Journal of Applied Fluid Mechanics, Vol. 17, No. 4, pp. 816-827, 2024.
Available online at www.jafmonline.net, ISSN 1735-3572, EISSN 1735-3645.
https://doi.org/10.47176/jafm.17.4.2248

Closed-loop Flow Control Method Based on Deep Reinforcement Learning using a Co-flow Jet

Y. R. Zhao, H. Y. Xu† and Z. Y. Xie

National Key Laboratory of Science and Technology on Aerodynamic Design and Research, Northwestern Polytechnical University, Xi'an, 710072, China

†*Corresponding Author Email: xuheyong@nwpu.edu.cn*

ABSTRACT

A closed-loop control framework is developed for the co-flow jet (CFJ) airfoil by combining the numerical flow field environment of a CFJ0012 airfoil with a deep reinforcement learning (DRL) module called tensorforce integrated in Python. The DRL agent, which is trained through interacting with the numerical flow field environment, is capable of acquiring a policy that instructs the mass flow rate of the CFJ to make the stalled airfoil at an angle of attack (AoA) of 18 degrees reach a specific high lift coefficient set to 2.0, thereby effectively suppressing flow separation on the upper surface of the airfoil. The subsequent test shows that the policy can be implemented to find a precise jet momentum coefficient of 0.049 to make the lift coefficient of the CFJ0012 airfoil reach 2.01 with a negligible error of 0.5%. Moreover, to evaluate the generalization ability of the policy trained at an AoA of 18 degrees, two additional tests are conducted at AoAs of 16 and 20 degrees. The results show that, although using the policy gained under another AoA cannot help the lift coefficient of the airfoil reach a set target of 2 accurately, the errors are acceptable with less than 5.5%, which means the policy trained under an AoA of 18 degrees can also be applied to other AoAs to some extent. This work is helpful for the practical application of CFJ technology, as the closed-loop control framework ensures good aerodynamic performance of the CFJ airfoil, even in complex and changeable flight conditions.

Article History

Received August 17, 2023
Revised October 29, 2023
Accepted November 18, 2023
Available online January 30, 2024

Keywords:

Co-flow jet
Closed-loop control
Flow control
Lift enhancement
Deep reinforcement learning

1. INTRODUCTION

The co-flow jet is an active flow control (AFC) method that was first developed by Zha and Paxton (2004). The method is to create an injection slot near the leading edge and a suction slot near the trailing edge to form a jet channel on the upper surface of the airfoil, as shown in Fig. 1.

A high-energy jet is injected through the injection slot to the flow field along the surface of the airfoil, and the same amount of mass flow is withdrawn into the suction slot. With the injection of high-energy jets into the external flow field, the energy of the boundary layer of the suction surface surges; that is, strong jet mixing and energy exchange occur with the main flow to ensure that the energy of the boundary layer is sufficient to overcome the inverse pressure gradient at high AoAs, thereby inhibiting flow separation. The streamline of the baseline airfoil and CFJ airfoil for NACA2415 is shown in Fig. 2.

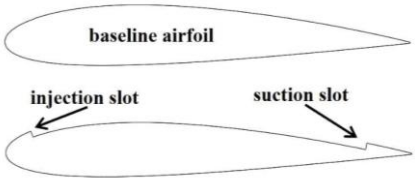


Fig. 1 Baseline airfoil and CFJ airfoil for NACA2415

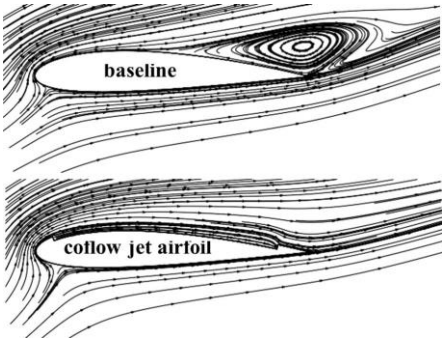


Fig. 2 Streamline of the baseline airfoil and CFJ airfoil for NACA2415

NOMENCLATURE

a	actions	S	reference area
C_d	drag coefficient	t	discrete time step
C_l	lift coefficient	T	final time step of a trajectory
C_p	pressure coefficient	V_j	velocity of the jet
C_μ	jet momentum coefficient	V_∞	velocity of the freestream
D_j^2	variance of the pressure per surface grid point	x_{ij}	pressure value of the j^{th} grid point under the i^{th} flight states
H	height	$\overline{x_j}$	average pressure value of the j^{th} grid point under 8400 flight states
i	serial number of the flight states	α	learning rate
j	serial number of the grid point on the surface of the airfoil	β	parameter for adjusting variation smoothness of the mass flow rate
J	expected return	γ	discount factor
l	probability ratio of the new policy and the old policy	ε	hyperparameter which limits l to a certain range
L	loss function	θ	parameter of the policy
$\dot{m}$	mass flow rate of the jet	μ	average value of a Gaussian distribution
Ma	Mach number	π	policy
r	rewards	ρ_∞	density of the freestream
R	discounted return	σ	variance of a Gaussian distribution
s	states	τ	trajectory, which is a sequence of states, actions and rewards

Since the concept of the CFJ was proposed, researchers have developed the technique from the mechanism study of 2D CFJ airfoils (Zha et al., 2006b) to the parametric study of CFJ airfoils with different shapes and various freestream conditions (Wang et al., 2008; Lefebvre & Zha, 2014; Ma & Xu, 2022; Ma et al., 2023). Using different study methods, such as numerical simulation (Wang & Zha, 2011; Gan et al., 2013), wind tunnel tests (Zha et al., 2006a; Zha et al., 2007a; Zha et al., 2018) and application research (Yang & Zha, 2019; Lei & Zha, 2021; Xu & Zha, 2021), CFJ technology has been demonstrated to have potential in lift enhancement and drag reduction with low energy consumption. However, despite years of development, CFJ technology is still an open-loop control method, which implies that the control strategy is fixed and nonadjustable. This approach is not sufficient to effectively address the complex and volatile nature of practical flight situations. Therefore, developing closed-loop methods for CFJ technology is significant.

There are many closed-loop control methods that can be selected for research. In general, closed-loop control methods can be divided into two categories: traditional PID control and model predictive control (MPC) (Franklin et al., 2019). Although PID control is widely used in engineering for its mature and simple technology, adjusting the PID parameters is still challenging, as it requires prior experience and multiple trials to acquire satisfactory parameters. Additionally, the control effect of PID is not ideal for high-dimensional and nonlinear systems. MPC approaches are better suited to solve nonlinear and high-dimensional control problems (Siegel et al., 2003; Glauser et al., 2004; Ahuja & Rowley, 2010), but they are comparatively difficult to implement. This is because researchers need to observe the entire system to derive the reduced-order model (ROM), which requires large amounts of data from experiments or numerical simulations and can restrict the feasibility of the method.

In the field of artificial intelligence (AI), a method called reinforcement learning (RL) is capable of finding control strategies automatically by interacting with the environment that needs to be controlled. The method has been used for solving low-dimensional problems since the 1980s (Sutton, 1988). Since the performance of computers has been greatly enhanced, deep learning has become a popular machine learning method using artificial neural networks (ANNs), which have worked well in perceptual problems such as speech and image recognition since 2012 (Krizheysky et al., 2017). Therefore, deep reinforcement learning (DRL) was devised by combining deep learning and RL to utilize the great ability of ANNs in data representation (Akbiyik & Yavuz, 2021; Moshtaghzadeh & Aligoodarz, 2022). The DRL method has been widely applied in different domains, including games, such as Atari games (Mnih et al., 2013) and AlphaGo (Silver et al., 2016), industrial applications, such as unmanned vehicles (Kiran et al., 2022), and the field of robot control (Gu et al., 2017).

Recently, researchers in the field of flow control have shown interest in utilizing DRL technology in AFC systems due to its ability to handle problems with nonlinearity and high dimensionality. Two typical objects were mainly discussed: 2D flow past circular cylinders and flow past airfoils. By using various DRL algorithms, such as proximal policy optimization (PPO) and deep Q-network (DQN), control strategies for different purposes, i.e., drag reduction (Rabault et al., 2019; Rabault & Kuhnle, 2019; Tang et al., 2020; Ren et al., 2021a; Paris et al., 2021) and hydrodynamic stealth (Ren et al., 2021b) of a circular cylinder and flow separation control of an airfoil (Shimomura et al., 2020), were successfully developed. Rabault et al. (2019) used the PPO algorithm to search for the best mass flow rate of jets implemented symmetrically on the upper and lower surfaces of a circular cylinder to reduce drag. This was the first time that DRL was used to perform AFC in a simple computational

fluid dynamics (CFD) simulation, where the Reynolds number was relatively low. To speed up obtaining control strategies, Rabault and Kuhnle (2019) were inspired by parallel numerical simulation and developed an advanced version of the method discussed before. They achieved this by adapting the DRL algorithm for parallelization, which involved running several independent simulations in parallel to gather experiences more rapidly. To further improve this work, several research studies have been conducted to enhance the robustness of DRL methods (Tang et al., 2020; Ren et al., 2021a) and optimize the sensor layout (Paris et al., 2021). For the purpose of hiding traces of bluff bodies, Ren et al. (2021b) used the DRL algorithm to find the control policy of windward-suction-leeward-blowing actuators implemented on a bluff body. As a result, compared to the uncontrolled bluff body, the velocity deficit in the wake of the controlled bluff body was significantly reduced. To control the flow past 2D airfoils, Shimomura et al. (2020) carried out the first trial. They used a dielectric barrier discharge plasma actuator (DBDPA) installed on the upper surface of an NACA0015 airfoil to prevent flow separation. By applying the DQN algorithm, the authors searched for the optimal frequency of the actuator at AoAs of 12° and 15°. The effectiveness of this strategy was experimentally investigated in a low-speed wind tunnel at a chord Reynolds number of 6.3×10^4. This study further confirmed the potential of DRL and paved the way for applying DRL-learned policies to other AFC-controlled airfoils. In addition to the field of flow control, Lou et al. (2023) applied DRL to aerodynamic optimization. Using a deep Q-network (double DQN), the authors realized the goal of finding the best airfoil geometry to attain a high lift-drag ratio, and this trial also confirmed the power of DRL in the field of applied fluid mechanics.

Therefore, the aim of this research is to combine the power of the DRL algorithm with CFJ to establish a closed-loop control framework for a stalled airfoil experiencing significant flow separation. The main objective of this control framework is to enhance the lift coefficient of the airfoil when it is in a stalled condition, thereby mitigating flow separation. To accomplish this goal, ten strategically placed pressure sensors are utilized to provide the control framework with crucial information about the flow field. The DRL agent interacts with the environment and learns its policy through these sensor data. A designed reward function is used as feedback to guide the agent's learning process, which allows the agent to determine the appropriate mass flow rate for the CFJ, ultimately achieving the desired lift coefficient. This research is the first endeavor in utilizing a closed-loop control method for CFJ airfoil systems. This trial holds potential in using CFJ technology in complex, changeable flying situations, making it more feasible for practical engineering applications of CFJ technology.

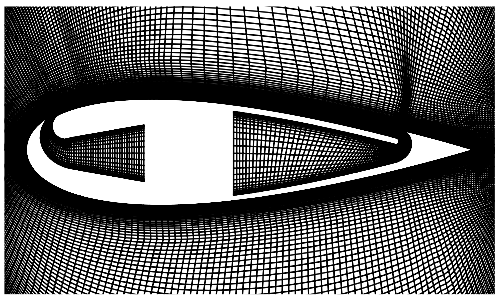


Fig. 3 2-D mesh of the CFJ0025 airfoil for CFD calculation

2. METHODOLOGY

2.1 Flow Solver and 2-D Mesh of CFJ0012 Airfoil for CFD Calculation

In this paper, numerical simulations of the flow field are conducted using Ansys Fluent software, which is based on the finite volume method. The Reynolds-averaged Navier–Stokes (RANS) equations, along with the Spalart-Allmaras (SA) turbulence model, are employed to validate the numerical method for the CFJ0025 airfoil. The CFJ0025 airfoil is a modified version of the NACA0025 airfoil, specifically designed to incorporate a CFJ.

The geometry model and computational conditions are based on the referenced paper (Zha et al., 2007a). The freestream Mach number is set to 0.11, and the Reynolds number, calculated based on the chord length, is 3.8×10^5. A two-dimensional mesh, as displayed in Fig. 3, is employed for the CFD calculation.

Figure 4 illustrates a comparison of the present computed value, the experimental value (Zha et al., 2007a) and the computed value calculated by Zha et al. (2007b) at various AoAs. In Fig. 4 (a), the present calculated lift coefficient matches well with the experimental result

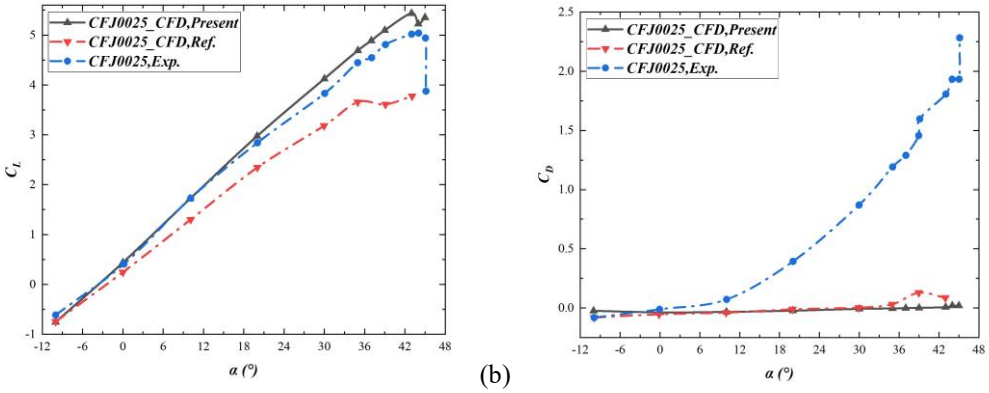
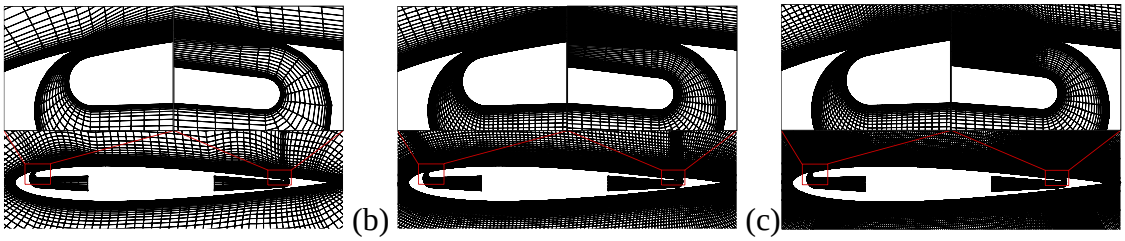


Fig. 4 Computed results compared with the experimental results at different angles of attack: (a) lift coefficient and (b) drag coefficient

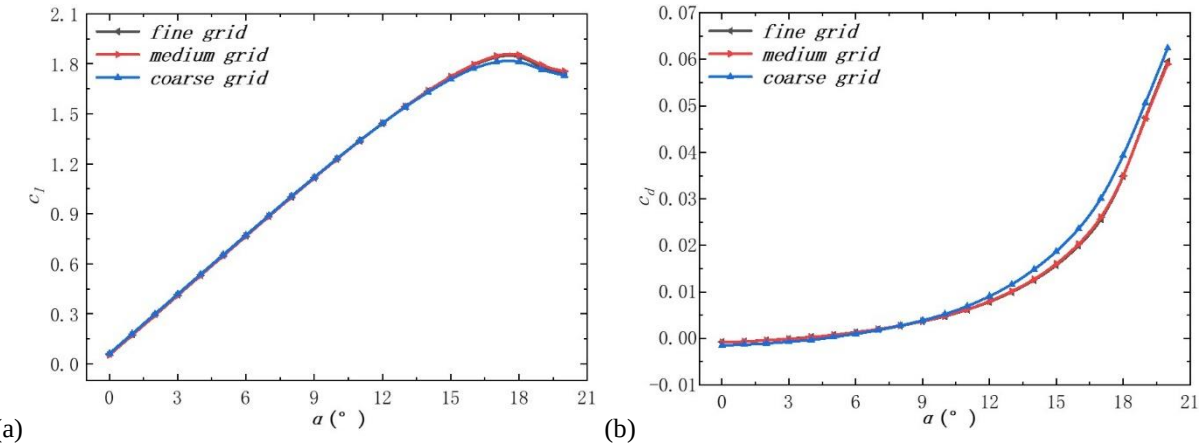
Table 1 Detailed information of three grids at different densities

Densities	Parameters				
	Wrap-around points	Normal layers	Growth rate	Size of the slot mesh	Size of the external flow field
Coarse	154	81	1.20	99×30	154×80
Medium	315	151	1.15	230×50	315×150
Fine	649	191	1.10	297×70	645×190



(a) (b) (c)

Fig. 5 2-D mesh for the CFJ0012 airfoil at three different densities: (a) coarse grid, (b) medium grid, and (c) fine grid



(a) (b)

Fig. 6 Computed result for the CFJ0012 airfoil in three different grid densities: (a) lift coefficient and (b) drag coefficient

when compared to the referenced computed values. This indicates that the present calculation method accurately predicts the lift coefficient of the airfoil, providing satisfactory agreement with the experimental data. However, there is a significant discrepancy between the experimental and computed values of the drag coefficient, as shown in Fig. 4 (b). This discrepancy could be attributed to the probable experimental measurement error or inadequate turbulence simulation by the RANS model (Zha et al., 2007b). In the present research, the lift coefficient of the airfoil is specifically focused on, and hence, the calculation method is sufficient for the study.

In this study, the CFJ0012 airfoil is selected as the subject of investigation, and the chord length of the airfoil is 1 m. The injection and suction slots are located at 7% and 85% of the chord from the leading edge, and the widths of the injection and suction slots are 0.45%c and 0.9%c, respectively. To ensure grid independence and evaluate grid sensitivity, three different density grids are created, as depicted in Fig. 5. The grids are set in an O-type structure. The total number of grids is 15290, 57400 and 144100, respectively. The far-field is set to 50c for all three grids. To ensure $y^+<1$, the heights of the first layer are 1.37×10^{-5}, 7×10^{-6}, and 1×10^{-6}. Other detailed information on the grids is listed in Table 1. The freestream Mach number is 0.2, and the Reynolds number based on the chord length is 4.68×10^6. The jet momentum coefficient is 0.04. The computed lift coefficient and drag coefficient for these three grid sizes are presented in Fig. 6. The results indicate that the computed values obtained from the medium and fine grids exhibit good agreement. However, it is observed that the computed values from the coarse grid differ significantly from those of the medium and fine grids. Considering both computational efficiency and accuracy, the medium grid is chosen for subsequent studies.

2.2 DRL-Based Closed-Loop Control Framework Using the PPO Algorithm Applied to the CFJ0012 Airfoil

The training process of DRL involves learning and optimizing the policy of the DRL agent by continually interacting with the environment. This training process can be conceptualized as an episode, denoted as τ in Eq. (1), which records how the agent interacts with and receives feedback from the environment.

$$\tau = (s_0, a_0, r_0, s_1, a_1, r_1, ..., s_T, a_T, r_T) \quad (1)$$

where s represents the state of the environment; a represents the action given by the agent; and r represents

the reward value calculated through a defined reward function.

At the start of an episode, the agent observes the initial state of the environment. Based on this information, the untrained agent will choose an action randomly from its action set. The chosen action is then implemented in the environment, causing it to transition to a new state. The environment provides the agent with a reward value, which serves as feedback on the quality of the action taken. The agent then updates its policy using this information, with the goal of finding a policy that maximizes the discounted return defined in Eq. (2).

$$R(\tau)=\sum_{t=0}^{T}\gamma^{t}r_{t} \tag{2}$$

where γ is a discount factor ranging from 0 to 1, which is used to decrease the importance of future rewards.

Various algorithms have been proposed in the development history of DRL, including REINFORCE, TRPO (Trust Region Policy Optimization) (Schulman et al., 2015), PPO, etc. Of these algorithms, the PPO algorithm, proposed by Schulman et al. (2017), has been widely recognized for its remarkable efficiency and robustness in practical applications. Therefore, a PPO agent is utilized in this paper as a part of the closed-loop control system.

The detailed framework based on the PPO algorithm is shown in Fig. 7, and the training parameters of PPO are listed in Table 2. The PPO agent used in this paper includes an actor network, whose structure is automatically defined by tensorforce according to the input types and shapes. The PPO agent will optimize its policy by interacting with the environment of a numerical flow field of the controlled object: the CFJ0012 airfoil with a very weak C_μ of 0.0001 at an angle of attack of 18°. The freestream Mach number is 0.2, and the Reynolds number based on the chord is 4.68×10^6. At the beginning of the training process, the actor network takes the initial state "s" as input and then outputs the average value μ and the variance σ to form a Gaussian distribution. From this distribution, an action value "a" within a set range will be randomly chosen. After implementing a jet with a mass flow rate of "a", the flow field transitions to the next state "s'". Simultaneously, a reward value "r" is computed, which provides immediate feedback for the agent's actions. To improve the agent's performance, a dataset consisting of the state-action-reward-next state tuple (s, a, r, s') is stored in an experience replay buffer. This buffer allows the agent to learn from past experiences by sampling mini batches of transitions. During training, the parameter of the actor network θ is updated by minimizing the loss function shown in Eq. (3).

$$L_{actor}=\hat{\mathrm{E}}_t[\min(l_t(\theta)R_t, clip(l_t(\theta),1-\varepsilon,1+\varepsilon)R_t)] \tag{3}$$

where $l_t(\theta)=\pi_{\theta_{new}}(a_t\mid s_t)/\pi_{\theta_{old}}(a_t\mid s_t)$ is the probability ratio of the new policy and the old policy and ε is a hyperparameter with a value of 0.25, which limits the value of $l_t(\theta)$ to the range of $[1-\varepsilon,1+\varepsilon]$.

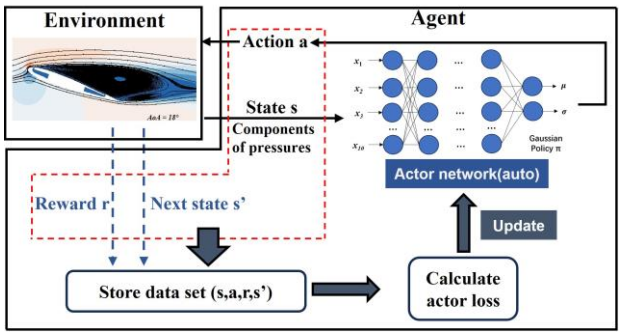


Fig. 7 Closed-loop control framework built with the PPO algorithm

Table 2 Training parameters for PPO

Parameters	Value
Experience memory size	50000
Batch size	22
Update frequency	2
Learning rate	0.003
Discount factor	0.97
Exploration	0.1
Episode timesteps	40

The above introduction to the DRL framework shows that it is crucial to carefully design the state space, the action space and the reward function. Hence, the design of these three parts is given as follows.

A. Determination of the State Space

The state space of this framework involves the pressure data collected from ten carefully selected points located on the surface of the airfoil. The choice of these specific points is determined by two crucial criteria (Samy et al., 2010):

a. Steep Pressure Gradient

This criterion focuses on identifying points with a significant and noticeable change in the pressure distribution along the airfoil's surface. These points are selected based on the presence of a steep pressure gradient, indicating a rapid transition in the pressure distribution.

b. Substantial Variance in the Pressure Values

The second criterion involves the identification of points that have considerable variation in the pressure values as the flight state changes. These points reflect sensitivity to alterations in the parameters such as the height *H*, Mach number *Ma*, and angle of attack (*AoA*).

To identify regions that meet the first criterion, the C_p distribution on the CFJ0012 airfoil is calculated. Figs. 8 (a)-(d) show the C_p distribution on the CFJ0012 airfoil under various flight conditions. Each figure within the set illustrates the C_p curves with the variation of a single parameter. Upon observing Figs. 8 (a)-(d), all figures consistently reveal that regions near the leading edge and the trailing edge, as well as the locations of the injection and suction slots, exhibit noteworthy changes in the pressure distribution along the surface. These areas display steep pressure gradients, indicating rapid transitions in the pressure values along the surface of the

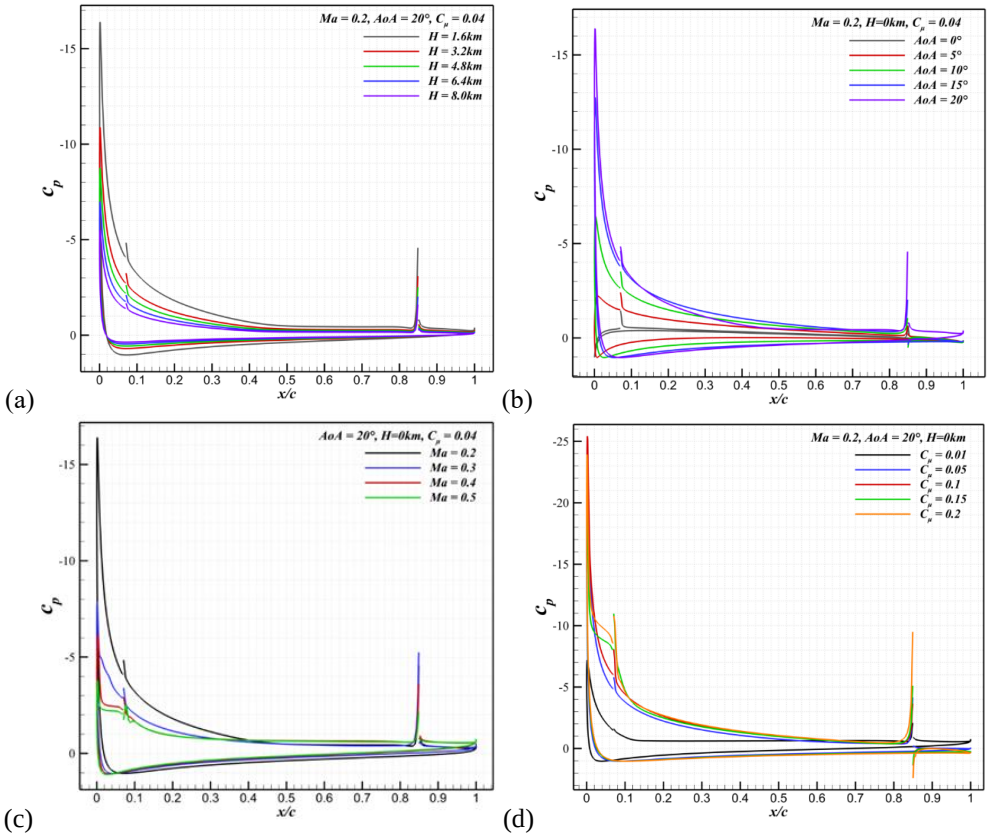


Fig. 8 C_p distribution of the CFJ0012 airfoil under different flight conditions with only one flight parameter change: (a) C_p curves with only H changes, (b) C_p curves with only AoA changes, (c) C_p curves with only Ma changes, and (d) C_p curves with only C_μ changes

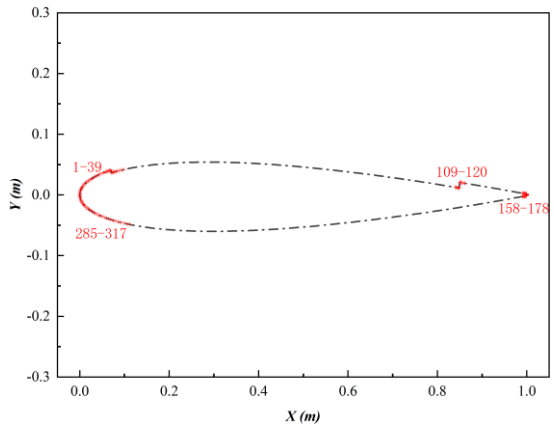


Fig. 9 Locations of the pressure probes where the pressure gradient is steep

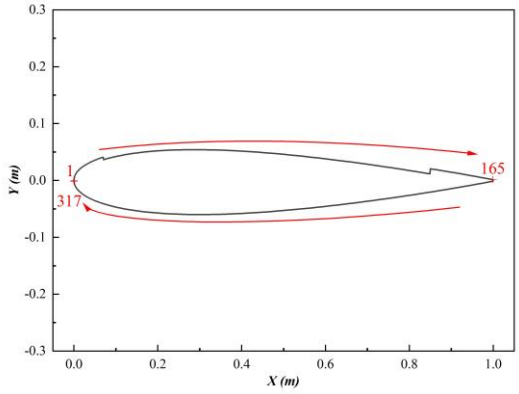


Fig. 10 Numbering of the airfoil surface grid nodes

airfoil. As a result, the range of pressure probes can be narrowed down to the red area on the airfoil, as shown in Fig. 9 (starting from the leading edge point, a total of 317 grid nodes of the airfoil surface are numbered clockwise, as shown in Fig. 10).

To further refine and determine the appropriate locations of the pressure probes, Criterion 2 should be considered. In the case of the CFJ airfoil, four key flight parameters are taken into account: the height H (ranging from 0 km to 8 km with an interval of 2 km), Mach number Ma (ranging from 0.2 to 0.5 with an interval of 0.1), jet momentum coefficient C_μ (ranging from 0.01 to 0.2 with an interval of 0.01), and angle of attack AoA (ranging from 0° to 20° with an interval of 1°). Among these parameters, C_μ is a particular flight parameter for the CFJ airfoil. It is a dimensionless parameter that quantifies the intensity of the CFJ, and the definition equation is shown in Eq. (4).

$$C_\mu = \dot{m} V_j / (\frac{1}{2} \rho_\infty V_\infty^{\ 2} S) \tag{4}$$

where $\dot{m}$ is the mass flow rate of the jet; V_j is the velocity of the jet; ρ_∞ is the density of the freestream; V_∞ is the velocity of the freestream; and S is the reference area. Therefore, 8400 (5×4×20×21) groups of flight states are obtained through the method shown in Fig. 11.

Then, the variance of the pressure per point is calculated through Eq. (5).

$$D_j^{\ 2} = \frac{\sum_{i=1}^{8400} (x_{ij} - \overline{x_j})^2}{8400} \tag{5}$$

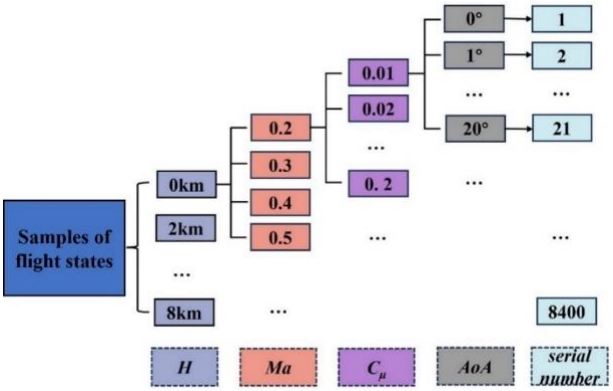


Fig. 11 8400 samples of flight states

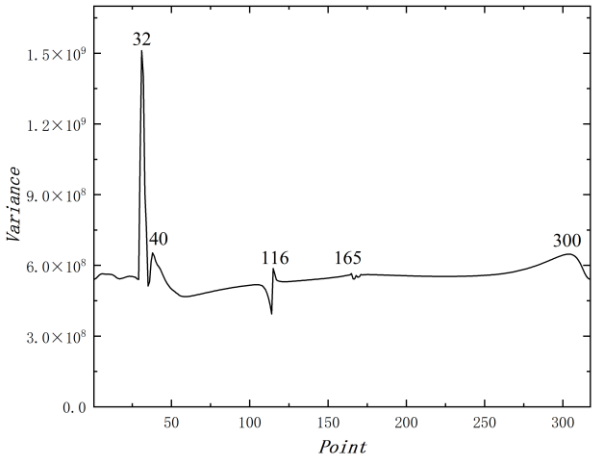


Fig. 12 Variances of the pressure values per point

where $i \in \{1,2,3,...,8400\}$ is the serial number of the flight states; $j \in \{1,2,3,...,317\}$ is the serial number of the grid point on the surface of the airfoil; x_{ij} denotes the pressure value of the j^{th} grid point under the i^{th} flight states; and $\overline{x_j}$ is the average pressure value of the j^{th} grid point under 8400 flight states.

Figure 12 displays the variances of the pressure values for each grid point. P32, P40, P116, P165 and P300 exhibit larger variances compared to the other points. These specific locations are indicated in Fig. 13. Interestingly, these points fall within the red area of the airfoil, as depicted in Fig. 9.

In the final selection of the pressure probes for the DRL framework, a set of pressure values of 10 pressure probes (Fig. 14) is chosen as members of the state space, and the coordinates of the pressure probes are listed in Table 3. Except for P57 and P104, all other points in the set satisfy at least one of the previously mentioned criteria. P57 and P104 are intentionally chosen because they are located within the jet slot. These points have the potential to provide information regarding the intensity of the jet.

B. Determination of the Action Space

In this framework, the range of the action space is exactly the range of the jet momentum coefficient. To facilitate the efficiency of the agent in obtaining the desired policy within the DRL framework, a preliminary experiment is conducted to determine a rough range of the

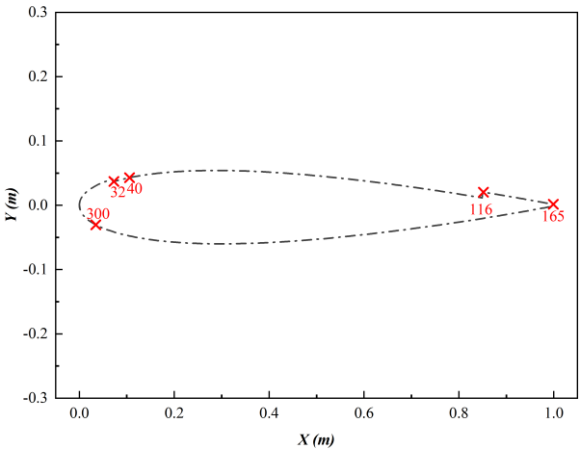


Fig. 13 Locations of the points where the pressure variance is large

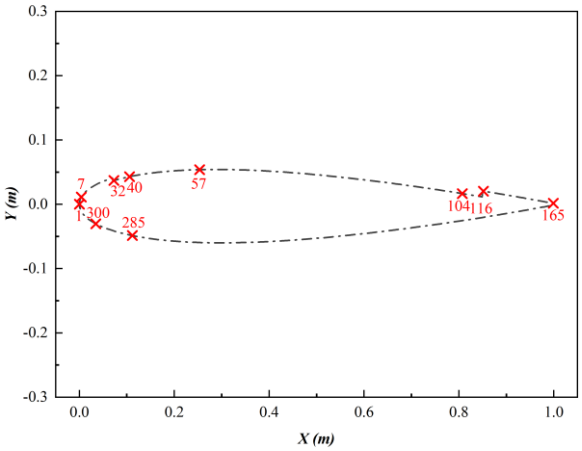


Fig. 14 Ten selected pressure points

Table 3 Coordinates of the ten selected pressure points

Probe number	Coordinate	
	X(m)	Y(m)
1	0	0
7	0.00416	0.01117
32	0.07293	0.0369
40	0.10598	0.04298
57	0.25351	0.05384
104	0.80718	0.01661
116	0.8516	0.02034
165	0.99892	0.0014
285	0.11169	-0.04869
300	0.03470	-0.03031

momentum coefficient. The experimental results, shown in Fig. 15, show that at an AoA of 18 degrees, the lift coefficient of the airfoil varies between 1.7 (C_μ= 0.04) and 2.3 (C_μ= 0.1). Therefore, in the present study, 2.0 is set as the control target, and the goal of the DRL training is to find a corresponding C_μ to help the CFJ airfoil reach the lift coefficient of 2.0. To simplify the numerical simulation process, the boundary conditions of the injection slot and the suction slot are modified to the mass flow inlet and mass flow outlet, respectively. Based on Eq. (4), the range of the mass flow rate is supposed to be [0.8

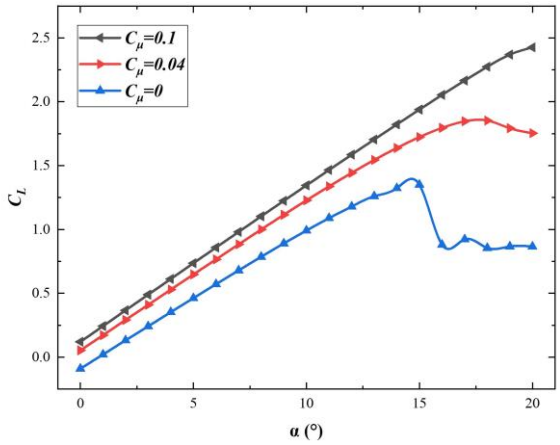


Fig. 15 Lift coefficient curves under different jet momentum coefficients

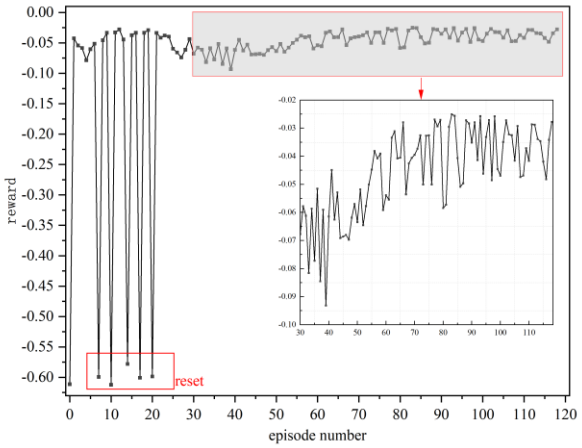


Fig. 16 Reward progression over the episodes at *AoA*=18°

kg/s, 1.2 kg/s]. However, to further display the decision-making ability of the DRL algorithm, the range is slightly broadened to [0.6 kg/s, 1.2 kg/s].

C. Determination of the reward function

The reward value should increase as the lift coefficient of the airfoil approaches the target value of 2.0. Therefore, the reward function of this case is designed and shown in Eq. (6).

$$r_t = -|C_l - 2| \tag{6}$$

3. RESULTS AND DISCUSSIONS

The whole training process starts from an initial flow field where C_μ is very small ($C_\mu = 0.0001$), i.e., the initial state. To obtain the instant lift coefficient according to the change in C_μ, a transient solver is used in this section. One important indicator for assessing the performance of the DRL training process is the variations in the reward values over the episodes. Here, one episode includes 40 actions. The mass flow rate of the jet changes gradually to the action value given by the agent in 25 numerical time steps according to Eq. (7) to ensure that the CFD solver will not crash.

$$\dot{m}_t = \dot{m}_{t-1} + \beta(a - \dot{m}_{t-1}) \tag{7}$$

where $\dot{m}_t$ is the current mass flow rate of the CFJ; $\dot{m}_{t-1}$ is the mass flow rate at the last numerical step; a is the action value given by the agent; and $\beta = 1/3$ is a parameter for adjusting the variation smoothness of the mass flow rate.

The reward curve over the episodes is depicted in Fig. 16. There are five points where the reward is extremely low, which can be attributed to the "reset" process. The "reset" process is that the beginning state of an episode will be reset to the initial state mentioned before, and its possibility is less than 20%. Otherwise, a new episode will start from the final state of the last episode. There is a noticeable increase in the reward value after the 40th episode, indicating that the agent begins to discover a favorable policy. However, the growth rate becomes

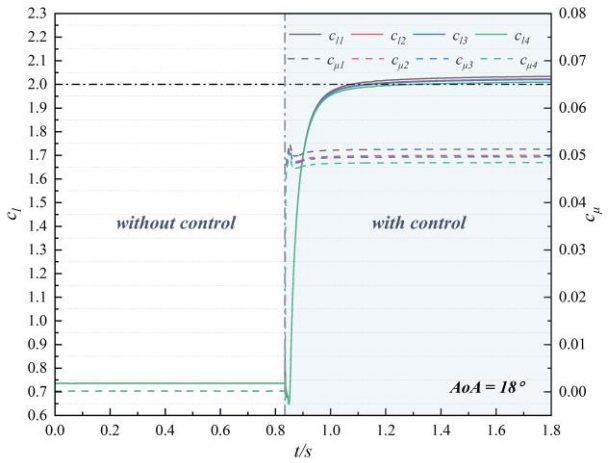
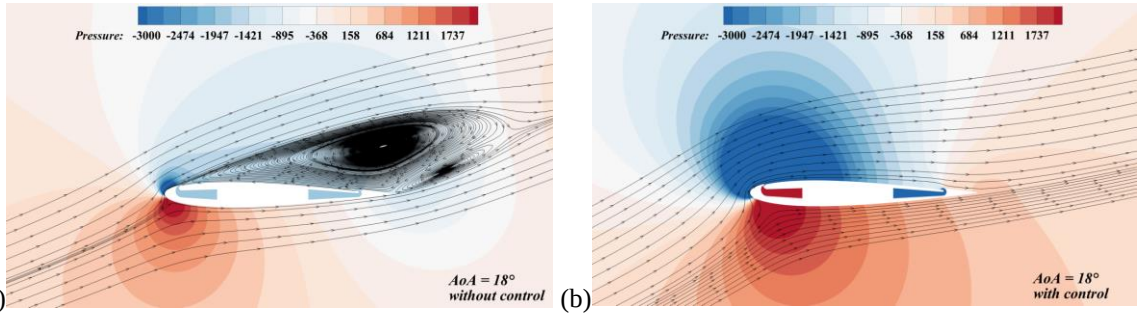


Fig. 17 Policy test at *AoA*=18°

stable after the 90th episode, suggesting that the agent's policy variation becomes less pronounced.

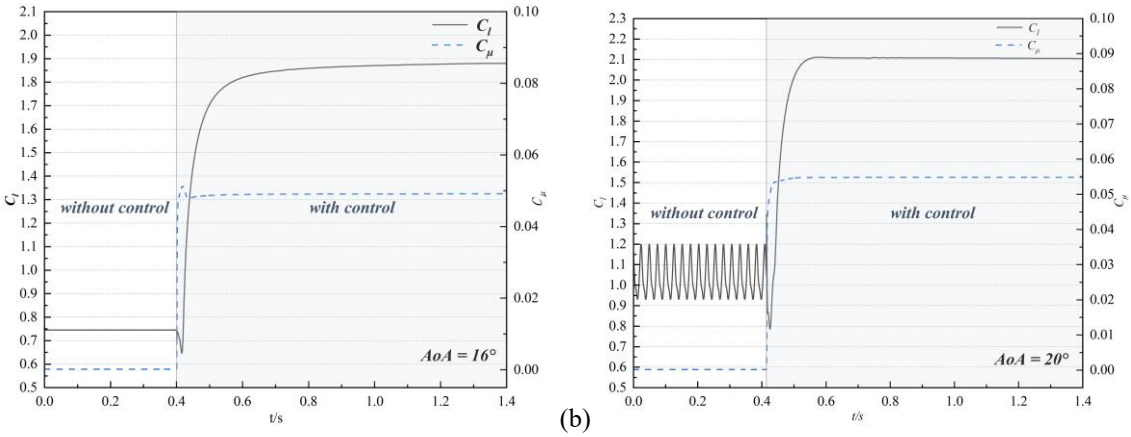
Four tests are carried out after Episodes 60, 80, 100 and 120, as shown in Fig. 17, and each test lasts for one second. In the "without control" part of Fig. 17, the CFJ0012 airfoil with a very weak C_μ of only 0.0001 suffers a very low lift coefficient of approximately 0.74. However, when the CFJ0012 airfoil is controlled according to the agent's policy, its lift coefficient experiences a significant increase and eventually approaches the target value of 2.0. Among the four tests, the policy acquired after Episode 120 exhibits the best performance. The lift coefficient becomes stable at approximately 2.01, with a negligible error of only 0.5%. In addition, the policy has physical meaning: when the lift coefficient is much less than the target value 2, the agent tends to output large action values to help the airfoil rapidly increase its lift coefficient, whereas when the lift coefficient approaches 2.0, the jet momentum coefficient gradually decreases and settles at approximately 0.049.

Figure 18 provides a visual representation of the pressure contours at an AoA of 18° for two scenarios: flow fields with a weak jet (C_μ=0.0001) and a much stronger jet instructed by the agent. In Fig. 18 (a), it is clearly observed



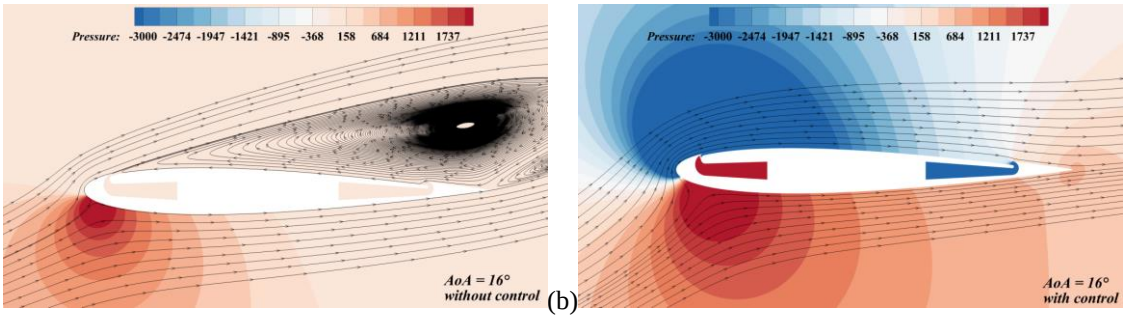
(a) (b)

Fig. 18 Comparison of the pressure contours with and without control at *AoA*=18°: (a) without control, and (b) with control



(a) (b)

Fig. 19 Generalization ability test of policy: (a) *AoA*=16°, and (b) *AoA*=20°



(a) (b)

Fig. 20 Comparison of the pressure contours with and without control at AoA=16°: (a) the flow field without control, and (b) the flow field with control

that flow separation takes place on the upper surface of the airfoil. This separation creates an adverse pressure gradient and reduces the overall lift generated by the airfoil. However, in Fig. 18 (b), when the CFJ is activated according to the agent's policy, the flow separation is effectively suppressed. This is evidenced by the pressure contours and streamlines showing a smooth and attached flow along the upper surface of the airfoil. The successful suppression of flow separation and transition to attached flow, as depicted in Fig. 18 (b), leads to a significant enhancement in lift, as shown in the "with control" portion in Fig. 17.

To test the policy's generalization ability, two more tests are conducted: one test at *AoA*=16° and the other test at *AoA*=20°. The test results are shown in Figs. 19 (a) and (b), respectively. At *AoA*=16°, the lift coefficient with a very weak jet (C_μ=0.0001) is approximately 0.75. After implementing the jet based on the policy, the lift coefficient increases significantly and reaches 1.89 within one second. However, this value deviates from the target value of 2.0, resulting in an error of 5.5%. At *AoA*=20°, when C_μ is very weak, the lift coefficient (C_l) fluctuates between 0.94 and 1.2. When the agent instructs a value of C_μ=0.055, the C_l curve rises to 2.1 within one second, with an error of 5% compared to the target value of 2. Therefore, the policy trained at 18° can instruct the airfoils at 16° and 20° to realize the control target to some extent, which means that the policy has a generalization ability. Moreover, the agent will provide different action values according to different states, for instance, C_μ=0.049 for *AoA*=16° and C_μ=0.055 for *AoA*=20°.

The pressure contours corresponding to the tests are shown in Fig. 20 and Fig. 21. Fig. 20 (a) displays the pressure contour of the flow field at *AoA*=16° with a very weak jet (C_μ=0.0001). In this case, obvious flow separation occurs on the upper surface of the airfoil..

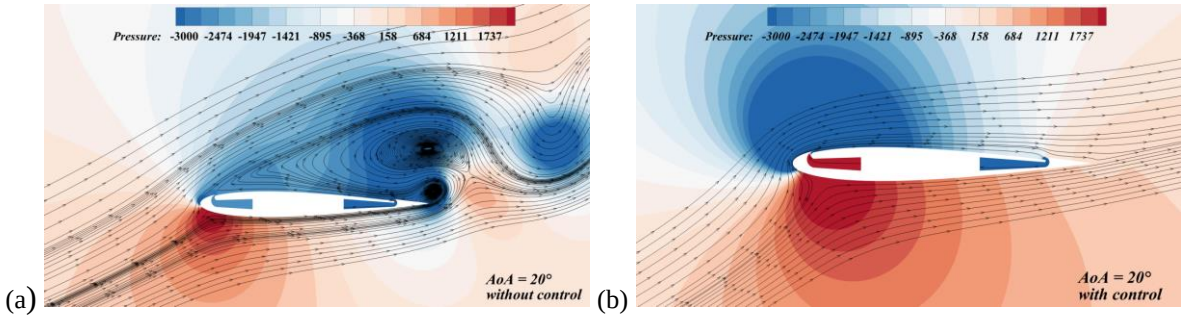


Fig. 21 Comparison of the pressure contours with and without control at AoA=20°: (a) the flow field without control, and (b) the flow field with control

However, when a stronger jet is applied according to the agent's policy, the flow separation completely disappears, as shown in Fig. 20 (b). At *AoA*=20°, vortex shedding occurs on the upper surface of the uncontrolled airfoil (Fig. 21 (a)). This vortex shedding leads to fluctuations in the lift coefficient, as observed in the "without control" portion in Fig. 19 (b). However, when the jet with C_μ=0.055 is implemented, this phenomenon is effectively suppressed (Fig. 21 (b)), resulting in no fluctuation in the lift coefficient

4. CONCLUSION

In this paper, a closed-loop control framework based on the PPO algorithm is developed for the CFJ airfoil. The framework exhibits accurate control performance, successfully achieving the settled control target of enhancing the lift coefficient to a specific value of 2.0, with an error of only 0.5%, for the stalled CFJ airfoil at an AoA of 18°. This achievement is made possible by the policy derived from the PPO agent, which interacts continuously with the environment through the state-action-reward-next state tuple. Additionally, generalization ability tests are conducted to evaluate the performance of the policy at different AoAs of 16° and 20°. The results indicate that the policy can be applied to other AoAs as well. By effectively suppressing flow separation and vortex shedding on the upper surface of the airfoil, the lift coefficient is significantly improved.

Future studies are planned to focus on further enhancing the robustness of this closed-loop control framework and increasing the control efficiency in terms of energy conservation. The study of the robustness of the framework includes investigating whether the framework can facilitate the attainment of different control targets and examining the applicability of the policy in various flight situations, such as different Mach numbers and Reynolds numbers. In terms of energy conservation, penalty terms such as the power consumption coefficient and drag coefficient could be added to the reward function to conduct further research.

ACKNOWLEDGMENTS

This work was supported by the National Natural Sci ence Foundation of China (Grant No. 11972306) and the 111 Project of China (B17037).

CONFLICT OF INTEREST

The authors have no conflicts to disclose.

AUTHORS CONTRIBUTION

Yiran Zhao: Conceptualization, Methodology, Investigation, Writing. Heyong Xu: Conceptualization, Supervision, Writing-Review. Zhengyao Xie: Methodology, Formal Analysis, Software.

REFERENCES

Ahuja, S., & Rowley, C. W. (2010). Feedback control of unstable steady states of flow past a flat plate using reduced-order estimators. *Journal of Fluid Mechanics*, *645*, 447–478. https://doi.org/10.1017/S0022112009992655

Akbiyik, H., & Yavuz, H. (2021). Artificial neural network application for aerodynamics of an airfoil equipped with plasma actuators. *Journal of Applied Fluid Mechanics*, *14*(4), 1165–1181. https://doi.org/10.47176/jafm.14.04.32133

Franklin, G. F., Powell, J. D., & Emami-Naeini, A. (2019). *Feedback Control of Dynamic Systems (8th ed.)*. Boston, MA: Pearson.

Gan, W. B., Zhou, Z., Xu, X. P., & Wang, R. (2013). Delayed detached-eddy simulation and application of a coflow jet airfoil at high angle of attack. *Applied Mechanics and Materials*, *444–445*, 270–276. https://doi.org/10.4028/www.scientific.net/AMM.444-445.270

Glauser, M., Higuchi, H., Ausseur, J., Pinier, J., & Carlson, H. (2004, June 28). *Feedback control of separated flows*. 2nd AIAA Flow Control Conference. 2nd AIAA Flow Control Conference, Portland, Oregon. https://doi.org/10.2514/6.2004-2521

Gu, S., Holly, E., Lillicrap, T., & Levine, S. (2017). *Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates.* 2017 IEEE International Conference on Robotics and Automation (ICRA), 3389–3396. https://doi.org/10.1109/ICRA.2017.7989385

Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Sallab, A. A. A., Yogamani, S., & Pérez, P. (2022). Deep reinforcement learning for autonomous driving: a survey. *IEEE Transactions on Intelligent Transportation Systems*, *23*(6), 4909–4926. https://doi.org/10.1109/TITS.2021.3054625

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://api.semanticscholar.org/CorpusID:195908774

Lefebvre, A. M., & Zha, G. (2014, June 16). *Co-flow jet airfoil trade study part I: Energy consumption and aerodynamic efficiency*. 32nd AIAA Applied Aerodynamics Conference, Atlanta, GA. https://doi.org/10.2514/6.2014-2682

Lei, Z., & Zha, G. (2021). *Lift enhancement of supersonic thin airfoil at low speed by co-flow jet active flow control.* AIAA AVIATION 2021 FORUM. https://doi.org/10.2514/6.2021-2591

Lou, J., Chen, R., Liu, J., Bao, Y., You, Y., & Chen, Z. (2023). Aerodynamic optimization of airfoil based on deep reinforcement learning. *Physics of Fluids*, *35*(3), 037128. https://doi.org/10.1063/5.0137002

Ma, C. Y., & Xu, H. Y. (2022). Parameter-based design and analysis of wind turbine airfoils with conformal slot co-flow jet. *Journal of Applied Fluid Mechanics*, *16*(2), 269–283. https://doi.org/10.47176/jafm.16.02.1318

Ma, C. Y., Xu, H. Y., & Qiao, C. L. (2023). Comparative study of two combined blowing and suction flow control methods on pitching airfoils. *Physics of Fluids*, *35*(3), 035120. https://doi.org/10.1063/5.0138962

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013, December 19). *Playing atari with deep reinforcement learning*. ArXiv.Org. https://arxiv.org/abs/1312.5602v1

Moshtaghzadeh, M., & Aligoodarz, M. R. (2022). Prediction of wind turbine airfoil performance using artificial neural network and CFD approaches. *International Journal of Engineering and Technology Innovation*, *12*(4), 275–287. https://doi.org/10.46604/ijeti.2022.9735

Paris, R., Beneddine, S., & Dandois, J. (2021). Robust flow control and optimal sensor placement using deep reinforcement learning. *Journal of Fluid Mechanics*, *913*, A25. https://doi.org/10.1017/jfm.2020.1170

Rabault, J., & Kuhnle, A. (2019). Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach. *Physics of Fluids*, *31*(9), 094105. https://doi.org/10.1063/1.5116415

Rabault, J., Kuchta, M., Jensen, A., Réglade, U., & Cerardi, N. (2019). Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *Journal of Fluid Mechanics*, *865*, 281–302. https://doi.org/10.1017/jfm.2019.62

Ren, F., Rabault, J., & Tang, H. (2021a). Applying deep reinforcement learning to active flow control in weakly turbulent conditions. *Physics of Fluids*, *33*(3), 037121. https://doi.org/10.1063/5.0037371

Ren, F., Wang, C., & Tang, H. (2021b). Bluff body uses deep-reinforcement-learning trained active flow control to achieve hydrodynamic stealth. *Physics of Fluids*, *33*(9), 093602. https://doi.org/10.1063/5.0060690

Samy, I., Postlethwaite, I., Gu, D. W., & Green, J. (2010). Neural-Network-Based flush air data sensing system demonstrated on a mini air vehicle. *Journal of Aircraft*, *47*(1), 18–31. https://doi.org/10.2514/1.44157

Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P. (2015). *Trust region policy optimization*. Proceedings of the 32nd International Conference on Machine Learning. https://proceedings.mlr.press/v37/schulman15.htm

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017, July 20). *Proximal Policy Optimization Algorithms.* ArXiv.Org. https://arxiv.org/abs/1707.06347v2

Shimomura, S., Sekimoto, S., Oyama, A., Fujii, K., & Nishida, H. (2020). Closed-Loop flow separation control using the deep q network over airfoil. *AIAA Journal*, *58*(10), 4260–4270. https://doi.org/10.2514/1.J059447

Siegel, S., Cohen, K., & McLaughlin, T. (2003, June 23). *Feedback control of a circular cylinder wake in experiment and simulation (Invited)*. 33rd AIAA Fluid Dynamics Conference and Exhibit, Orlando, Florida. https://doi.org/10.2514/6.2003-3569

Silver, D., Huang, A., Maddison, C. Guez, A., Sifre, L., Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Kalchbrenner, N., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature 529*, 484–489. https://doi.org/10.1038/nature16961

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, *3*(1), 9–44. https://doi.org/10.1007/BF00115009

Tang, H., Rabault, J., Kuhnle, A., Wang, Y., & Wang, T. (2020). Robust active flow control over a range of Reynolds numbers using an artificial neural network trained through deep reinforcement learning. *Physics of Fluids*, *32*(5), 053605. https://doi.org/10.1063/5.0006492

Wang, B., & Zha, G. C. (2011). Detached-Eddy simulation of a coflow jet airfoil at high angle of attack. *Journal of Aircraft*, *48*(5), 1495–1502. https://doi.org/10.2514/1.C000282

Wang, B., Haddoukessouni B., Levy, J., & Zha, G. C. (2008). Numerical investigations of injection-slot-size effect on the performance of coflow jet airfoils. *Journal of Aircraft*, *45*(6), 2084-2091. https://doi.org/10.2514/1.37441

Xu, K., & Zha, G. (2021). High control authority three-dimensional aircraft control surfaces using coflow jet. *Journal of Aircraft*, *58*(1), 72-84. https://doi.org/10.2514/1.C035727

Yang, Y., & Zha, G. (2019, January 7). *Conceptual design of the co-flow jet hybrid electric regional airplane*. AIAA Scitech 2019 Forum. AIAA Scitech 2019 Forum, San Diego, California. https://doi.org/10.2514/6.2019-1584

Zha, G. C., Paxton, C. D., Conley, C. A., Wells, A., & Carroll, B. F. (2006a). Effect of injection slot size on the performance of coflow jet airfoil. *Journal of aircraft*, *43*(4), 987-995. https://arc.aiaa.org/doi/abs/10.2514/1.16999

Zha, G., Gao W., & Paxton, C. (2006b, January 9). *Numerical simulation of co-flow jet airfoil flows*. 44th AIAA Aerospace Sciences Meeting and Exhibit. Reno, Nevada. https://doi.org/10.2514/6.2006-1060

Zha, G., Yang, Y., Ren, Y., & McBreen, B. (2018, June 25). *Super-lift and thrusting airfoil of coflow jet actuated by micro-compressors.* 2018 Flow Control Conference, Atlanta, Georgia. https://doi.org/10.2514/6.2018-3061

Zha, G. C., & Paxton, C. (2004, June 28). *A novel airfoil circulation augment flow control method using CFJ*. 2nd AIAA Flow Control Conference, Portland, Oregon. https://doi.org/10.2514/6.2004-2208

Zha, G. C., Carroll, B. F., Paxton, C. D., Conley, C. A., & Wells, A. (2007a). High-performance airfoil using coflow jet flow control. *AIAA Journal*, *45*(8), 2087–2090. https://doi.org/10.2514/1.20926

Zha, G. C., Gao, W., & Paxton, C. D. (2007b). Jet effects on coflow jet airfoil performance. *AIAA Journal*, *45*(6), 1222–1231. https://doi.org/10.2514/1.23995