort reward.

3.3.1. Power Performance Reward Function Design

During gear shifting, when the clutch disengages and re-engages, the vehicle speed decreases due to the interruption of power. If the decrease is too large, it will affect the vehicle's power performance. Additionally, during the stage where the primary and secondary discs are in contact but not yet synchronized, the friction between the friction discs generates sliding friction power, which not only affects the clutch's service life but also the efficiency of power transmission. Therefore, to meet the power performance requirements, the power performance reward function is established as follows:

$$R_p(t,a,b) = \begin{cases} 1 & \Delta v_{\max} < \Delta v'_{\max} \\ -0.5 & \Delta v_{\max} \geq \Delta v'_{\max} \\ 1 & W_{\text{loss}} < W'_{\text{loss}} \\ -0.5 & W_{\text{loss}} \geq W'_{\text{loss}} \end{cases} \tag{21}$$

where $\Delta v_{\max}$ represents the difference between the vehicle speed before clutch disengagement and the lowest speed after disengagement in the current calculation round, $\Delta v'_{\max}$ represents the difference between the vehicle speed before clutch disengagement and the lowest speed after disengagement in the previous calculation round, W_{loss} represents the sliding friction power generated during the sliding phase in the current calculation round, and W'_{loss} represents the sliding friction power generated during the sliding phase in the previous calculation round. According to the calculation method in reference [27], the sliding friction power generated by the clutch during the sliding phase can be expressed as follows:

$$W_{\text{loss}} = \int_{t_s}^{t_e} (\omega_1 - \omega_2) T_c dt \tag{22}$$

where ω_1 represents the angular velocity of the primary disc, ω_2 represents the angular velocity of the secondary disc, T_c represents the clutch friction torque, t_s represents the time when the primary and secondary discs of the clutch begin to contact, and t_e represents the time when the clutch is fully locked.

3.3.2. Fuel Economy Reward Function Design

During gear shifting, when the clutch disengages and re-engages, the interruption of power transmission between the primary and secondary discs of the clutch leads to a longer duration of power interruption. The longer the interruption, the higher the additional fuel injection amount of the engine. Therefore, to meet the fuel economy requirements, the power interruption time and the fuel consumption per unit distance

during the simulation process are used as observation indicators to establish the sub-objective function as follows:

$$R_e(t,a,b) = \begin{cases} 1 & \Delta t_{\text{brk}} < \Delta t'_{\text{brk}} \\ -0.5 & \Delta t_{\text{brk}} > \Delta t'_{\text{brk}} \\ 1 & \text{fuel} < \text{fuel}' \\ -0.5 & \text{fuel} > \text{fuel}' \end{cases} \tag{23}$$

where Δt_{brk} represents the difference between the time when the clutch disengages and the time when the vehicle speed returns to the level before disengagement; fuel represents the fuel consumption per unit distance during the simulation process.

3.3.3. Comfort Reward Function Design

In the study of automotive clutches, jerk is an important indicator for evaluating the smoothness of the gear-shifting process. Fundamentally, it is the disturbance of torque that causes the appearance of shocks. Jerk represents the rate of change of the vehicle's forward direction, so to meet the comfort requirements, the sub-objective function is established as follows:

$$R_e(t,a,b) = \begin{cases} 1 & J_v < J'_v \\ -0.5 & J_v > J'_v \end{cases} \tag{24}$$

According to the calculation method in reference, the vehicle speed jerk during the clutch closing process can be expressed as:

$$J_v = \frac{\partial a_v}{\partial t} \tag{25}$$

where a_v represents the vehicle's longitudinal acceleration, J_v represents the vehicle speed jerk, and its magnitude reflects the strength of the jolt felt by the driver during the driving process.

In summary, considering power performance, fuel economy, and comfort, three sub-objective functions are designed. The first sub-objective function focuses on the speed drop and sliding friction power during the gear-shifting process to ensure that power performance is not affected. The second sub-objective function focuses on the power interruption time and fuel consumption per unit distance during the clutch disengagement and re-engagement process to improve fuel economy. The third sub-objective function focuses on the vehicle speed jerk to evaluate the smoothness of the gear-shifting process and reduce the jolt felt by the driver.

4. Simulation Experiments and Result Comparison

4.1. Design of Experimental Schemes

The reference speed profile for the driver model was a uniform acceleration, and training focused solely on optimizing the control actions for a single upshift event occurring at approximately 68 km/h. To normalize the reward function in Equation (20) and introduce a weighting mechanism to emphasize the importance of different optimization goals, this study focuses on commercial vehicles, which prioritize power performance and fuel economy while also considering comfort. Therefore, the weight for the comfort objective is fixed at 0.2, and the weights for the power performance and fuel economy objectives are varied systematically to explore vehicle performance under different requirements. This section designs five scenarios to investigate the impact of weight changes on optimization results. Scenarios 1 to 4 primarily address different weight combinations for power perfor-

mance and fuel economy to assess their effects on commercial vehicle performance. Each scenario has distinct weight settings to reflect varying performance needs and priorities. Additionally, Scenario 5 serves as a comparative scenario with a higher weight for the comfort objective to compare vehicle performance under high comfort requirements with the other scenarios. The specific weight settings are detailed in Table 2.

Table 2. Weight Settings for Each Scheme.

Scheme	Power (P)	Economy (E)	Comfort (C)
P1E7	0.1	0.7	0.2
P3E5	0.3	0.5	0.2
P5E3	0.5	0.3	0.2
P7E1	0.7	0.1	0.2
P2E2C6	0.2	0.2	0.6

Based on the five scenarios mentioned above, the reward function weights are set accordingly, and the DDPG algorithm is deployed. After iterating for 300 rounds for each scenario, the cumulative average reward values during the iteration process are recorded, as shown in Figure 6.

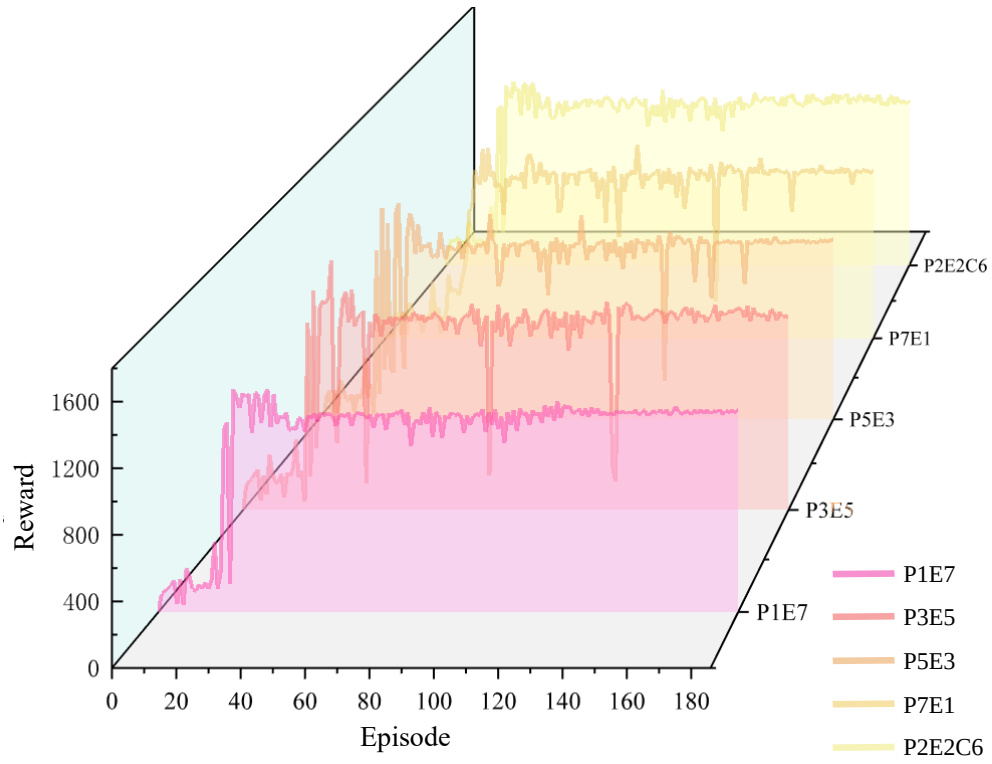


Figure 6. Original Closing Curve of Clutch and Throttle.

The figure indicates that the cumulative average reward values have converged, signifying that the optimization is essentially complete. The optimized closure curves are illustrated in Figure 7.

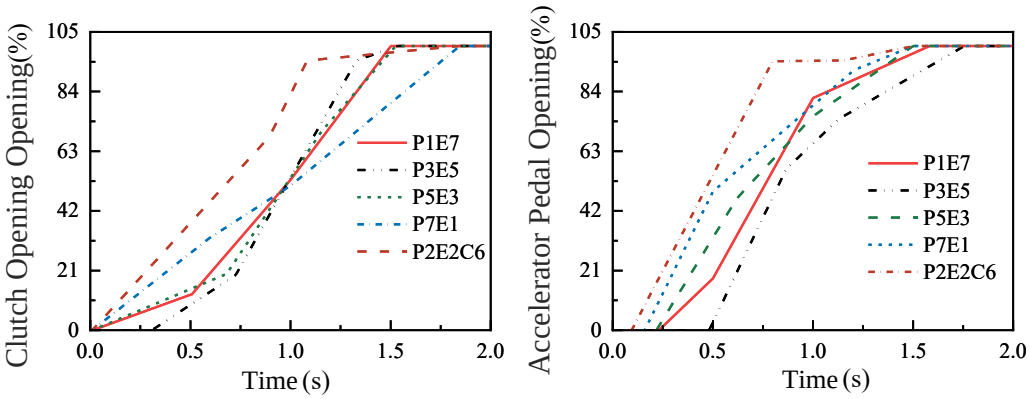


Figure 7. Clutch and Throttle Pedal Curves for Different Optimization Schemes.

4.2. Comparison of Engagement Characteristics After Optimization

The optimized curves were substituted into the dynamic simulation model of the clutch to simulate the separation and engagement process of the clutch during gear shifting, in order to verify the effectiveness of the optimization results. The simulation data from the model are shown in Figure 8, and the main indicators for each scheme are summarized in Table 3.

Table 3. Comparison of Key Performance Indicators for Each Scheme.

Scheme	Original	P1E7	P3E5	P5E3	P7E1	P2E2C6
Power-interruption duration (s)	3.84	4.94	4.49	4.32	4.02	3.49
Speed during Power Interruption (m/s)	0.16	0.19	0.18	0.146	0.145	0.13
Maximum Slip Friction Work (J)	9764.5	573.39	548.91	578.85	569.75	584.05
Jerk count	4	3	4	3	2	2

Figure 8a illustrates the changes in vehicle speed during the gear shifting process. At 60.86 s, the clutch begins to disengage, temporarily interrupting the power transmission and causing a drop in vehicle speed. Once the gear shifting operation is completed, the clutch re-engages according to different closure curves, restoring power and gradually increasing the vehicle speed. The time required for the vehicle speed to return to its original level after the clutch disengages is defined as the power interruption time. It can be seen from the figure that as the weight of the power performance objective function gradually increases, the power interruption time gradually decreases but is still longer than before optimization. The scheme with the highest power performance weight, P7E1, has a power interruption time that is 4.6% longer than before optimization. However, the speed fluctuation during the gear shifting process is reduced by 9.4% compared to before optimization. The preliminary analysis suggests that this is mainly because scheme P7E1 takes comfort into account, resulting in a slower clutch closure rate than before, which slightly delays the acceleration time but reduces speed fluctuations, thereby meeting the power performance requirements. The results for scheme P7E1, which prioritizes power performance, warrant further discussion. Although the power interruption time increased slightly compared to the original strategy, this does not indicate a degradation in overall power performance. Analysis of the optimized curves (Figure 7) reveals a control strategy of 'slower clutch engagement paired with faster throttle recovery.' This trade-off was made to achieve more significant improvements in other critical power-related metrics: a substantial reduction in speed fluctuation (9.4%) and a drastic decrease in slip friction work (94%). The former ensures smoother power transmission and better drivability, while the latter improves powertrain efficiency and durability. Thus, the P7E1 scheme enhances the

overall quality and efficiency of power performance, demonstrating the multi-objective optimizer's ability to make intelligent compromises between competing goals.

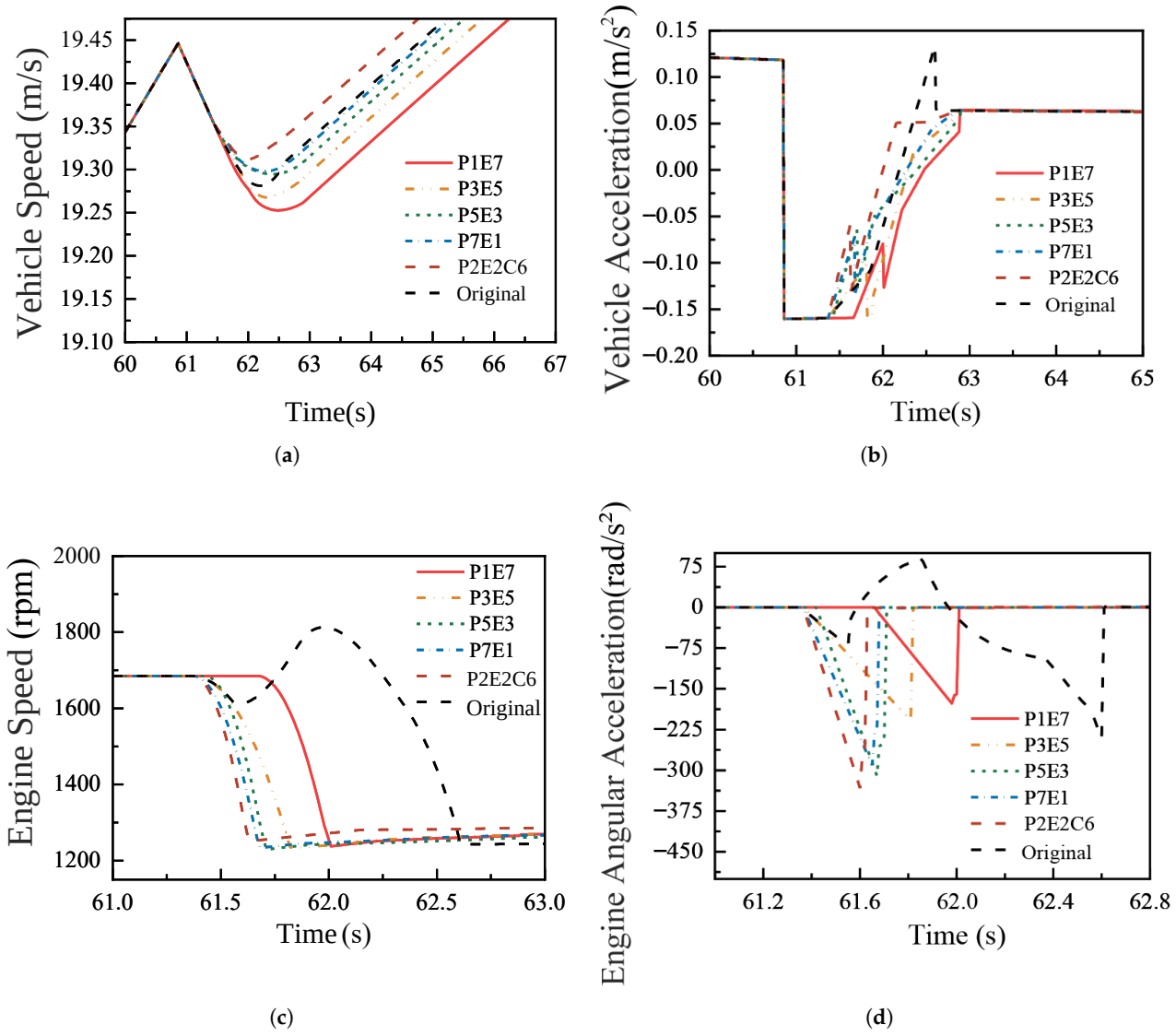


Figure 8. Simulation results of clutch engagement process: (**a**) vehicle speed; (**b**) vehicle acceleration; (**c**) engine speed; (**d**) engine angular acceleration.

Figure 8b records the changes in vehicle acceleration during the gear shifting process. Without optimization, there is a significant inflection point in vehicle acceleration at 62.57 s. After optimization, this inflection is replaced by a smoother acceleration curve, with lower power performance weights resulting in smaller curve slopes. This indicates that the optimization process effectively smoothed the acceleration changes during gear shifting, thereby improving driving comfort.

Figure 8c shows the changes in engine angular velocity. In the unoptimized state, the engine speed has a significant peak of about 1800 rpm, which may be due to the clutch closing too early after gear shifting, causing the load speed transmitted by the drive shaft to exceed the engine speed and thus causing engine speed fluctuations. However, after optimization, the changes in engine speed are much smoother, with no significant fluctuations.

Figure 8d records the changes in engine angular acceleration. In the unoptimized state, the engine angular acceleration has three fluctuations, but after algorithm optimization, the number of fluctuations is reduced to one, significantly improving the vehicle's smooth-

ness. Specifically, the optimized closure curve of scheme P1E7 reduces the maximum fluctuation amplitude of angular acceleration by 26.7%, while the optimized closure curve of scheme P5E3, which is inclined towards power performance, increases the maximum fluctuation amplitude of angular acceleration by 28.1% compared to before optimization but reduces the number of fluctuations, thus balancing driving comfort to some extent.

Figure 9b shows the amount of sliding friction power during the clutch closure process. The optimized results indicate a significant reduction in sliding friction power, from the original 9764.5 J to about 600 J. Scheme P3E5 has the largest reduction, with sliding friction power reduced to 5.62% of the original, while the optimized closure curve of scheme P7E1, which is inclined towards power performance, reduces sliding friction power to 5.88% of the original. As the weight of the power performance objective function increases, the sliding friction power generated during clutch closure gradually decreases. This may be because the relative speed between the primary and secondary discs of the clutch is lower when they engage after optimization, thereby reducing the generation of sliding friction power. By integrating vehicle speed to obtain the distance traveled and calculating the ratio of engine fuel injection to distance traveled, the fuel consumption per unit distance can be obtained.

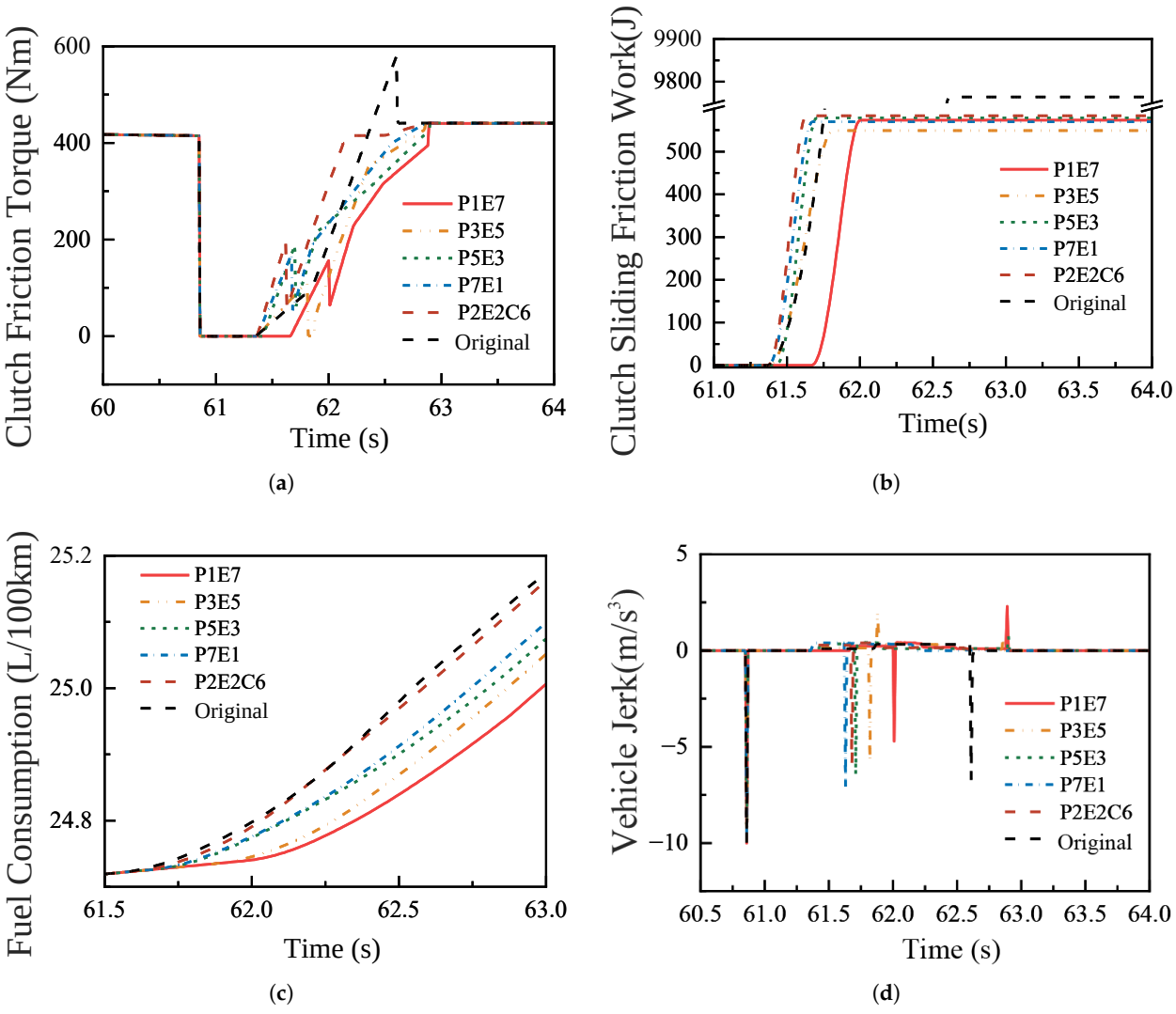


Figure 9. Simulation results of the clutch engagement process: (**a**) clutch friction torque; (**b**) clutch friction work; (**c**) fuel consumption per 100 km; (**d**) vehicle jerk.

Figure 9c records the fuel consumption per unit distance under different optimized curves. The results show that after applying the optimization algorithm, the vehicle's fuel consumption per unit distance is reduced compared to the unoptimized state, and the optimized curve with a preference for fuel economy achieves lower fuel consumption per unit distance.

In summary, by setting five different weight reward functions, five sets of closure curves were obtained, and curves with different weights have different performances in terms of power performance, fuel economy, and comfort, all of which are better than before optimization.Considering all three aspects, scheme P5E3 is chosen as the final optimized "fuel–clutch coordination" curve, as shown in Figure 10. The optimal weight combination (P5E3) was identified through a comparative analysis of the results from several pre-defined weighting schemes, rather than through a formal optimization of the weights themselves. The P5E3 scheme was selected because it achieved the most favorable and balanced trade-off across all three performance objectives—power, economy, and comfort—as evidenced by the comprehensive data in Table 2 and Figures 8 and 9.

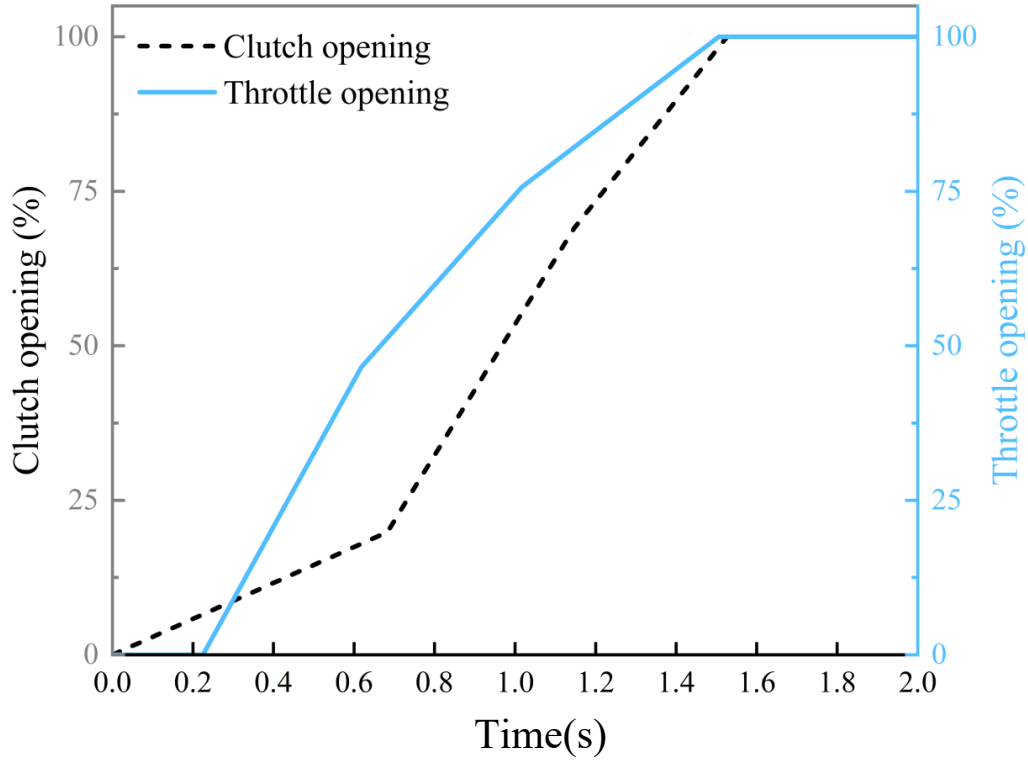


Figure 10. The final optimized results of the throttle-clutch coordination curve.

After applying the best optimized "fuel–clutch coordination" curve, the fuel consumption is reduced by an average of 0.39%, the maximum vehicle impact is reduced by 35.6%, the maximum sliding friction power is reduced to 5.93% of the original, and the power interruption speed is reduced by 8.75%. Therefore, it has a good performance in terms of fuel economy, comfort, and power performance.

5. Conclusions and Future Work

This study proposes a multi-objective optimization framework for clutch engagement strategies based on the Deep Deterministic Policy Gradient (DDPG) algorithm, aiming to enhance vehicle power performance, fuel economy, and driving comfort in a coordinated manner. By integrating a high-fidelity vehicle longitudinal dynamics model with a detailed clutch state transition mechanism, a reinforcement learning simulation environment was

constructed to accurately simulate real-world driving conditions and driver behavior. A hybrid reward function was designed, incorporating multiple competing objectives such as sliding friction power, power interruption, and jerk in a weighted form to achieve effective balance among different performance metrics.

Experimental results indicate that the proposed DDPG-based control strategy significantly improves the overall performance of the clutch engagement process: sliding friction power is reduced by 94.07%, power interruption speed is decreased by 8.75%, maximum jerk is lowered by 35.6%, and average fuel consumption per 100 km is also reduced by 0.39%. This demonstrates the effectiveness of the strategy in multi-objective optimization, achieving a good balance among key performance indicators.

Through systematic sensitivity analysis of the reward function weights, it was found that the overall performance is optimal when the fuel economy weight is 0.3 and the power performance weight is 0.5 (Scheme P5E3), achieving excellent comprehensive performance without significantly sacrificing any single performance indicator. This conclusion provides a practical reference for the calibration of performance targets in the clutch control systems of commercial vehicles, especially in scenarios where both driving quality and fuel efficiency are key requirements.

The main contributions of this study include: first, the development of a novel DDPG-driven optimization framework integrating throttle-clutch coordination mechanisms; second, the design of a multi-objective reward structure that fully reflects real-world driving requirements; and third, the validation of the proposed method's consistent improvement in multiple performance metrics through high-fidelity simulation. It is important to note the scope of the training conditions for this study. The DDPG agent was trained and evaluated on a specific, standardized shift scenario to validate the core methodology. The reference speed profile for the driver model was a uniform acceleration, and training focused solely on optimizing the control actions for a single upshift event occurring at approximately 68 km/h. This approach allowed for a controlled analysis of the algorithm's capability to improve shift quality. While this demonstrates the potential of the framework, it is recognized that the generalizability of the policy to diverse driving cycles (e.g., urban, highway) and varying initial states requires further investigation. Ensuring robustness across a wider operating envelope is a key objective for subsequent research.

Despite the positive experimental results, the current study still has certain limitations, including the idealized assumption of actuators, incomplete coverage of dynamic uncertainties, and external disturbances (such as changes in road gradient and extreme thermal conditions). While the simulation results demonstrate the strong potential of the proposed method, this study has certain limitations that must be addressed for practical implementation. The current simulation environment does not account for critical real-world factors such as clutch temperature variations (which affect friction coefficient and thermal deformation) and changes in road gradient (which alter vehicle load torque). These disturbances could significantly impact the performance and robustness of the learned control policy. Therefore, future work will explicitly focus on enhancing the system's robustness. This will involve: (1) integrating a high-fidelity thermal model of the clutch into the simulation environment; (2) introducing external disturbances, including varying road profiles and load conditions, during training; and (3) exploring robust or adaptive reinforcement learning techniques to ensure consistent performance under uncertainty. The final validation will be conducted through Hardware-in-the-Loop (HIL) testing and ultimately, real vehicle experiments. A limitation of the current study is the absence of a direct quantitative comparison with other control strategies, such as finely-tuned PID controllers, Model Predictive Control (MPC), or alternative reinforcement learning algorithms. The primary focus of this work was to establish the feasibility of the DDPG-based

framework as a novel alternative to traditional experimental calibration for clutch–throttle coordination. Future work will include a comprehensive comparative analysis against these benchmark strategies to rigorously evaluate the relative advantages, computational efficiency, and performance trade-offs of the proposed method.

In summary, this study provides an innovative and feasible technical pathway for intelligent clutch control based on deep reinforcement learning, offering both theoretical significance and engineering value in the optimization of vehicle powertrains. The methodology presented in this study, centered on a DDPG-driven optimization framework, demonstrates significant potential for generalizability and scalability. While the current validation is focused on a specific dry clutch model, the core architecture—comprising a vehicle simulation environment, a multi-objective reward function, and an agent optimizing continuous control parameters—is inherently flexible. This suggests its potential applicability to other types of transmission systems (e.g., automated manual transmissions) and to the coordinated control of power sources in hybrid electric vehicles. The practical validation of this potential on diverse powertrain systems constitutes a key and necessary direction for our future work. in addition, future research will explore systematic methodologies, such as Pareto front analysis or automated hyperparameter optimization, to more rigorously determine the optimal reward function weights for specific vehicle performance targets.

Author Contributions: Conceptualization and methodology, Y.L., C.X. and T.A.; formal analysis, Y.Z. and C.Z.; writing—review and editing, Y.H. and L.Y.; data curation, Y.L. and C.X.; validation, Y.Z., C.Z. and T.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Early Career Youth Science and Technology Talent Project of Jiangxi Province under Grant Nos. 20244BCE52130 (Y.L.) and 20244BCE52124 (T.A.), and by Natural Science Foundation of Jiangxi Province under Grant Nos. 20242BAB20248 (T.A.), 20252BAC240364 (T.A.) and 20242BAB20059 (C.Z.).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

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

References

1. Hong, J.; Gao, B.; Yue, H.; Chen, H. Dry Clutch Control of Two-Speed Electric Vehicles by Using an Optimal Control Scheme With Persistent Time-Varying Disturbance Rejection. *IEEE Trans. Transp. Electrif.* **2021**, *7*, 2034–2046. [CrossRef]
2. Xue, Q.; Zhang, X.; Chen, H.; Yue, M.; Teng, T.; Yu, J. Dynamic Coordinated Control Strategy of a Dual-Motor Hybrid Electric Vehicle Based on Clutch Friction Torque Observer. *Heliyon* **2024**, *10*, e27255. [CrossRef]
3. Sciuto, L. G.; Kowol, P.; Capizzi, G. Modeling and Experimental Characterization of a Clutch Control Strategy Using a Magnetorheological Fluid. *Fluids* **2023**, *8*, 145. [CrossRef]
4. Darrell, R. Coordinated Torque, Energy and Clutch Control Strategy for Downshifts in P2 Parallel xHEV Powertrains. *SAE Int. J. Adv. Curr. Pract. Mobil.* **2021**, *3*, 1841–1861. [CrossRef]
5. Deng, T.; Hu, B.F.; He, Y.Z.; Yin, Y.L. Simulation, Experimental Testing and Optimization of Starting and Shifting Control Strategies of DCT Wet Dual Clutches with Respect to Sliding Friction. *Iran. J. Sci. Technol. Trans. Mech. Eng.* **2019**, *43*, 693–705. [CrossRef]
6. Li, X.; Yan, H.; Wei, D.; Wu, J.; Yin, M.; Xiao, Y.; Li, K. Study on the Effect of Operating Conditions on the Friction Pair Gap in a Wet Multi-Disc Clutch in a Helicopter Transmission System. *Appl. Sci.* **2024**, *14*, 7751. [CrossRef]
7. Gao, A.; Fu, Z.; Tao, F. An Adaptive Fuzzy Coordinated Control Strategy for Hybrid Electric Vehicles Considering Clutch Wear and Engine Temperature Variation. *Energy Sci. Eng.* **2024**, *12*, 3631–3646. [CrossRef]
8. Singh, M. D.; Prasanna, S.; Baskar, P. Comprehensive Review on Dual Clutch Transmission. *Int. J. Veh. Perform.* **2022**, *8*, 74–108. [CrossRef]
9. Cerone, V.; Canale, M.; Regruto, D.; Tomalino, T. Actuator Position Control of Dual Clutch Transmission Systems Using Model Predictive Control Techniques. Ph.D. Thesis, Politecnico di Torino, Torino, Italy, 2018. Available online: http://webthesis.biblio.polito.it/id/eprint/9556 (accessed on 13 February 2025).

10. Škugor, B.; Deur, J.; Ivanović, V. E-Clutch Torque Control Including Compensation of Thermal Expansion Effects. *IEEE Trans. Veh. Technol.* **2020**, *69*, 246–257. [CrossRef]
11. Pisaturo, M.; Senatore, A.Thermal Compensation Control Strategy in Automated Dry Clutch Engagement Dynamics and Launch Manoeuvre. *Int. J. Automot. Technol.* **2019**, *20*, 1089–1101. [CrossRef]
12. Oh, J.; Choi, S. Real-time Estimation of Transmitted Torque on Each Clutch for Ground Vehicles with Dual Clutch Transmission. *IEEE/ASME Trans. Mechatron.* **2015**, *20*, 24–36. [CrossRef]
13. Zhang, L.; Wei, C.; Hu, J. Model for the Prediction of Drag Torque Characteristics in Wet Clutch with Radial Grooves. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2019**, *223*, 3043–3056. [CrossRef]
14. Yan, Z.; Wu, B.; Gao, X.; Bai, X. Engagement/Disengagement Characteristics of Pull-Type Diaphragm Spring Clutch for Heavy-Duty Commercial Vehicles. *Chin. J. Mech. Eng.* **2024**, *37*, 533–543. [CrossRef]
15. Zhao, Z.; Li, X.; He, L.; Wu, C.; Hedrick, J.K. Estimation of Torques Transmitted by Twin-Clutch of Dry Dual-Clutch Transmission During Vehicle's Launching Process. *IEEE Trans. Veh. Technol.* **2017**, *66*, 4727–4741. [CrossRef]
16. Wu, P.; Qiang, P.; Pan, T.; Zang, H. Double-Loop Control for Torque Tracking of Dry Clutch. *Machines* **2022**, *10*, 1142. [CrossRef]
17. Vu, T.M.; Moezzi, R.; Cyrus, J.; Hlava, J.; Petru, M. Automatic Clutch Engagement Control for Parallel Hybrid Electric Vehicle. *Energies* **2021**, *14*, 7256. [CrossRef]
18. Park, J.; Choi, S. Optimization Method of Reference Slip Speed in Clutch Slip Engagement in Vehicle Powertrain. *Int. J. Automot. Technol.* **2021**, *22*, 55–67. [CrossRef]
19. Zhu, C.; Chen, Z.; Shi, Z.; Zhang, Y. Study on the Engagement Characteristics and Control Strategy of High Speed Difference Dry Friction Clutch. *Machines* **2023**, *11*, 407. [CrossRef]
20. Geng, X.; Liu, W.; Liu, X.; Wen, G.; Xue, M.; Wang, J. Optimal Torque Control of the Launching Process with AMT Clutch for Heavy-Duty Vehicles. *Machines* **2024**, *12*, 363. [CrossRef]
21. Li, L.; Zhu, Z.; Chen, Y.; He, K.; Li, X.; Wang, X. Engagement Control of Automated Clutch for Vehicle Launching Considering the Instantaneous Changes of Driver's Intention. *J. Dyn. Syst. Meas. Control* **2017**, *139*, 021011. [CrossRef]
22. Zhang, J.; Liu, X.; Wei, H.; Liu, M.; Huang, W.; Yan, X. Study on Shifting Performance of Tractor Multi-Clutch Under Different Engagement Rules. *Agriculture* **2024**, *14*, 254. [CrossRef]
23. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous Control with Deep Reinforcement Learning. *arXiv* **2015**, arXiv:1509.02971. [CrossRef]
24. Romero, C.A.; Correa, P.; Ariza Echeverri, E.A.; Vergara, D. Strategies for Reducing Automobile Fuel Consumption. *Appl. Sci.* **2024**, *14*, 910. [CrossRef]
25. Ibrahim, S.; Mostafa, M.; Jnadi, A.; Salloum, H.; Osinenko, P. Comprehensive Overview of Reward Engineering and Shaping in Advancing Reinforcement Learning Applications. *IEEE Access* **2024**, *12*, 175473–175500. [CrossRef]
26. He, J. Overall Evaluation of Drive-Off Procedures in a Mild Hybrid Powertrain. *Prim. Publ.* **2024**, *16*, 149. [CrossRef]
27. Shin, S.; Choi, S. B. Control-Oriented Modeling of Wet Clutch Friction Considering Thermal Dynamics. *Mechatronics* **2024**, *99*, 103146. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.