

Deep-reinforcement-learning-based separation control in a two-dimensional airfoil

Garcia, Xavier; Miró, Arnau; Suárez, Pol; Alcántara-Ávila, Francisco; Rabault, Jean; Font, Bernat; Lehmkuhl, Oriol; Vinuesa, Ricardo

DOI

[10.1016/j.ijheatfluidflow.2025.109913](https://doi.org/10.1016/j.ijheatfluidflow.2025.109913)

Publication date

2025

Document Version

Final published version

Published in

International Journal of Heat and Fluid Flow

Citation (APA)

Garcia, X., Miró, A., Suárez, P., Alcántara-Ávila, F., Rabault, J., Font, B., Lehmkuhl, O., & Vinuesa, R. (2025). Deep-reinforcement-learning-based separation control in a two-dimensional airfoil. *International Journal of Heat and Fluid Flow*, 116, Article 109913. <https://doi.org/10.1016/j.ijheatfluidflow.2025.109913>

Important note

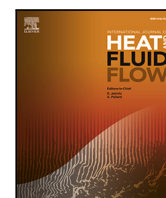
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Deep-reinforcement-learning-based separation control in a two-dimensional airfoil

Xavier Garcia^{a,*,}, Arnau Miró^{c,b}, Pol Suárez^a, Francisco Alcántara-Ávila^a, Jean Rabault^d, Bernat Font^e, Oriol Lehmkuhl^b, Ricardo Vinuesa^a

^a FLOW, Engineering Mechanics, KTH Royal Institute of Technology, Stockholm, Sweden

^b Barcelona Supercomputing Center, Barcelona, Spain

^c Universitat Politècnica de Catalunya, Barcelona, Spain

^d Independent Researcher, Oslo, Norway

^e Faculty of Mechanical Engineering, Technische Universiteit Delft, The Netherlands

ARTICLE INFO

Keywords:

Deep reinforcement learning
Active flow control
Drag Reduction
Flow Separation Control
Computational fluid dynamics
Fluid Mechanics
Airfoil
Energy Efficiency

ABSTRACT

The aim of this study is to discover new active-flow-control (AFC) techniques for separation mitigation in a two-dimensional NACA 0012 airfoil at a Reynolds number of 3000. To find these AFC strategies, a framework consisting of a deep-reinforcement-learning (DRL) agent has been used to determine the action strategies to apply to the flow. The actions involve blowing and suction through jets at the airfoil surface. The flow is simulated with the numerical code Alya, which is a low-dissipation finite-element code, on a high-performance computing system. Various control strategies obtained through DRL led to 43.9% drag reduction, while others yielded an increase in aerodynamic efficiency of 58.6%. In comparison, periodic-control strategies demonstrated lower energy efficiency while failing to achieve the same level of aerodynamic improvements as the DRL-based approach. These gains have been attained through the implementation of a dynamic, closed-loop, time-dependent, active control mechanism over the airfoil.

1. Introduction

Active flow control (AFC) research is crucial not only for the development of aerodynamics and aircraft design but also for all transport-related systems. This is primarily due to the critical relationship between drag reduction and overall system efficiency. Flow control strategies, whether passive (implemented through geometric modifications or additions) or active (utilizing devices that dynamically alter the flow over time), enable the manipulation of flow behavior to enhance aerodynamic performance.

Active flow control facilitates key aerodynamic benefits such as drag reduction, fuel consumption minimization, and noise reduction. Additionally, AFC can enhance lift, improve maneuverability — particularly in low-speed vehicles — and even contribute to radar signature reduction by strategically disturbing the flow.

Over the past decades, various passive flow control techniques have been developed. The earliest advancements in this field, including those during World War II, are reviewed by McLellan and Ladson (1988). A notable example is the surface-roughness technique studied by Beratlis et al. (2017), where different surface textures alter the boundary layer,

delaying flow separation and thereby reducing drag. Another significant contribution is the work of Bechert and Bartenwerfer (1989), who investigated the aerodynamic effects of fine longitudinal ribs in viscous flows.

Following the exploration of passive methods, the field moved toward active techniques. The first significant approach involved periodic excitations, as demonstrated by Amitay et al. (1998). Later, Greenblatt and Wygnanski (2000) extended these efforts to prevent flow separation using similar excitation strategies. A major breakthrough was the implementation of closed-loop controllers, as demonstrated by Muddada and Patnaik (2010), who used simple actuators to reduce drag in a low Reynolds number (Re) cylinder flow, achieving a 53% drag reduction.

More relevant to the present study, blowing and suction techniques have been widely explored for boundary layer control. Kametani and Fukagata (2011) investigated their effects on flow stability, while Voevodin et al. (2019) and Yousefi and Saleh (2015) studied suction and ejection mechanisms for drag reduction and flow control optimization.

In recent years, modern AFC strategies have increasingly incorporated deep reinforcement learning (DRL), utilizing artificial neural

* Corresponding author.

E-mail address: xgarcia@kth.se (X. Garcia).

networks (ANNs) for autonomous control strategy discovery. DRL has been recognized for its ability to optimize flow control problems in a computationally efficient manner. The first demonstration of DRL in AFC was provided by Rabault et al. (2019), where a DRL agent controlled the 2D flow around a cylinder at $Re = 100$, achieving an 8% drag reduction using a proximal-policy-optimization (PPO) algorithm. Later, Rabault and Kuhnle (2019) enhanced the framework through parallelization, significantly accelerating training efficiency. Tang et al. (2020) extended DRL-based AFC to higher Reynolds numbers using a more complex setup involving four synthetic jets.

Building on these early advancements, an increasing amount of research has confirmed the effectiveness of DRL for AFC applications. Ren et al. (2021) explored weakly turbulent 2D cylinder flows, while Varela et al. (2022) expanded these findings to higher Reynolds number cylinders. Cavallazzi et al. (2024) applied DRL to Couette flow, showcasing its versatility across different flow types. Font et al. (2025) further demonstrated DRL's adaptability in controlling turbulent separation bubbles. Additionally, DRL applications have been extended to Rayleigh–Bénard convection cases, as examined by Vignon et al. (2023a) and Vasanth et al. (2024).

Given the rapid progress in this field, interested readers may refer to broader review articles such as Vignon et al. (2023b) for an extensive overview of DRL-based AFC.

Among the most relevant prior studies, Suárez et al. (2024b) explored the first DRL-controlled three-dimensional (3D) cylinder flow, addressing transitional turbulent regimes in cylinder wakes by incorporating multiple control jets in the spanwise direction and managing complex flow structures. Building on this, Suárez et al. (2024a) extended the study to higher Reynolds numbers, where turbulence becomes even more intricate, further increasing the complexity of the active flow control (AFC) problem.

For AFC applications on airfoils, Zong et al. (2024) applied Q-learning to regulate plasma synthetic jets at an airfoil's trailing edge, aiming to suppress flow separation. Additionally, Wang et al. (2022) explored DRL-based synthetic jet control on a NACA 0012 airfoil at $Re_D = 3000$, employing three jets on the upper surface. Their results demonstrated a 27% drag reduction and a 27.7% lift enhancement. However, their strategy utilized a static jet intensity, lacking real-time adaptability as they use an open-loop structure.

This study aims to control the recirculation bubble that forms at the trailing edge of a two-dimensional NACA 0012 airfoil using a DRL-based computational framework. The methodology builds upon the approaches of Rabault et al. (2019), Suárez et al. (2024b), and Suárez et al. (2024a), integrating a proximal-policy-optimization (PPO) algorithm (Schulman et al. (2017)) for active flow control. The computational experiments are conducted at $Re_D = 3000$ using three synthetic jets placed on the upper airfoil surface. These jets are dynamically controlled by a DRL agent to optimize aerodynamic performance by reducing drag and increasing aerodynamic efficiency. The simulations are executed on the Nord 3 Supercomputer at the Barcelona Supercomputing Center (BSC-CNS) using Alya, a high-performance computational fluid dynamics (CFD) solver optimized for large-scale simulations (Vazquez et al., 2014).

The present work focuses on demonstrating DRL-based control efficacy in a consistent numerical framework, rather than absolute agreement with other datasets in the literature.

The remainder of this paper is organized as follows. Section 2 presents the methodology, detailing the deep reinforcement learning (DRL) framework, the numerical setup, and the validation of the Alya solver. Section 3 discusses the DRL training results for active flow control (AFC) on a two-dimensional airfoil. Finally, Section 4 summarizes the key findings and provides concluding remarks on the study's contributions and potential future directions.

2. Methodology

2.1. Domain, mesh, and numerical method

This study investigates DRL-based active flow control in a NACA 0012 airfoil immersed in a two-dimensional channel flow. The selected configuration facilitates both validation and direct comparison with the findings of Wang et al. (2022). All geometric parameters are normalized with respect to the airfoil chord length (D),¹ which serves as the reference scale throughout the study.

The computational domain, depicted in Fig. 1, extends $6D$ in the streamwise direction and $1.4D$ in the perpendicular direction (y). The coordinate system's origin is located at the leading edge of the airfoil, positioned at $(x = 1.5D, y = 0)$. The airfoil is oriented at an angle of attack of 10° , achieved through the appropriate rotational transformation.

The inlet boundary condition Γ_{in} is defined by a parabolic velocity profile, given by

$$U_{in}(y) = \frac{4U_m(0.7D - y)(0.7D + y)}{1.4D^2}, \quad (1)$$

where $U_m = 0.45$ represents the maximum velocity at the profile's center. The Reynolds number is defined as

$$Re_D = \frac{\bar{U}D}{\nu}, \quad (2)$$

where ν is the kinematic viscosity, and \bar{U} represents the mean velocity, computed as

$$\bar{U} = \int_{-0.7D}^{+0.7D} U_{in}(y) dy = \frac{2}{3}U_m. \quad (3)$$

From this, the mean velocity at the inlet is determined to be $\bar{U} = 0.3$.

A no-slip condition is applied at the airfoil surface and channel walls Γ_{wall} . At the outlet boundary Γ_{out} , a zero-gradient condition is enforced on velocity, while maintaining a constant pressure.

To enhance numerical stability, a buffer zone is incorporated near the outlet, wherein the fluid viscosity is artificially increased by an order of magnitude. This technique, adapted from Ferziger et al. (2020), effectively dissipates vortices, prevents recirculation instabilities, and ensures mass conservation. The absence of this stabilization mechanism could lead to numerical divergence and simulation failure due to unphysical recirculation effects at the outflow boundary.

The mesh used in the simulations is an unstructured triangular mesh with refinement near the airfoil, around the control jets, and in the wake region.

Three active control jets are positioned on the upper surface of the airfoil at $x/D = 0.2, 0.3$, and 0.4 , respectively. The actuators, labeled as jet₁, jet₂, and jet₃, are configured to ensure that actuation occurs perpendicular to the local airfoil surface. This setup ensures that any observed improvements in aerodynamic performance are due to active control rather than direct streamwise momentum injection. The jets produce a parabolic velocity profile, which ensures zero velocity at both edges.

The mass flow rate profile of each jet is defined by the following equation:

$$Q_i(x) = Q_{a_i} \sin\left(\frac{\pi(x - x_{2_i})}{x_{1_i} - x_{2_i}}\right), \quad (4)$$

where $i = 1, 2, 3$ corresponds to the three jets. Here, Q_i represents the mass flow rate profile of each jet, Q_{a_i} is the value selected by the agent for each jet, and x_{1_i} and x_{2_i} denote the initial and final streamwise coordinates of each jet, respectively.

¹ To maintain consistency with Wang et al. (2022), we use D to represent the airfoil chord length. This differs from the common convention where c is used to represent the chord.

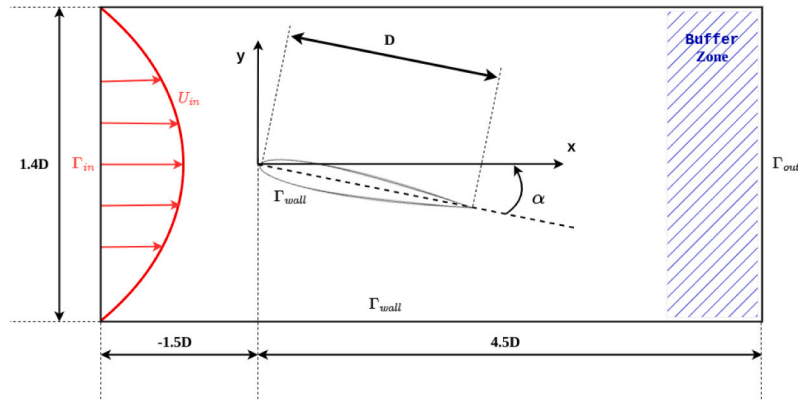


Fig. 1. Computational domain dimensions, normalized by the airfoil chord length D . The angle of attack α is illustrated, with boundary conditions represented by Γ . The buffer zone is highlighted in dark blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

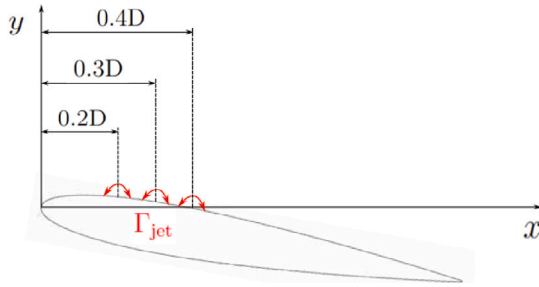


Fig. 2. Jets location with respect to the leading edge of the airfoil.

To ensure that blowing or suction always occurs perpendicular to the local airfoil surface, each jet is oriented normally to the surface by applying a rotation angle θ , given by:

$$\theta = 2 \tan^{-1} \left(\frac{x_1 - x_2}{(y_2 - y_1) + \sqrt{(y_2 - y_1)^2 + (x_1 - x_2)^2}} \right). \quad (5)$$

Thus, the velocity components in the streamwise and perpendicular directions are expressed as:

$$Q_x = Q \cos(\theta), \quad (6)$$

$$Q_y = Q \sin(\theta). \quad (7)$$

The numerical method used by Alya to compute the forces on the airfoil surface solves the incompressible viscous Navier–Stokes equations. These equations are formulated for the entire domain (Ω) as follows:

$$\partial_t u + (u \cdot \nabla)u - \nabla \cdot (2\nu \epsilon) + \nabla p = f, \quad (8)$$

$$\nabla \cdot u = 0, \quad (9)$$

where u represents the velocity field, and ϵ denotes the velocity strain-rate tensor, which is defined as a function of the velocity \mathbf{u} by

$$\epsilon = \frac{1}{2} (\nabla \mathbf{u} + \nabla^T \mathbf{u}). \quad (10)$$

The term f corresponds to the external body force acting on the flow, which in the present study is set to zero since no external body forces are applied.

The convective component in the non-linear term, $C_{\text{nonc}}(\mathbf{u}) = (\mathbf{u} \cdot \nabla)\mathbf{u}$, is formulated to ensure conservation of energy, momentum, and angular momentum, as detailed in Charnyi et al. (2017, 2019).

For time discretization, a semi-implicit Runge–Kutta scheme of second order is employed for the convective term, while a Crank–Nicolson scheme is applied for the diffusive term. During time integration, Alya

utilizes an eigenvalue-based time-stepping method, as implemented by Trias and Lehmkuhl (2011).

At each time step, the numerical solution of these equations is obtained, and the drag force (F_D) and lift force (F_L) are computed by integrating over the entire surface (S) of the airfoil:

$$\mathbf{F} = \int (\zeta \cdot \mathbf{n}) \mathbf{e}_j dS, \quad (11)$$

where ζ represents the Cauchy stress tensor, \mathbf{n} is the unit normal vector to the surface, and \mathbf{e}_j is a unit vector aligned with the main flow velocity for drag computation, and perpendicular to it for lift computation. The aerodynamic coefficients are then determined as:

$$C_d = \frac{2F_d}{\rho \bar{U}^2 D}, \quad (12)$$

$$C_l = \frac{2F_l}{\rho \bar{U}^2 D}. \quad (13)$$

2.2. DRL characteristics

The DRL framework is illustrated in Fig. 3, which depicts each step in the training process. Initially, the solver, Alya, performs a simulation without any jet interaction. This reference simulation, referred to as the baseline, is executed from 0 to 200 time units (TU). Once the baseline simulation is completed, the episodes are executed sequentially, with each episode starting from the final step of the baseline.

Each episode consists of a sequence of actions (a_t), corresponding to the actuation of the jets. The selection of these actions is performed by the agent, which makes decisions based on the received state (s_t) and the corresponding reward (r_t).

Once an episode is completed, the next episode begins. At the start of each new episode, the agent's policy and networks retain their weights learned during the preceding episode, thus exploiting prior learning to accelerate convergence. Each new episode starts from the last time step of the baseline simulation. When all episodes have been simulated, a complete training cycle is concluded.

The DRL environment communicates with the agent through three distinct channels. The first channel is the observation state, s_t , which represents the pressure values at various predefined locations in the computational domain, referred to as probes or witness points. These points are strategically placed in critical areas around the airfoil and in the wake region, as illustrated in Fig. 4. The extracted pressure values are normalized to ensure that the data received by the agent remains within the range $[-1, 1]$. This approach follows Rabault et al. (2019), wherein a single heuristic scalar (based on prior knowledge of expected pressure magnitudes) is applied uniformly to all sensor readings to map the raw values into $[-1, 1]$. This uniform normalization (i) simplifies the interpretation of state fluctuations across different

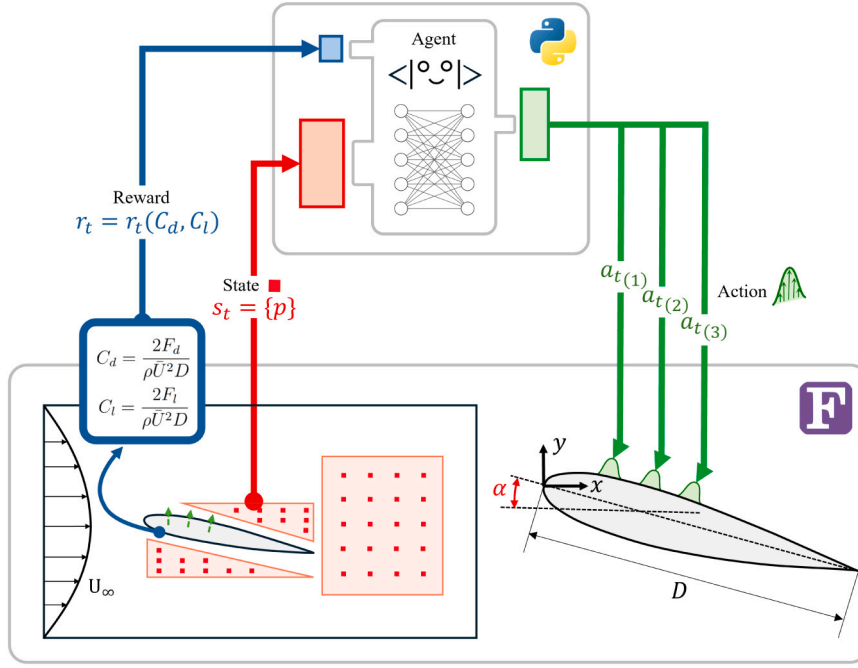


Fig. 3. Schematic representation that illustrates the reinforcement-learning framework applied to a two-dimensional airfoil, showing communication channels between two main actors. In this case, the direction of the information is clockwise. At the top, we show the agent architecture featuring a shared neural network. At the bottom, the computational-fluid-dynamics (CFD) environment is depicted, with the airfoil chord D as the reference length.

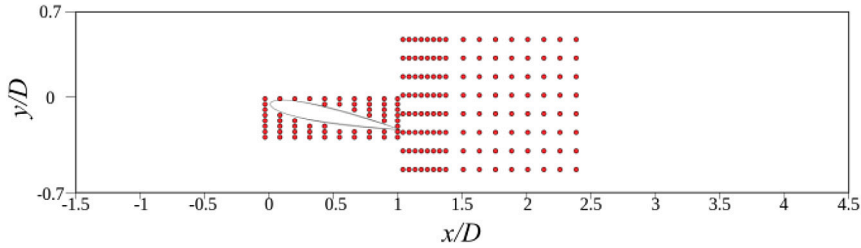


Fig. 4. Schematic representation of the computational domain, where red dots indicate the locations of probes used to define the state s_t . The probes are strategically placed around the wing and in the wake region to capture relevant flow information.

sensors and (ii) conforms to Tensorforce recommendations, improving optimizer efficiency by reducing input precision.

The second communication channel is the action, a_t , which corresponds to the value determined by the agent and applied to the control jets for active flow manipulation. Although the action a_t is initially bounded within $[-1, 1]$ to suit the DRL algorithm, the actuation intensity Q_i is subsequently refined via Eqs. (6), (7), and (14), with the agent selecting Q_i through its policy $\pi(a_t | s_t)$, thus ensuring a smooth temporal transition. The selected value of Q_i is determined by the agent based on the policy function $\pi(a_t | s_t)$.

$$Q(t) = Q_i + g(t)(Q_{i+1} - Q_i), \quad (14)$$

where $t = (t - t_i)/(t_{i+1} - t_i)$, and the function $g(t)$ is defined as

$$g(t) = \frac{f(t)}{f(t) + f(1 - t)}, \quad (15)$$

with $f(t)$ given by an exponential function that ensures a smooth transition:

$$f(t) = \begin{cases} e^{-\frac{1}{t}}, & \text{if } t > 0, \\ 0, & \text{if } t \leq 0. \end{cases} \quad (16)$$

Maintaining a smooth transition is crucial to facilitating the learning process for the agent while avoiding abrupt changes that could introduce instability. Previous experiments have demonstrated that using

step or linear functions poses challenges for effective learning and computational burden.

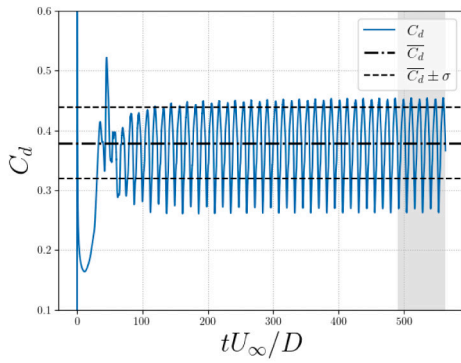
The final communication channel between the environment and the agent is the reward, r_t , which serves as the primary criterion for the PPO agent to decide future actions. The reward function can be formulated based on different objectives, depending on the specific focus of each training scenario. In this study, a single reward function is used, integrating both drag reduction and lift modulation through tunable weights, and offset by σ to ensure the reward starts at zero:

$$r = \mu \left[\sigma + w_d (C_{d_0} - C_d) + w_l (C_l - C_{l_0}) \right], \quad (17)$$

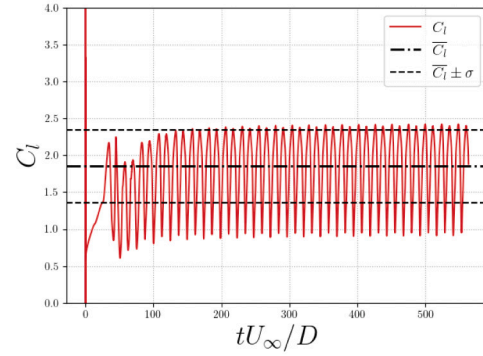
where:

- μ is a scaling factor,
- σ is an offset constant chosen so that $r = 0$ when $(C_d, C_l) = (C_{d_0}, C_{l_0})$,
- w_d and w_l are the weights for drag and lift contributions, respectively,
- C_{d_0} and C_{l_0} are the reference drag and lift coefficients.

Finally, Table 1 provides a summary of all the parameters considered, highlighting the key characteristics of the two main cases examined in this study.



(a) Drag coefficient (C_d) during the baseline simulation, with its average value ($\overline{C_d}$).



(b) Lift coefficient (C_l) during the baseline simulation, with its average value ($\overline{C_l}$).

Fig. 5. Baseline aerodynamic coefficients for drag and lift.

Table 1

Comparison of parameters used in the simulations and DRL framework for the two cases: (1) C_d reduction and (2) C_l/C_d enhancement. Shared parameters apply to both cases, while case-specific parameters differ between the two configurations.

Shared Parameters		
Mesh Cells	184 790	
Witness Points	178	
Q_{norm}	1.25	
Time per Action	0.275	
Actions per Episode	300	
Batch Size	33	
CPUs - Environment	16	
Environments	33	
Total CPUs	544	
Network	{[Dense, 512], [Dense, 512]}	
Update Frequency	1.0	
Learning Rate	0.001	
Multi-Step	25	
Subsampling Fraction	0.2	
Likelihood Ratio Clipping	0.2	
Discount	0.99	
Optimizer Steps	5	
Optimizer Type	Adam	
Optimizer Learning Rate	0.001	
Case-Specific Parameters		
Parameter	Case 1 C_d Reduction	Case 2 C_l/C_d Enhancement
μ	1.0	1.2
λ	0.3	N/A
σ	0.0	-3.88
w_1	N/A	0.3
w_2	N/A	0.7
Q_{max}	± 0.5	± 1.5
Number of Episodes	957	1188

2.3. CFD validation

The framework has been previously validated for flow control around cylinders in studies such as Varela et al. (2022) and Suárez et al. (2024b). To extend this validation to an airfoil configuration, it is necessary to perform a baseline simulation without actuation. This baseline allows for a direct comparison with the results of Wang et al. (2022) in terms of lift and drag coefficients.

The baseline results are presented in Figs. 5(a) and 5(b). The computed average drag coefficient is $C_d = 0.376$, while the average lift coefficient is $C_l = 1.822$, yielding an aerodynamic efficiency of $C_l/C_d = 4.846$.

When compared to the results reported by Wang et al. (2022), our baseline drag coefficient differs by approximately 15%, while the lift coefficient shows a larger variation of about 45%. This increased lift is

a well-known consequence of channel (ground-effect) confinement: the proximity of the lower wall accelerates flow beneath the airfoil, adding a “virtual” lift component (Vinueza et al., 2015). As illustrated in Figs. 6(a) and 6(b), the compression of streamlines near the leading edge and beneath the airfoil (Fig. 6(a)) and the upstream symmetry loss before airfoil interaction (Fig. 6(b)) vividly demonstrate this ground-effect phenomenon.

To verify that our results accurately capture this effect and are not artifacts of mesh or solver settings, we conducted a mesh-convergence study using three graded meshes. Across these meshes, both C_D and C_L varied by less than $\pm 2\%$, demonstrating mesh independence. Furthermore, we performed an open-domain (unconfined) simulation at $Re_D = 3000$ using the same solver settings; the resulting coefficients agree within 3% of established benchmarks (Swanson and Langer, 2016), confirming the solver’s accuracy. These consistency checks substantiate the reliability of our present simulations and isolate channel confinement as the source of the lift increase.

3. Results and discussion

This section is divided into two parts. The first part illustrates the results obtained from intensity-fixed solutions following the same approach as the study conducted by Wang et al. (2022). The final part discusses the outcomes of applying deep reinforcement learning (DRL) for synthetic jet control in airfoil training.

3.1. Application of fixed-intensity control in synthetic jets

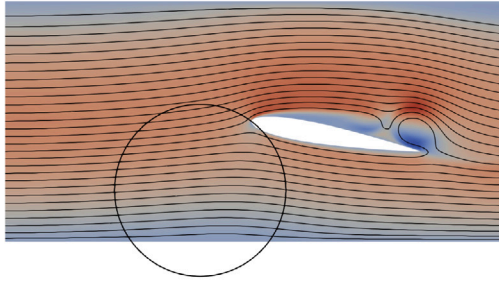
This subsection presents the results obtained by applying a fixed-intensity control strategy, following the approach described in Wang et al. (2022). The primary objective is to validate state-of-the-art techniques and assess variations in intensity magnitude to determine whether different values lead to improved performance.

To establish a direct comparison, it is necessary to translate the mass flow rate values from Wang et al. (2022) into the framework used in this study. Their work employs a dimensionless parameter, Q^* , defined based on a reference mass flow rate, $Q_{\text{ref}} = 0.3$, which can be expressed in terms of the normal mass flow rate as follows:

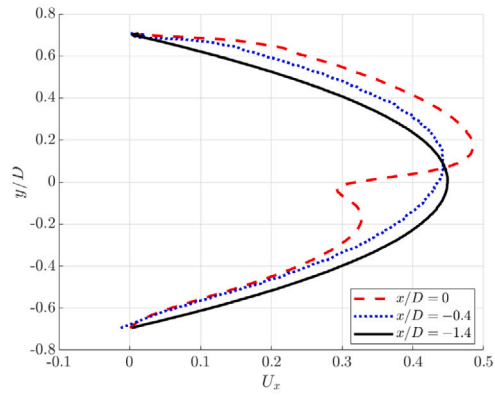
$$Q_{\text{jet}_i} = Q^* Q_{\text{ref}}. \quad (18)$$

The numerical framework in this study allows for tuning the values selected by the agent, which must be appropriately translated for use in the computational fluid dynamics (CFD) solver. This transformation is achieved using Eq. (19):

$$Q_{\text{jet}_i} = Q_{\text{agent}} Q(x), \quad (19)$$



(a) Flow streamlines depicting the ground effect.

(b) Velocity distribution in the x -direction at locations $(x/D = -1.4D, -0.4D, 0D)$, highlighting ground effect asymmetry before interaction with the airfoil.**Fig. 6.** Representation of the ground effect: (a) Flow streamlines and (b) velocity profile in the x -direction at selected locations.**Table 2**

Jet intensities used in Wang et al. (2022) translated into this study's framework.

Q_{jet_1}	Q_{jet_2}	Q_{jet_3}
-1.25	1.25	-0.125

where $Q(x)$ is given by:

$$Q(x) = \int_{x_1}^{x_2} \sin \left[\pi \frac{x - x_2}{x_1 - x_2} \right] \times \cos \left[2 \arctan \left(\frac{x_1 - x_2}{(y_2 - y_1) + \sqrt{(y_2 - y_1)^2 + (x_1 - x_2)^2}} \right) \right] dx, \quad (20)$$

where x_1 and x_2 correspond to the start and end coordinates of the jet in the x -direction, while y_1 and y_2 represent the corresponding coordinates in the y -direction.

The values of Q_{agent} corresponding to the same mass flow rate as in Wang et al. (2022) are presented in Table 2.

The results of the simulation, after reaching convergence, are shown in Table 3. The evolution of the flow behavior and the aerodynamic coefficients can be observed in Figs. 7 and 8. The baseline averages are computed from $t = 50$ to 200 because the uncontrolled simulation attains a statistically stationary state only after $t = 50$, excluding initial transients. The sudden decline in both coefficients at $t = 200$ corresponds to the instant when actuation is activated, demonstrating the control effect.

We first apply the exact same fixed-intensity control inputs (jet positions and amplitudes) as in Wang et al. (2022) to both benchmark our solver fidelity and isolate numerical effects. A comparison between actuated and non-actuated flows reveals that vortex shedding generates a recirculation bubble near the trailing edge, causing localized separation and increased skin friction. The achieved drag reduction is approximately 28%, closely matching the 27% reduction reported by Wang et al. (2022). However, the lift coefficient exhibits a -3% change in this study compared to a +27% increase in their results, primarily due to channel (ground)-effect confinement. Despite this, aerodynamic efficiency improves by approximately 34%, indicating a significant gain even in the absence of lift augmentation. In this section, we then manually vary the steady jet intensities to assess potential further improvements; time-varying control via reinforcement learning is introduced in Section 5.2.

An analysis of the flow dynamics reveals that once the jets are activated using a fixed-intensity control strategy, the flow stabilizes

Table 3

Results obtained using the jet intensities from Table 2 and comparison with baseline coefficients. Values marked with an asterisk (C_d^* , C_l^*) correspond to results reported by Wang et al. (2022).

C_d	C_l	C_l/C_d	ΔC_l	ΔC_d	$\Delta C_l/C_d$	C_d^*	C_l^*	C_l/C_d^*
0.27	1.77	6.5	-28.23%	-2.89%	+34.18%	0.23	1.27	5.45

Table 4

Definition of cases for the fixed-policy implementation, ranging from low to high mass flow rate intensities.

Case	Q_{jet_1}	Q_{jet_2}	Q_{jet_3}
#1	-0.625	0.7	-0.075
#2	-0.8	0.9	-0.1
#3	-0.9	0.9	-0.09
#4	-1	1.1	-0.1
#5	-1.2	1.2	-0.12
#6	-1.25	1.25	-0.125
#7	-1.35	1.5	-0.15
#8	-1.5	1.35	-0.15
#9	-1.65	1.5	-0.15
#10	-1.8	2	-0.2
#11	-2.25	2.5	-0.225

completely beyond 115 TU , where $TU = tU_\infty/D$. Beyond this point, vortex shedding ceases to occur. The ability to achieve such flow stabilization and predictability using a simple and adaptable policy highlights the potential of this approach.

Once the strategy has been reproduced, an additional objective of this study is to explore alternative fixed-intensity control policies and evaluate their outcomes. To ensure a comprehensive analysis, a range of intensities from low to high has been examined. In total, 11 cases, including the reference case, have been simulated. These cases are summarized in Table 4.

The aerodynamic coefficients resulting from these implementations are shown in Fig. 9. The trends observed indicate that increasing the intensity while maintaining the same policy leads to an increase in both the drag coefficient C_d and the lift coefficient C_l , while the aerodynamic efficiency remains relatively constant. This suggests that the selection of a particular control policy should depend on the specific optimization objective. For instance, if the primary goal is to reduce drag while maintaining lift, case #5 would be a suitable choice. Conversely, if the objective is to maximize lift, case #10 appears to be the optimal strategy.

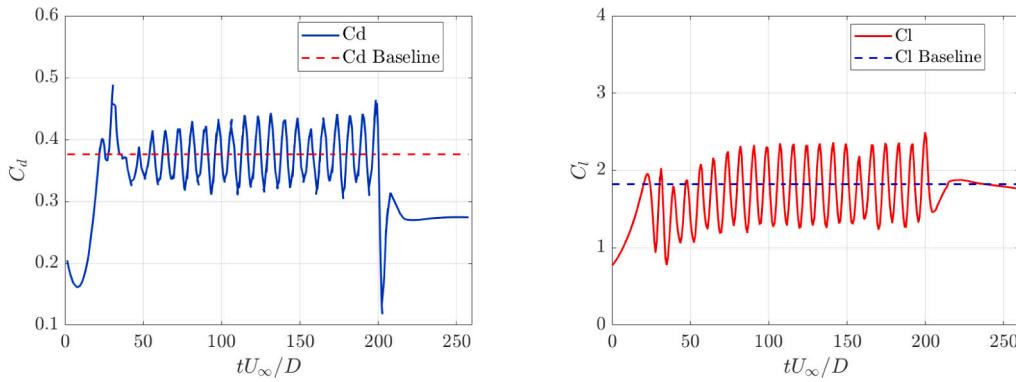


Fig. 7. Evolution of drag and lift coefficients when applying the fixed-intensity strategy from Wang et al. (2022) compared to baseline values. Baseline averages are taken over $t = 50$ – 200 once stationarity is reached; the sharp drop at $t = 200$ marks the onset of actuation.

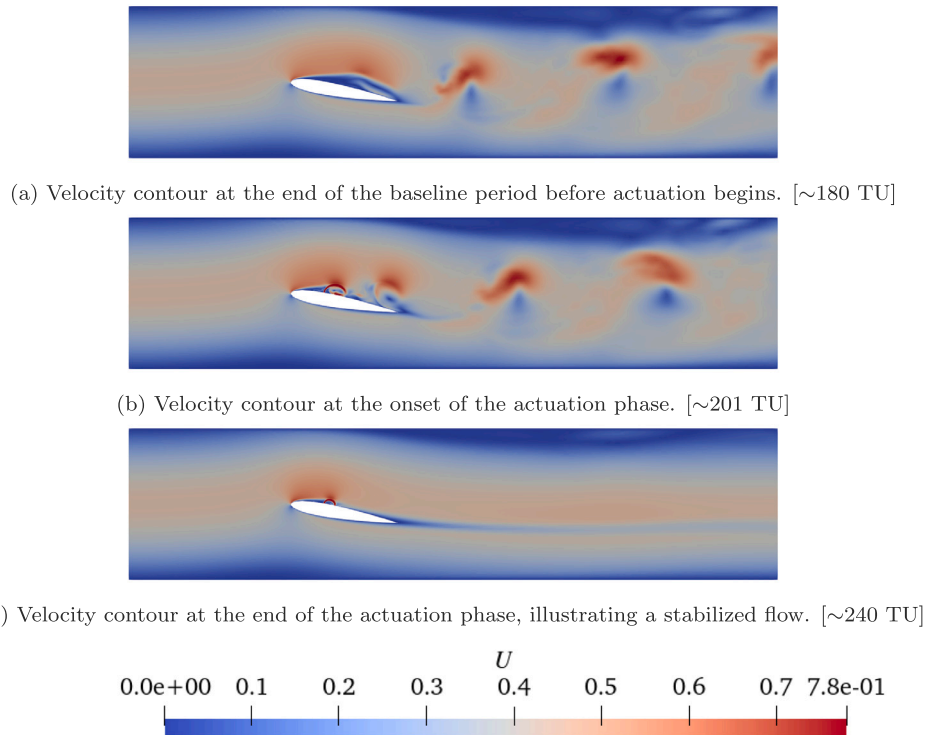


Fig. 8. Comparison of flow conditions between baseline, onset of actuation, and stabilized flow using the fixed-intensity policy strategy from Wang et al. (2022). Actuator positions (jet_1 , jet_2 , and jet_3) are defined in Fig. 2 and shown in the computational-domain overview (Fig. 3); they are omitted here to preserve the resolution of the depicted flow features.

A critical issue with excessive mass flow rate injection is the abrupt and unpredictable drop in aerodynamic performance. When the injected mass flow rate becomes too large, it disrupts the aerodynamic structures, leading to a loss of lift generation.

As shown in Fig. 9, the case achieving the greatest drag reduction is #1, with a reduction of 38.1%. The highest increase in lift is observed in case #10, with a 42.7% increase. The most efficient overall case is #8, yielding a 40.1% improvement in aerodynamic efficiency.

To further illustrate the flow behavior under different fixed-intensity actuation strategies, Fig. 10 presents a comparison of the stabilized flow fields for low, medium, and high intensity cases. In the low-intensity case, vortex shedding persists, though with a significantly reduced amplitude. In the medium-intensity case, the flow stabilizes completely, and no vortex shedding is observed. In the high-intensity case, the flow over the upper surface of the airfoil accelerates, leading to increased lift generation due to a higher pressure gradient.

3.2. DRL training results of a 2D airfoil at Re_D 3000 using synthetic jets

This section presents the main findings of the study, where the complete implementation of the framework is utilized to identify optimal strategies for flow control in two distinct directions. As previously explained, the objectives are to achieve drag reduction on one side and efficiency enhancement on the other.

3.2.1. Drag reduction strategy case

In the first case, focusing on drag reduction, a total of 957 episodes were used to train the agent. The evolution of the training process is presented in Fig. 11. The final deterministic case obtained from this training is depicted in Figs. 12(a) and 12(b). Here, “deterministic” refers to the evaluation mode of the DRL agent: after training with exploration noise to discover optimal control policies, all exploration is disabled and the policy’s mean action is applied at each time step, yielding a repeatable control law.

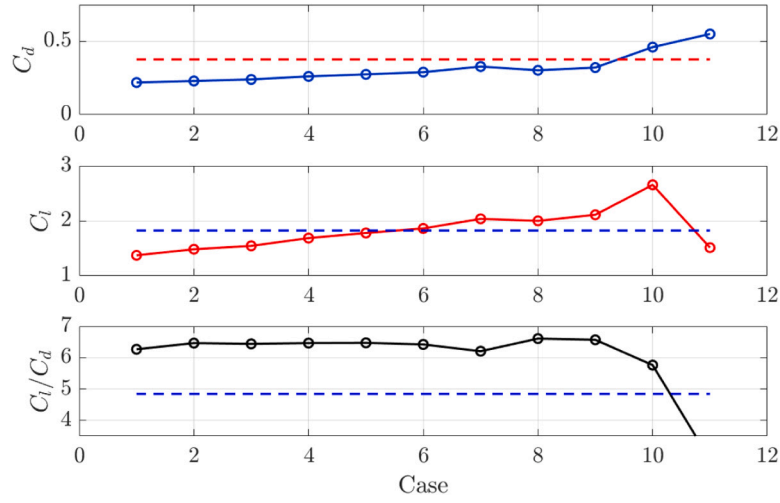


Fig. 9. Aerodynamic coefficients resulting from the fixed-intensity policies listed in Table 4, compared to the baseline values (represented by the dashed line).

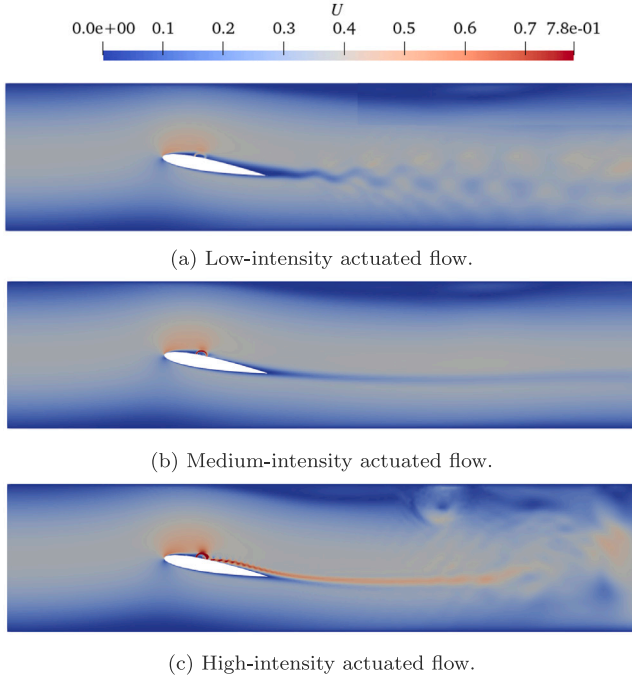


Fig. 10. Comparison of flow behavior for low, medium, and high-intensity actuation cases using the fixed-intensity policy.

By implementing the control actions illustrated in Fig. 13(a), with further details provided in Fig. 13(b), the results indicate an average drag reduction of 43.89% and a lift reduction of 35.89% during the actuated phase. Compared to the best drag reduction case from Section 3.1, the DRL-based strategy achieves a 6% greater drag reduction.

The control actions employed by the agent exhibit a distinct pattern rather than remaining constant. The jets alternate between blowing (adding momentum) and suction (reducing momentum) throughout the control process. Jets 1 and 2 exhibit similar actuation patterns, indicating that they perform nearly identical actions. In contrast, Jet 3 consistently executes the opposite action to Jets 1 and 2, ensuring that the net mass flow rate remains zero.

This result can be further analyzed by examining Fig. 14, which presents the pressure coefficient distribution along the airfoil for both the actuated drag reduction case and the baseline. The figure illustrates

how the jets decelerate the flow over the upper surface, leading to a reduction in lift. Additionally, it is important to highlight that the small oscillations near the trailing edge diminish in magnitude and smoothly converge with the lower surface, thereby forming a weaker wake. This process effectively reduces the primary source of drag.

The primary mechanism responsible for drag reduction is the control of vortex shedding and the associated recirculation bubble. From the perspective of vortex shedding, Fig. 15(a) presents the pressure signal recorded at a witness point located at $[x/D, y/D] = [2, 0]$ in the wake region. The results indicate that the jet-induced strategy reduces the amplitude of oscillatory shedding. The magnitude of these oscillations is decreased by approximately 62.5%, allowing the dissipation of discontinuities at a faster rate and preventing the formation of strong eddies in the wake, which are otherwise observed in the non-actuated flow field.

Another major contributor to drag is the boundary layer and the resulting skin friction Örlü and Vinuesa (2020). It is essential to assess whether the boundary layer remains attached or detaches and whether the jets exert sufficient control authority to influence its behavior. This is illustrated in Fig. 15(b), which shows the streamwise shear stresses around the airfoil. The figure highlights how the jets energize the boundary layer, with the first jet actively injecting momentum. Beyond the jet region, the oscillatory patterns seen in the C_p plot are also visible in the shear stress distribution, indicating multiple instances of detachment and reattachment before reaching the trailing edge. A comparison between actuated and non-actuated flows reveals that in the uncontrolled (baseline) case, vortex shedding generates a recirculation bubble near the trailing edge, causing localized separation and increased skin friction. The clearest visualization of the recirculation bubble is provided in Fig. 14, where the non-actuated velocity field is depicted.

By analyzing the Reynolds stresses at different streamwise locations, valuable insights into the turbulent structures and flow characteristics can be obtained. Fig. 16a presents the Reynolds stresses along the streamwise direction, where the oscillations in both the baseline and actuated flows exhibit noticeably smaller magnitudes. At $x/D = 2.5$, in the middle of the wake, the expected streamwise turbulence is observed, as the inflow follows a parabolic profile. This suggests that the actuated wake more closely resembles the inflow compared to the baseline case, reinforcing the drag reduction mechanism through improved pressure distribution. A similar trend is observed in Figs. 16b and 16c, where wake positions exhibit lower stress values. When turbulence-related stresses are captured, they are primarily concentrated in the middle of the wake, as expected.

In Fig. 16b, the effects of jet actuation are evident, leading to significant turbulence mixing in the wall-normal direction. However, this

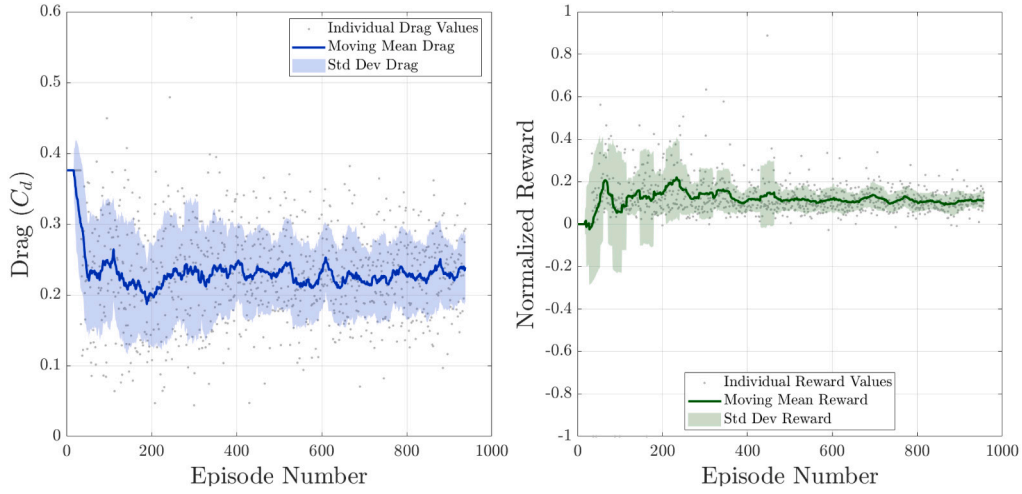
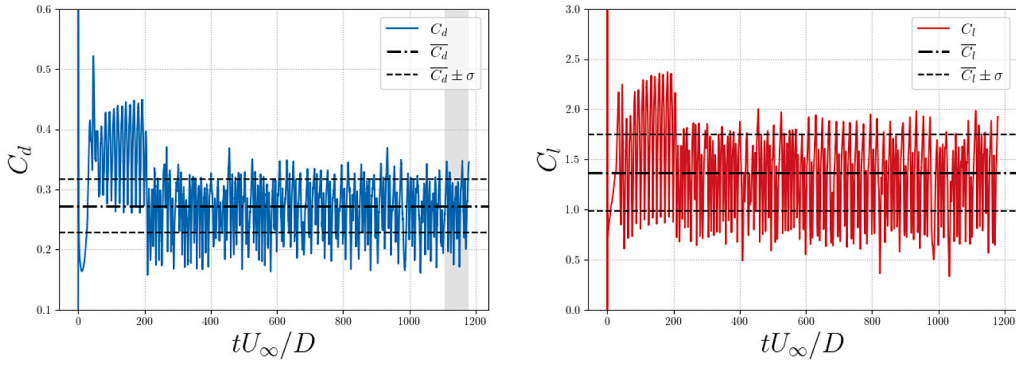


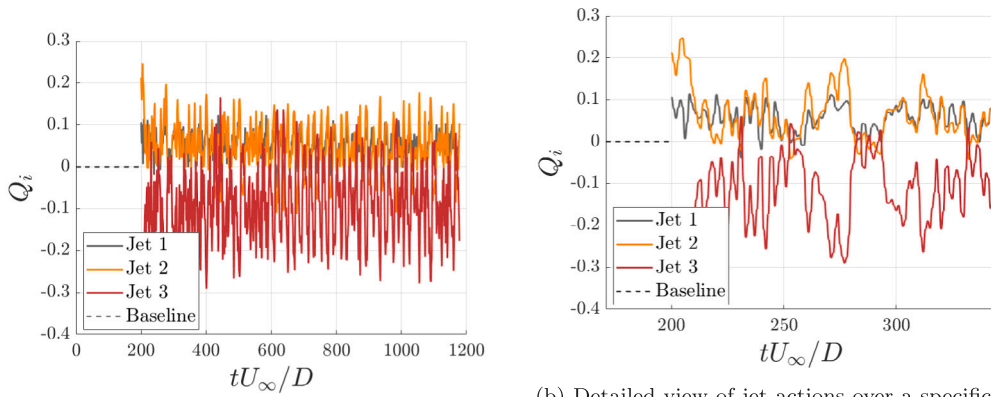
Fig. 11. Comparison of the evolution of the drag coefficient (C_d) and normalized rewards during training. The left subplot illustrates drag values with a moving average and standard deviation, highlighting the trend and variability over episodes. The right subplot presents the rewards, showing progress and fluctuations in the reward signal. The shaded regions indicate the standard deviation for each metric.



(a) Drag coefficient throughout the deterministic case. Actuation starts at 200 TU.

(b) Lift coefficient throughout the deterministic case. Actuation starts at 200 TU.

Fig. 12. Evolution of drag and lift coefficients throughout the deterministic case, with actuation commencing at 200 TU.



(a) Jet actions during the deterministic case for drag reduction.

(b) Detailed view of jet actions over a specific time range, providing insight into the agent's control decisions.

Fig. 13. Comparison of jet actions during the deterministic case for drag reduction, including a detailed analysis of a specific time range.

influence results in a smaller turbulent region along the upper surface of the airfoil compared to the non-actuated case. While turbulence mixing persists, its intensity is reduced to half the magnitude observed in the non-actuated flow, with this reduction becoming even more pronounced when analyzing turbulence in the wake region.

Fig. 16c exhibits a similar behavior along the airfoil. The first notable discrepancies arise around $x/D = 1$, indicating that the non-actuated flow field experiences interaction between the streamwise and wall-normal directions due to the recirculation bubble formed in that region on both the upper and lower surfaces. In contrast, for the

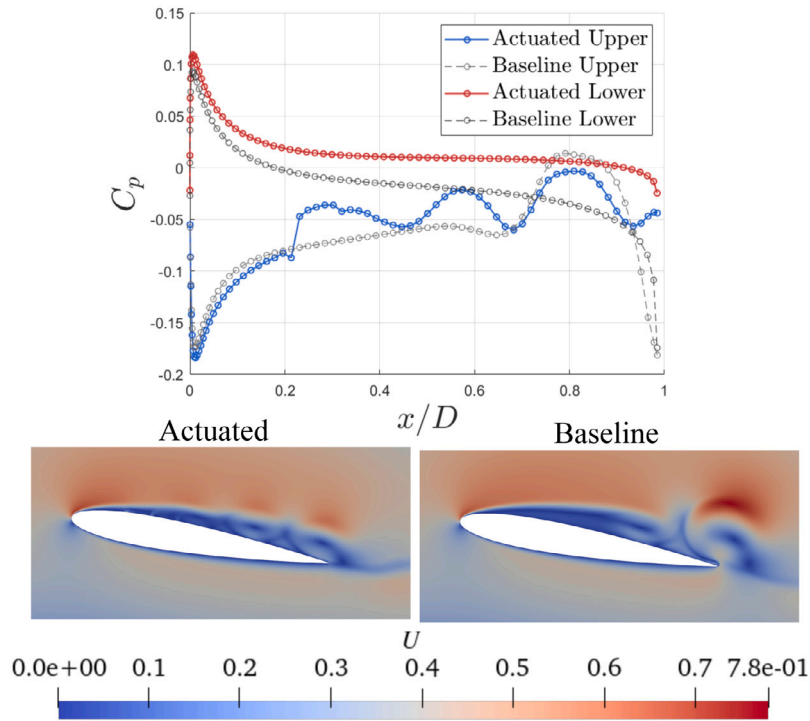
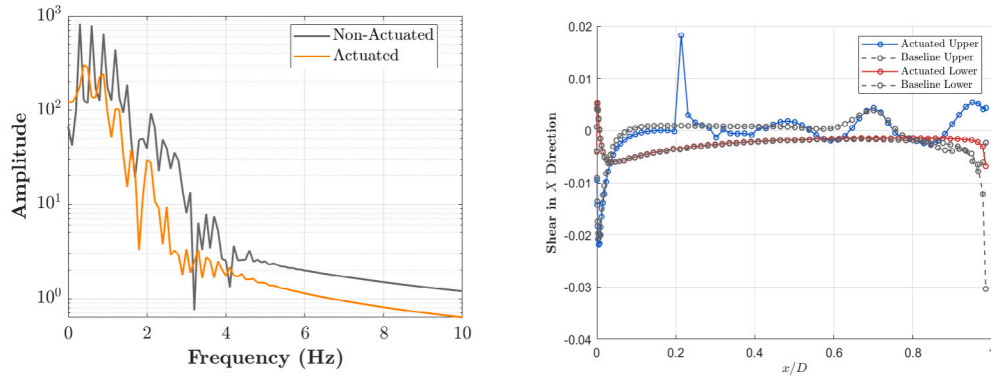


Fig. 14. Comparison of C_p along the airfoil for the baseline (right) and actuated (left) cases. Velocity contours for the instants used to extract the C_p distribution are also shown.



(a) Frequency spectrum comparison of the wake for non-actuated and actuated cases. Actuation significantly reduces the amplitude of dominant frequencies, leading to a weakened wake and reduced vortex shedding.

(b) Comparison of streamwise shear stress along the airfoil surface for actuated and baseline cases, highlighting the impact of actuation on boundary layer behavior and attachment.

Fig. 15. Comparison of frequency spectrum (left) and streamwise shear stress (right) for non-actuated and actuated cases, demonstrating the effect of actuation on wake dynamics and boundary layer behavior.

actuated case, turbulence mixing and eddy generation are restricted to a single peak, suggesting the absence of a recirculation bubble, as the flow field predominantly mixes in a single direction. Beyond this region, in the wake, the interaction between turbulent eddies and the main stream flow decays significantly in the actuated case compared to the non-actuated flow, indicating a weaker and more stable wake.

3.2.2. Efficiency enhancement strategy case

For the second case, different training strategies have been employed to optimize the agent's performance. As described in Section 2.2, two distinct reward formulations have been utilized. In the presented deterministic case, the agent is trained using the third reward function, which separates lift and drag while assigning different weights to each parameter. Specifically, weights of 0.3 and 0.7 were applied to drag and lift, respectively, following Eq. (17).

A total of 1188 episodes were required for training, with the evolution of the training process illustrated in Fig. 17. Fig. 17 summarizes the agent's learning progress over training episodes. Aerodynamic efficiency increases along the episodes starting from the baseline value (4.846), demonstrating that the control policy yields progressively better performance. The normalized reward, although negative in absolute value, rises towards zero (becoming less negative) indicating continual improvement of the learned strategy until stabilization. Note that the absolute sign of the reward is less important than its upward trend, which confirms convergence to an optimal policy.

The results of the deterministic strategy are presented in Figs. 18(a), 18(b), and 19, depicting the evolution of drag, lift, and aerodynamic efficiency, respectively. The final outcome yields an 11.97% reduction in drag, a 39.63% increase in lift, and an overall efficiency improvement of 58.64%. When compared to the best fixed-intensity case from

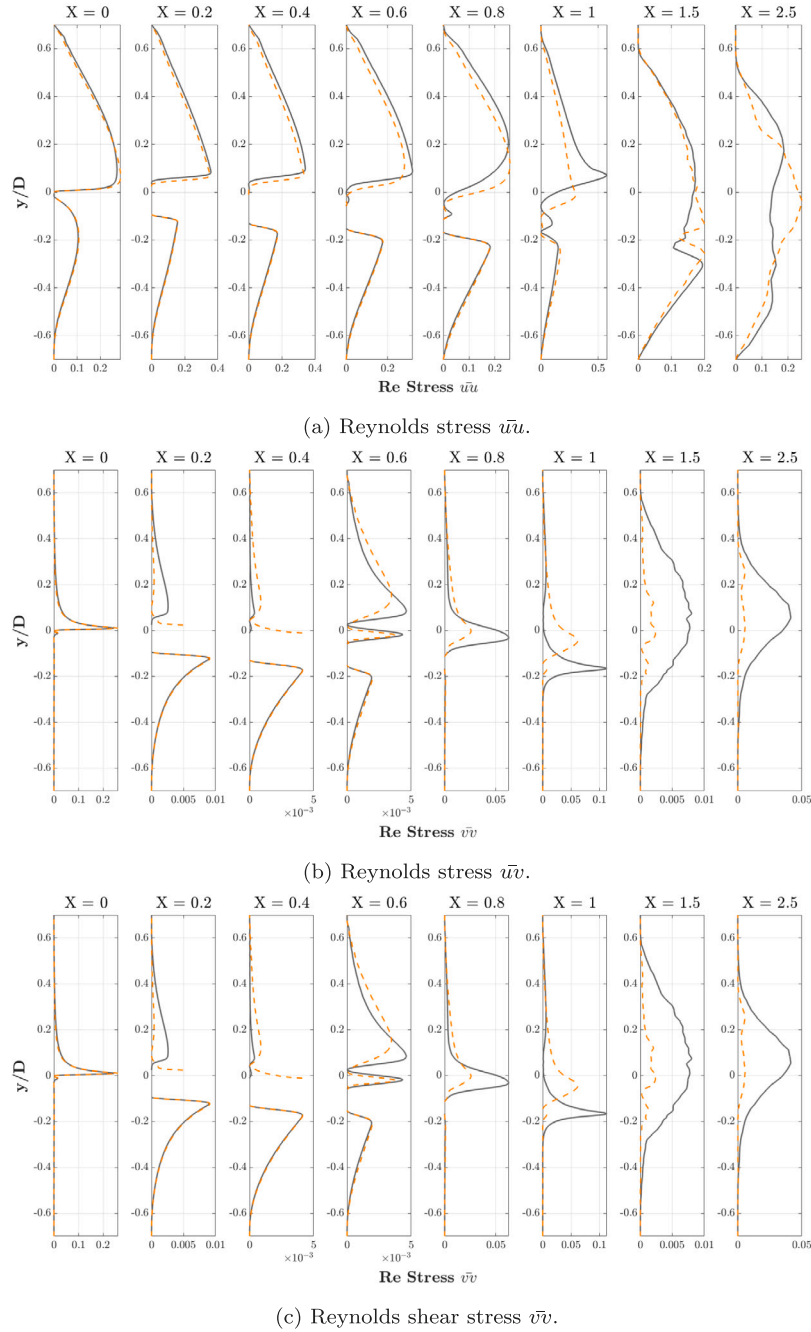


Fig. 16. Comparison of Reynolds stresses ($\bar{u}u$, $\bar{u}v$, and $\bar{v}v$) along the Y -coordinate at different streamwise locations (x/D). The dashed orange lines represent the actuated case, while the solid gray lines correspond to the baseline flow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Section 3.1, the DRL agent achieves an additional 18% improvement in aerodynamic efficiency.

As observed in Fig. 20, the control actions remain nearly constant throughout the actuation phase. This trend has been consistently identified in all efficiency-driven cases. Regardless of the specific reward function or parameter adjustments, the agent converges toward a steady-state strategy, with a transition phase lasting no longer than 15 to 20 actions.

Notably, the control actions in this case have peak actuation intensities approximately ten times greater in magnitude than those employed in the drag-reduction case discussed in Section 3.2.1. By employing a higher mass flow rate, the agent and jets effectively transform the

baseline dynamic state, characterized by vortex shedding and a recirculation bubble, into a completely steady-state flow, controlled by the induced jet structure. As illustrated in Figs. 20 and 21, jet₁ and jet₂ primarily function to counteract the mass flow introduced by jet₃. This interaction is particularly significant and warrants deeper investigation.

The resulting flow structure closely resembles that identified in Wang et al. (2022), though with a key difference: in their study, jet 2 was responsible for adding mass flow, while jet 1 facilitated the formation of a circular flow structure. In contrast, in the present study, the generated recirculation bubble slightly modifies the upper region of this structure, leading to an elliptical rather than a circular shape. This elliptical flow pattern accelerates in the direction of the airfoil before realigning with the main flow.

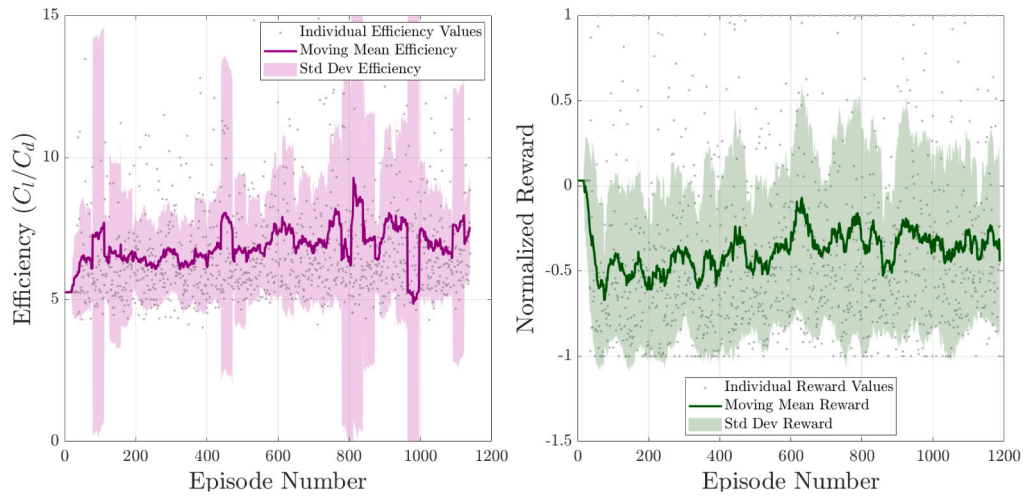
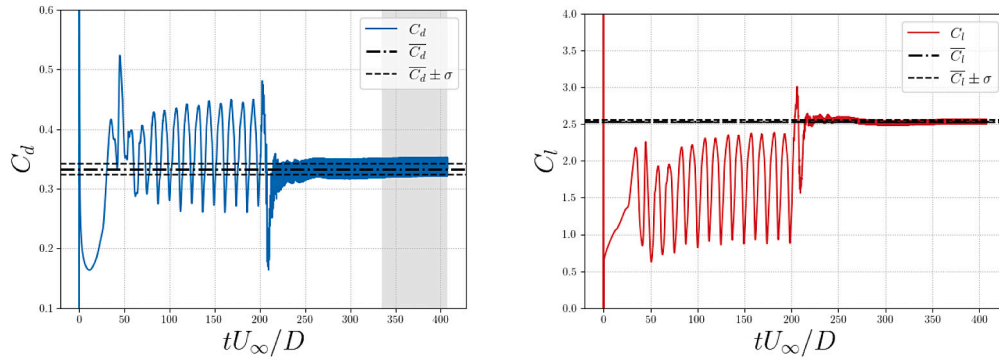


Fig. 17. Comparison of the evolution of aerodynamic efficiency (C_l/C_d) and normalized rewards during training. The left subplot illustrates efficiency values with a moving average and standard deviation, capturing trends and variability over episodes. The right subplot presents the reward evolution, highlighting progress and fluctuations. The shaded regions indicate standard deviation for each metric.



(a) Drag coefficient (C_d) evolution over time, showing oscillatory behavior during the transient phase and a reduction in drag in the steady-state region under DRL-based active flow control.

(b) Lift coefficient (C_l) evolution over time, illustrating initial transient oscillations and stabilization at a higher value due to jet actuation. The shaded region represents steady-state statistics.

Fig. 18. Time evolution of aerodynamic coefficients under DRL-based active flow control: (a) Drag coefficient (C_d), showing steady-state reduction; (b) Lift coefficient (C_l), highlighting stabilization at higher values.

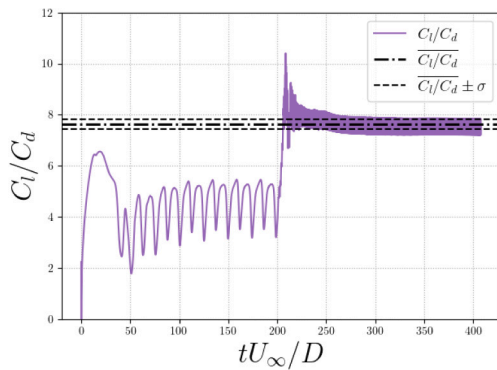


Fig. 19. Aerodynamic efficiency (C_l/C_d) over time, illustrating the improvement achieved through DRL-based active flow control. The shaded region represents time-averaged steady-state values with standard deviation.

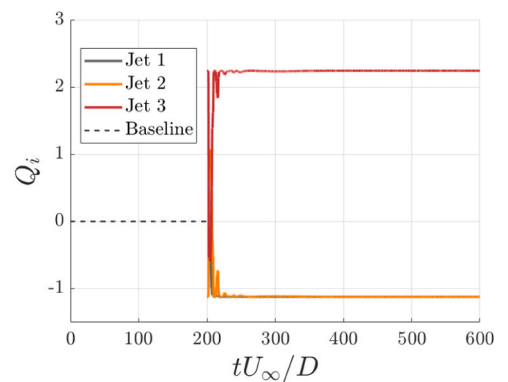


Fig. 20. Time evolution of the mass flow rate Q for the three jets, illustrating the transition from the baseline to the steady-state actuation phase.

A detailed examination of Fig. 21 reveals that, prior to jet actuation, the region of highest flow acceleration creates a near-zero velocity zone, effectively forming a void. The jets mitigate this effect

by generating a continuous flow structure that not only accelerates the flow in the initial region but also decelerates it downstream, as the last jet counteracts the main flow direction. This process reduces

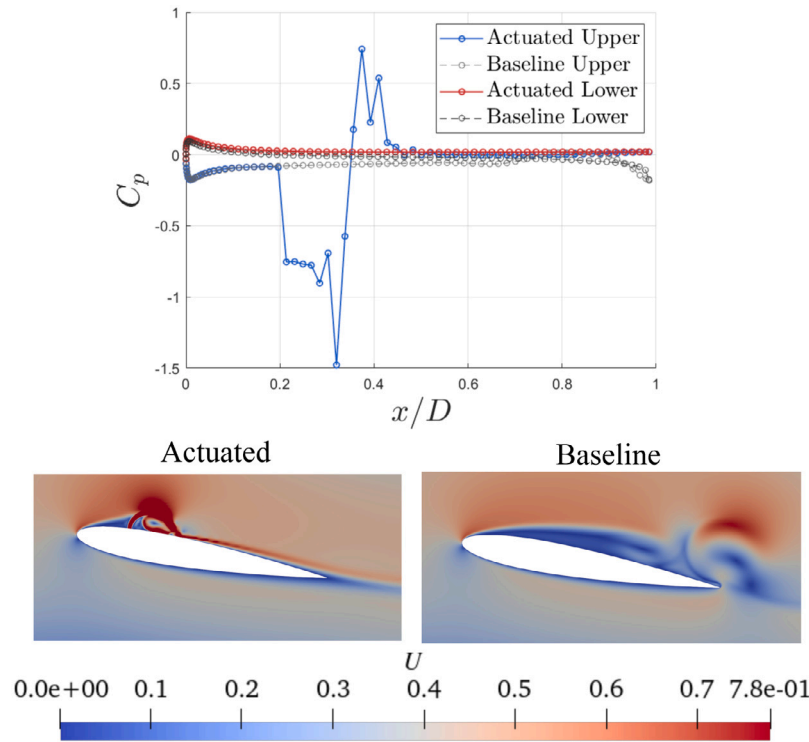


Fig. 21. Comparison of C_p along the airfoil for the baseline (right) and actuated (left) cases. Velocity contours are displayed for the instants used to extract the C_p distribution.

velocity variations and shear-layer roll-up, minimizing separation and eliminating abrupt velocity gradients (flow discontinuities) that the uncontrolled flow would otherwise exhibit as it converges toward the leading edge.

As a result, the flow acceleration leads to a substantial increase in the pressure differential between the front and rear sections of the airfoil, generating a significant lift enhancement. Simultaneously, the controlled deceleration downstream prevents excessive drag penalties, optimizing overall aerodynamic efficiency.

The specific scenario where the third jet is the primary blowing jet, while the other two jets counteract its effect with equal intensity, provides insight into why this strategy avoids discontinuities at high mass flow rates. Unlike the fixed-intensity cases analyzed in Section 3.1, where abrupt variations in flow properties were observed, the present DRL-based strategy achieves a smooth transition and maintains stability even under aggressive actuation conditions.

Fig. 21 illustrates the pressure coefficient distribution along the airfoil, highlighting the impact of jet actuation on the flow field. The results indicate that the jets accelerate the flow between positions 0.2 and 0.3, followed by a reduction in velocity and a corresponding increase in pressure. However, this pressure increase is short-lived, and the flow quickly stabilizes without discontinuities. At the leading edge, both the upper and lower surfaces converge with nearly identical pressures, ensuring a smooth wake formation. This behavior is facilitated by the presence of two distinct layers on the upper surface: a thin low-velocity layer adjacent to the surface and a higher-velocity layer above it. As a result, the trailing edge encounter remains smooth, with the thicker layer slightly separating from the surface. However, the increased velocity at the airfoil's trailing edge relative to the inflow parabolic profile leads to an associated increase in pressure drag.

The frequency spectrum confirms the absence of oscillations or dynamic behavior in the wake, reinforcing the steady nature of the flow. This can also be observed through the Reynolds stresses, as depicted in Fig. 22.

Fig. 22a reveals how the streamwise Reynolds stresses capture the two layers on the upper surface. The peak of the thicker layer begins after the jet region and extends into the wake. These peaks do not indicate a high-turbulence or separation region but rather an amplification of turbulence energy caused by external disturbances, in this case, the jets.

Fig. 22b further illustrates the influence of jet actuation on the flow, as indicated by increased momentum transfer in the wall-normal direction. The previously mentioned flow layers are also visible here. When comparing the baseline separation region around $X = 0.8$ with the steady actuated case, the wall-normal turbulence is significantly reduced and controlled. Additionally, at the trailing edge, no recirculation bubble is observed, as there are no signs of separation or discontinuities. This is further supported by the minimal stress levels at $X = 1$, where both upper and lower surface flows merge smoothly. In the wake region, turbulence mixing in the wall-normal direction is absent at $X = 2.5$, indicating the suppression of eddies.

Fig. 22c depicts the Reynolds shear stresses, which highlight the formation of the recirculation bubble induced by the jets. The presence of this bubble is particularly noticeable at $X = 0.4$, where substantial turbulence mixing and momentum transfer between the streamwise and wall-normal directions occur. The negative sign of the Reynolds shear stress suggests that this mixing is directed towards the wall, a phenomenon also observed in Fig. 21. Finally, at the wake profiles, no discernible eddies or discontinuities are present, indicating a significantly weaker wake compared to the baseline scenario. This can also be confirmed by examining Fig. 23, where the wake is notably smoother compared to both the non-actuated and drag reduction cases.

3.2.3. Comparison of the DRL results with periodic action strategies

Various periodic-control jet flow strategies have been tested and compared against the optimized DRL-based strategies. Among these, the best-performing periodic case has been selected for direct comparison.

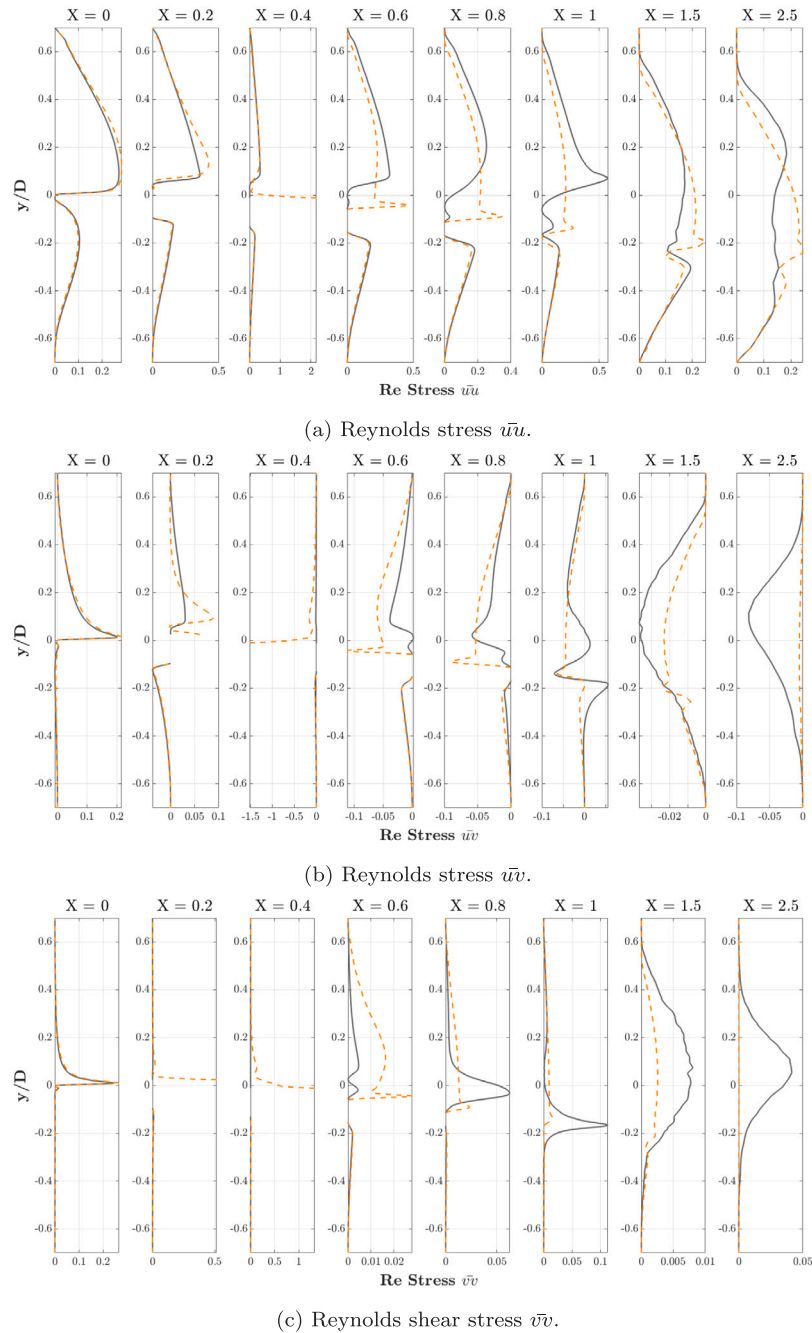


Fig. 22. Comparison of Reynolds stresses ($\bar{u}u$, $\bar{u}v$, and $\bar{v}v$) along the Y -coordinate at different streamwise locations (x/D). The dashed orange lines represent the actuated case, while the solid gray lines correspond to the baseline flow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The results indicate that the optimal periodic strategy occurs when jet_1 and jet_2 operate in phase, while jet_3 is phase-shifted by 180° . This strategy closely resembles the DRL-optimized strategy, as illustrated in Fig. 24, with the key distinction being that, in this case, the control actions follow a strictly periodic modulation at a frequency of 0.54 Hz—three times the vortex shedding frequency. Among all tested periodic strategies, this approach has proven to be the most effective for drag reduction, achieving a 42% decrease in drag. However, the DRL-optimized strategy outperforms it, yielding a 43.9% reduction.

Regarding lift reduction, the periodic strategy results in a 32% decrease, whereas the DRL strategy achieves a slightly greater reduction of 35%.

A comprehensive summary of all control strategies — including the open-loop fixed case of Wang et al. (2022), the periodic forcing benchmark, the best fixed-intensity cases (cases 1 and 8), and our DRL-based policies — is presented in Table 5. This table collates the percentage changes in drag (ΔC_d), lift (ΔC_l), and aerodynamic efficiency ($\Delta(C_l/C_d)$) relative to the uncontrolled baseline, facilitating a direct performance comparison.

It is important to note that the periodic strategy is significantly less energy-efficient, requiring 19.8% higher mass flow rate to achieve slightly inferior drag reduction. By contrast, our DRL-based control not only matches (and in some metrics exceeds) the performance of periodic forcing but does so with substantially lower actuation effort, translating into dramatic energy savings. Moreover, none of

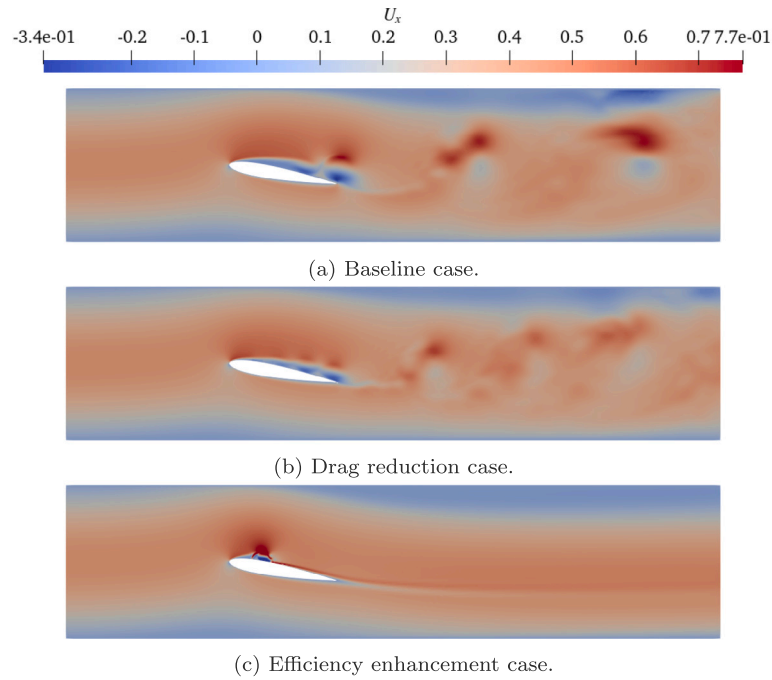


Fig. 23. Comparison of the flow field in terms of streamwise velocity (U_x) for three different cases: (a) Baseline, showing natural flow behavior without actuation; (b) Drag reduction case, demonstrating the effects of DRL-based active flow control in minimizing drag, evidenced by reduced wake size and vortex shedding; and (c) Efficiency enhancement case, where actuation optimizes the flow for improved aerodynamic performance, resulting in a streamlined wake and enhanced boundary layer attachment. The color scale represents velocity magnitude in the x -direction.

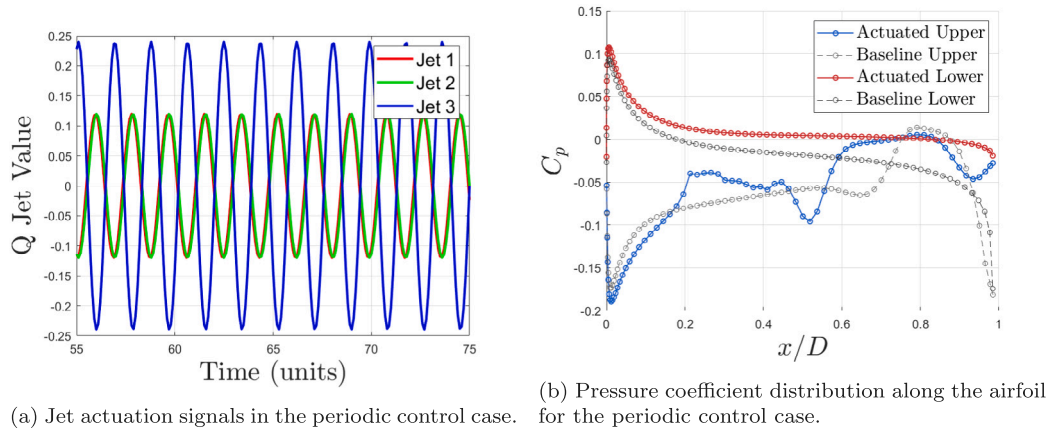


Fig. 24. Comparison of the pressure coefficient (C_p) along the airfoil for the baseline and periodic actuation cases.

the periodic strategies tested in prior work deliver efficiency gains comparable to those uncovered by our learned policy, underscoring the transformative potential of DRL for aerodynamic optimization.

4. Conclusions

This study has demonstrated the potential of deep reinforcement learning (DRL) for active flow control (AFC) in a two-dimensional NACA 0012 airfoil at a Reynolds number of 3000. The implementation of a DRL agent, capable of dynamic decision-making in real-time simulations, underscores its effectiveness in identifying optimal AFC strategies for both drag reduction and aerodynamic efficiency enhancement.

Initially, the DRL approach was designed to achieve significant aerodynamic drag reduction while incorporating additional regularization to maintain lift performance. This resulted in a 43.4% decrease in drag; however, it also led to a 35.9% reduction in lift, highlighting the

Table 5

Percentage changes in drag (ΔC_d), lift (ΔC_l), and aerodynamic efficiency ($\Delta(C_l/C_d)$) relative to the uncontrolled baseline ($C_d = 0.376$, $C_l = 1.822$, $C_l/C_d = 4.846$) for various control strategies on a NACA0012 airfoil at $\alpha \approx 10^\circ$, $Re = 3000$.

Control case	ΔC_d	ΔC_l	$\Delta(C_l/C_d)$
Wang et al. (2022)	-27.0%	+27.7%	+11.5%
Periodic control	-42.0%	-32.0%	+17.2%
DRL – drag focus	-43.4%	-35.9%	+13.4%
DRL – efficiency focus	-30.6%	+10.0%	+58.6%
Fixed-intensity (case 1 – best drag)	-41.7%	-24.7%	+29.3%
Fixed-intensity (case 8 – best eff.)	-19.4%	+9.7%	+36.3%

inherent trade-off when prioritizing drag minimization. This outcome was primarily achieved through the suppression of vortex shedding and the mitigation of the recirculation bubble at the trailing edge. The DRL agent successfully identified dynamic strategies to weaken wake dynamics, resulting in a smoother downstream flow. The observed

drag reduction highlights the adaptability and efficiency of reinforcement learning-based AFC. To achieve this reduction, the stabilization of the wake, characterized by decreased vortex shedding amplitudes, contributed to a reduction in pressure drag, while boundary layer modifications played a role in decreasing skin friction drag.

Conversely, when the focus shifted toward aerodynamic efficiency, the DRL agent exhibited its ability to optimally balance lift and drag, ultimately achieving a 58.6% improvement in aerodynamic efficiency. Unlike the drag-reduction strategy, the efficiency-enhancement approach relied on higher jet intensities. The DRL agent leveraged this increased actuation to dynamically transition the flow from an unactuated vortex-shedding regime to a controlled, steady-state configuration. This efficiency improvement was driven by both an increase in lift and a reduction in drag, demonstrating that both performance metrics can be simultaneously optimized. The underlying physics involved the formation of a controlled closed recirculation bubble, which facilitated flow acceleration while simultaneously stabilizing the flow by smoothing the effective geometry encountered by the freestream.

When compared to fixed-intensity jet strategies, the DRL approach exhibited superior adaptability. While fixed policies were effective to some extent, they lacked the responsiveness necessary to optimize performance under dynamic conditions. In contrast, DRL-based strategies dynamically modulated actuation intensities based on real-time feedback, ensuring sustained aerodynamic improvements. Even in the efficiency-enhancement case, where jet actuation eventually stabilized at constant intensities, the DRL approach optimized the transition phase to minimize the time required to reach a steady-state flow. Furthermore, in the best-performing drag reduction cases, the DRL strategy achieved an additional 6% drag reduction compared to fixed-intensity jets. Similarly, in efficiency-focused cases, DRL improved aerodynamic efficiency by 18% by leveraging all three jets instead of just two, as employed in the fixed-intensity strategy. These findings highlight the superior performance of DRL in optimizing complex AFC systems.

In comparison to periodic-control strategies, the DRL approach achieved both superior performance and efficiency. While the most effective periodic actuation strategy reduced drag by 42%, DRL further improved this to 43.9%, requiring 19.8% less mass flow rate. This highlights DRL's ability to optimize aerodynamic performance while minimizing energy consumption, in contrast to periodic actuation strategies, which lack real-time adaptability to evolving flow conditions.

Despite these notable gains, the study encountered certain limitations and challenges. First, the channel configuration introduced ground effects, which led to augmented lift, potentially influencing direct comparisons with baseline studies. Second, a mass flow rate constraint was imposed to ensure numerical stability in the computational fluid dynamics (CFD) framework, which may have restricted the exploration of more aggressive actuation strategies.

Future research will focus on extending the DRL framework to higher Reynolds numbers and three-dimensional domains. Additionally, hybrid strategies combining DRL with traditional control methodologies will be investigated to enhance robustness. Further studies will also explore the scalability of DRL agents in optimizing flows with multiple objectives and complex constraints, aiming to refine AFC methodologies for practical aerodynamic applications.

CRediT authorship contribution statement

Xavier Garcia: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Arnau Miró:** Writing – review & editing, Validation, Supervision, Software, Methodology, Investigation, Data curation. **Pol Suárez:** Writing – review & editing, Visualization, Software, Methodology, Investigation, Data curation. **Francisco Alcántara-Ávila:** Writing – review & editing, Software, Methodology. **Jean Rabault:** Methodology, Software, Writing – review & editing.

Bernat Font: Writing – review & editing, Software, Methodology. **Oriol Lehmkuhl:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Ricardo Vinuesa:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Ricardo Vinuesa acknowledges the financial support from ERC grant no. 2021-CoG-101043998, DEEPCONTROL. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

During the preparation of this work, the author(s) used Generative AI in order to enhance text clarity, improve formal language, and refine structure. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Data availability

Data will be made available on request.

References

- Amitay, M., Smith, D.R., Glezer, A., 1998. Aerodynamic flow control using synthetic jet technology. In: Proceedings of the 36th AIAA Aerospace Sciences Meeting and Exhibit. AIAA, Reno, Nevada, USA.
- Bechert, D.W., Bartenwerfer, M., 1989. The viscous flow on surfaces with longitudinal ribs. *J. Fluid Mech.* 206, 105–129.
- Berattis, N., Squires, K.D., Balaras, E., 2017. Separation control and drag reduction using roughness elements. In: Proceeding of Tenth International Symposium on Turbulence and Shear Flow Phenomena. Begellhouse, Swissotel Chicago, Chicago, Illinois, U.S.A., pp. 199–204.
- Cavallazzi, G.M., Guastoni, L., Vinuesa, R., Pinelli, A., 2024. Deep reinforcement learning for the management of the wall regeneration cycle in wall-bounded turbulent flows. *Flow, Turbul. Combust.*
- Charnyi, S., Heister, T., Olshanskii, M.A., Rebholz, L.G., 2017. On conservation laws of Navier–Stokes Galerkin discretizations. *J. Comput. Phys.* 337, 289–308.
- Charnyi, S., Heister, T., Olshanskii, M.A., Rebholz, L.G., 2019. Efficient discretizations for the EMAC formulation of the incompressible Navier–Stokes equations. *Appl. Numer. Math.* 141, 220–233.
- Ferziger, J.H., Peric, M., Street, R.L., 2020. *Computational Methods for Fluid Dynamics*, fourth ed. Springer.
- Font, B., Alcántara-Ávila, F., Rabault, J., et al., 2025. Deep reinforcement learning for active flow control in a turbulent separation bubble. *Nat. Commun.* 16, 1422.
- Greenblatt, D., Wygnanski, I.J., 2000. The control of flow separation by periodic excitation. *Prog. Aerosp. Sci.* 36 (7), 487–545.
- Kametani, Y., Fukagata, K., 2011. Direct numerical simulation of spatially developing turbulent boundary layers with uniform blowing or suction. *J. Fluid Mech.* 681, 154–172.
- McLellan, B.W., Ladson, C.L., 1988. A History of Suction-Type Laminar-Flow Control with Emphasis on Flight Research. Technical Report TM-4080, NASA.
- Muddada, S., Patnaik, B.S.V., 2010. An active flow control strategy for the suppression of vortex structures behind a circular cylinder. *Eur. J. Mech. B Fluids* 29 (2), 93–104.
- Örlü, R., Vinuesa, R., 2020. Instantaneous wall-shear-stress measurements: advances and application to near-wall extreme events. *Meas. Sci. Technol.* 31 (11), 112001. Publisher: IOP Publishing.
- Rabault, J., Kuchta, M., Jensen, A., Reglade, U., Cerardi, N., 2019. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *J. Fluid Mech.* 865, 281–302.
- Rabault, J., Kuhnle, A., 2019. Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach. *Phys. Fluids* 31 (9), 094105.
- Ren, F., Rabault, J., Tang, H., 2021. Applying deep reinforcement learning to active flow control in weakly turbulent conditions. *Phys. Fluids* 33, 037121.

- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O., 2017. Proximal policy optimization algorithms.
- Suárez, P., Alcántara-Ávila, F., Miró, A., Rabault, J., Font, B., Lehmkuhl, O., Vinuesa, R., 2024a. Active flow control for drag reduction through multi-agent reinforcement learning on a turbulent cylinder at $Re_D = 3900$. *Flow Turbul. Combust.* Accepted.
- Suárez, P., Alcántara-Ávila, F., Rabault, J., Miró, A., Font, B., Lehmkuhl, O., Vinuesa, R., 2024b. Flow control of three-dimensional cylinders transitioning to turbulence via multi-agent reinforcement learning.
- Swanson, R.C., Langer, S., 2016. Comparison of NACA 0012 Laminar Flow Solutions: Structured and Unstructured Grid Methods. Technical Report NF1676L-21405, NTRS Author Affiliations: NASA Langley Research Center, Deutsches Zentrum fuer Luft- und Raumfahrt e.V. NTRS Document ID: 20160003623 NTRS Research Center: Langley Research Center (LaRC).
- 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. *Phys. Fluids* 32 (5), 053605.
- Trias, F., Lehmkuhl, O., 2011. A self-adaptive strategy for the time integration of Navier-Stokes equations. *Numer. Heat Transfer Part B: Fundamentals*, 116–134.
- Varela, P., Suárez, P., Alcántara-Ávila, F., Miró, A., Rabault, J., Font, B., García-Cuevas, L.M., Lehmkuhl, O., Vinuesa, R., 2022. Deep reinforcement learning for flow control exploits different physics for increasing Reynolds number regimes. *Actuators* 11 (12).
- Vasanth, J., Rabault, J., Alcántara-Ávila, F., et al., 2024. Multi-agent reinforcement learning for the control of three-dimensional Rayleigh–bénard convection. *Flow Turbul. Combust.*
- Vazquez, M., Houzeaux, G., Koric, S., Artigues, A., Aguado-Sierra, J., Aris, R., Mira, D., Calmet, H., Cucchiatti, F., Owen, H., Taha, A., Cela, J.M., 2014. Alya: Towards exascale for engineering simulation codes.
- Vignon, C., Rabault, J., Vasanth, J., Alcántara-Ávila, F., Mortensen, M., Vinuesa, R., 2023a. Effective control of two-dimensional Rayleigh–bénard convection: invariant multi-agent reinforcement learning is all you need. *Phys. Fluids* 35 (6), 065146.
- Vignon, C., Rabault, J., Vinuesa, R., 2023b. Recent advances in applying deep reinforcement learning for flow control: perspectives and future directions. *Phys. Fluids* 35 (3), 031301.
- Vinuesa, R., Schlatter, P., Nagib, H.M., 2015. On minimum aspect ratio for duct flow facilities and the role of side walls in generating secondary flows. *J. Turbul.* 16 (6), 588–606, Publisher: Taylor & Francis.
- Voevodin, A.V., Korniyakov, A.A., Petrov, A.S., Petrov, D.A., Sudakov, G.G., 2019. Improvement of the take-off and landing characteristics of wing using an ejector pump. *Thermophys. Aeromechanics* 26 (1), 9–18.
- Wang, Y.-Z., Mei, Y.-F., Aubry, N., Chen, Z., Wu, P., Wu, W.-T., 2022. Deep reinforcement learning based synthetic jet control on disturbed flow over airfoil. *Phys. Fluids* 34 (3), 033606.
- Yousefi, K., Saleh, R., 2015. Three-dimensional suction flow control and suction jet length optimization of NACA 0012 wing. *Meccanica* 50 (6), 1481–1494.
- Zong, H., Wu, Y., Liang, H., Su, Z., Li, J., 2024. Experimental study on Q-learning control of airfoil trailing-edge flow separation using plasma synthetic jets. *Phys. Fluids* 36 (1), 015101.