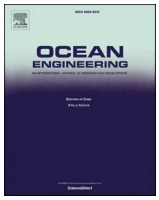
Ocean Engineering 287 (2023) 115775

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng



Active flow control on airfoils by reinforcement learning

Koldo Portal-Porras [a], Unai Fernandez-Gamiz [a,*], Ekaitz Zulueta [b], Roberto Garcia-Fernandez [a,c], Saioa Etxebarria Berrizbeitia [d]

[a] Nuclear Engineering and Fluid Mechanics Department, Faculty of Engineering of Vitoria-Gasteiz, University of the Basque Country, UPV/EHU, Nieves Cano 12, Vitoria-Gasteiz, 01006, Araba, Spain
[b] Automatic Control and System Engineering Department, Faculty of Engineering of Vitoria-Gasteiz, University of the Basque Country, UPV/EHU, Nieves Cano 12, Vitoria-Gasteiz, 01006, Araba, Spain
[c] Sunsundegui S.A., Poligono Ibarrea, s/n, 31800, Altsasu, Navarra, Spain
[d] Mechanical Engineering Department, Faculty of Engineering of Vitoria-Gasteiz, University of the Basque Country, UPV/EHU, Nieves Cano 12, Vitoria-Gasteiz, 01006, Araba, Spain

ARTICLE INFO

Handling Editor: Prof. A.I. Incecik

Keywords:
Artificial neural networks
Deep learning
Reinforcement learning
Computational fluid dynamics
Moving flap

ABSTRACT

Active flow control is a widespread practice for airfoil aerodynamic performance enhancement. Within active flow control, reactive strategies are very effective, but the adequate design of these strategies is often complex. This study proposes a reactive control strategy based on a Reinforcement Learning (RL) agent to effectively govern the motion of a rotating flap implemented on a NACA0012 airfoil. With this objective, first different Computational Fluid Dynamics (CFD) simulations are conducted to gather data about the tested case. Then, a numerical model based on Artificial Neural Networks (ANN) is developed to model the discussed case. Finally, the RL agent is trained and tested under different conditions. The results show that the trained RL agent is able to provide a fast and reliable response for every tested condition, setting the adequate position of the flap and obtaining an appropriate aerodynamic performance of the airfoil for all the tested conditions. In comparison with the optimum conditions, the absolute error in the position of the flap set by the agent is below 2.2 ° for all the angles of attack, resulting in an aerodynamic performance very close to the optimum, being only 0.39%–3.05% lower, depending on the case.

1. Introduction

Wind energy, in combination with other renewable energies, is becoming one of the main sources of energy around the world, gaining more and more importance in the global energy mix every year (IRENA, 2023). The reason behind this is that wind energy is a clean energy source without fossil fuel consumption and without greenhouse gas emissions during its operation. Nevertheless, in order to compete against traditional energy sources in terms of energy production and associated costs, an improvement of turbine performance is still required.

A widespread practice for wind turbine performance enhancement is the implementation of flow control devices. Flow control devices can be divided into two main groups, passive devices, which remain motionless, and therefore, external energy is not required for their operation; and active devices, which move depending on the established operating conditions, and hence, require an external energy source. Aramendia et al. (Aramendia-Iradi et al., 2016; Aramendia et al., 2017) provided an extensive review of both active and passive flow control devices, focusing on their application on wind turbine blades.

Several authors have used different active and passive flow control devices techniques in airfoils. The addition of passive flow control devices is a widespread practice, as these devices can significantly improve aerodynamic performance in a simple way and without requiring an external power source. For example, Fernandez-Gamiz et al. (2017) added microtabs on the Trailing Edge (TE) of a DU91W(2)250 airfoil, Aramendia et al. (2019) added Gurney flaps on the same airfoil, and Tunio et al. (2020) implemented a spanwise waviness on a NACA0021 airfoil. Due to the complexity of their design and the need of an external power source, active devices are less popular than passive devices. However, there are many studies in which active devices are used, as the achievable aerodynamic improvements are much higher. For example, Anzalotta et al. (2020) used plasma actuators for controlling the flow

* Corresponding author.
E-mail addresses: koldo.portal@ehu.eus (K. Portal-Porras), unai.fernandez@ehu.eus (U. Fernandez-Gamiz), ekaitz.zulueta@ehu.eus (E. Zulueta), rgarcia@sunsundegui.com (R. Garcia-Fernandez), saioa.etxebarriab@ehu.eus (S. Etxebarria Berrizbeitia).

https://doi.org/10.1016/j.oceaneng.2023.115775
Received 1 June 2023; Received in revised form 30 August 2023; Accepted 3 September 2023
Available online 6 September 2023

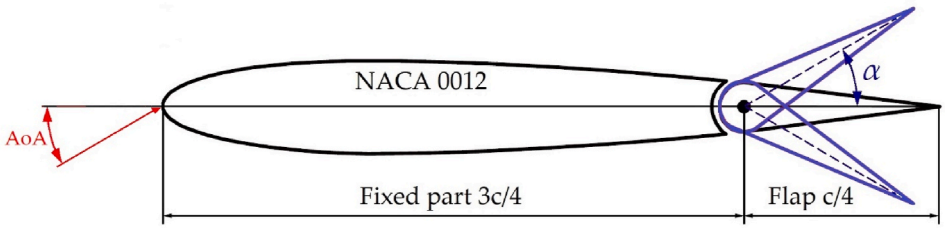


Fig. 1. Tested case.

through the tip gap of a NACA0065 airfoil, Julian et al. (2022) proposed a co-flow jet for improving the aerodynamic performance of a NACA0015 airfoil, Spens et al. (2023) added curved fluidic oscillators on the leading edge of a NACA0018 airfoil and Anılır et al. (Anılır et al.) used TE flaps on a NACA0012 airfoil, with a predetermined oscillation.

Among active techniques, two main strategies exist, depending on the behavior of the flow control device. When the behavior of the flow control device is previously defined, the device is considered to be predetermined. On opposite, when the behavior of the flow control device depends on the instantaneous flow state, the device is considered to be reactive. Predetermined strategies are very common, due to their simplicity in their design and implementation. However, due to their adaptation to the flow state, reactive strategies provide a greater aerodynamic improvement capability, see Gad-el- Hak (1996). Nevertheless, there are cases in which the development of reactive strategies may be overly complex, such as in systems affected by many parameters or where data acquisition is very complex.

In last years, with the growth in the knowledge in Artificial Intelligence and the increase in the capacity of computers, data-driven methods are becoming increasingly popular for fluid dynamics problem solving. This kind of methods provide a very fast and accurate solution of the analyzed case, avoiding the problems related with traditional experimental tests and Computational Fluid Dynamics (CFD) simulations, such as the influence of the user for setting up the case.

Many authors have implemented data-driven methods, mostly Deep Learning (DL) methods, for fluid dynamics problem solving. In areas related with wind turbines and airfoils, Sekar et al. (2019) proposed a two-step-consistent approach based on a Multi-Layer Perceptron (MLP) and a Convolutional Neural Network (CNN) for flow field prediction over an airfoil. Thuerey et al. (2020) approximated the velocity and pressure fields obtained by Reynolds-Averaged Navier-Stokes (RANS)--based Spalart and Allmaras (1992) turbulence model around different airfoils shapes. Zhang et al. (2018) and Chen et al. (2020) used CNN architectures for aerodynamic coefficient prediction of airfoils; and Hui et al. (2020) used a CNN architecture for pressure distribution prediction of airfoils. Kim and Yoon (2022) improved the aerodynamic performance of different NACA airfoils by a geometric modification performed by a CNN.

Other authors focused their efforts on flow control devices. For example, Portal-Porras et al. (2022) developed a CNN for velocity and pressure field prediction around airfoils with different flow control devices. Aramendia et al. (2019) and Rodriguez-Eguia et al. (2020) used Artificial Neural Networks (ANN) to model aerodynamic coefficients of airfoils with Gurney flaps and TE flaps, considering different geometrical modifications on those devices.

As reviewed by Vinuesa et al. (2022) and Ren et al. (2020), there are several data-driven methods for flow control device optimization, such as, Genetic Programming (GP) (Koza, 1992) or Bayesian regression based on Gaussian processes (Rasmussen et al., 2004). Among data-driven methods, Deep Reinforcement Learning (DRL) is a very promising option for control strategy formulation because of the two main reasons stated by Vinuesa et al. (2022). The first reason is that with DRL methods assumptions of the properties of the system are not made; and the second reason is that it uses Neural Networks for the decision making, which are very accurate for representing complex nonlinear functions. Although these techniques are very effective, they have certain limitations, such as the complexity of modeling the environment, or the time and computational resources required for their training.

In the pioneer work conducted by Rabault et al. (2019) the vortex shedding behind a cylinder at moderate Reynolds number (Re), $Re = 100$, is reduced by two jets located on top and bottom sides of the cylinder, which are controlled through DRL. Following that study, Ren et al. (2021) used DRL methods to control the vortex shedding on the same case at turbulent flow regime, $Re = 1000$. In both studies a remarkable drag reduction is achieved, around 8% and 30%, respectively. Fan et al. (2020) used DRL to control the rotation-rate of two small rotational cylinders located downstream and parallel to a cylinder, maximizing the power gain efficiency by selecting the proper rotational speed; and Han et al. (2022) defined a control strategy for vortex shedding suppression on the wake behind a rotational cylinder, whose rotational speed is controlled through DRL, achieving a significant reduction of the drag. Other authors have used these tools in airfoils and wind turbines. For example, Wang et al. (2023) optimized the geometrical parameters of the blade in a wind turbine by two DRL models.

In this study, a DRL-based reactive strategy is proposed for controlling a rotating flap on a NACA0012 airfoil. With this objective, firstly various CFD simulations are conducted to gather data on the evaluated case under different conditions. Next, ANNs are employed to model the problem and obtain fast and reliable results of drag (C_D) and lift (C_L) aerodynamic coefficients. Finally, a DRL agent is trained to develop a control strategy that determines the optimal angle of the flap for each Angle of Attack (AoA) of the flow.

The following of the manuscript is divided as follows: Section 2 provides a detailed explanation of the tested case, the methodology followed to set up and conduct the CFD simulations and the design and training methodology of the ANN-based model and the DRL agent; Section 3 shows the obtained results, testing the DRL agent with different AoA signals; and Section 4, summarizes the main findings of this study.

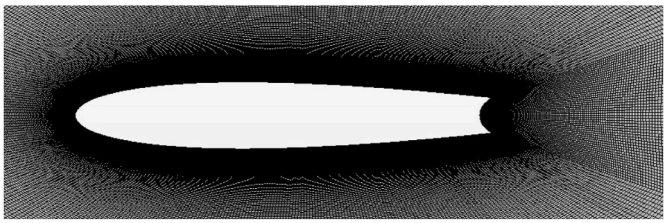
2. Methodology

2.1. Case

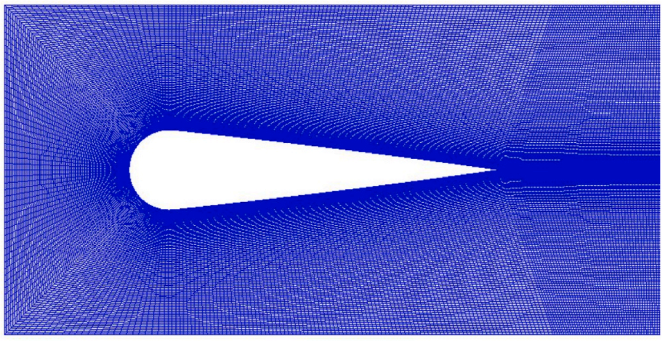
In this study a rotating active flap on a NACA0012 airfoil is analyzed. The front part of the airfoil is defined as the fixed part, which remains motionless; and the back side of the airfoil is defined as the rotating flap, which may rotate with a defined angular velocity (ω). The angle of this flap (α) can change between $\alpha = -30°$ and $= 30°$. The chord length (c) of the studied airfoil is equal to 1 m, and the rotation axis of the flap is located at $3c/4$ from the Leading Edge (LE) of the airfoil. The flow hits the airfoil with an Angle of Attack (AoA) between $AoA = -10°$ and $= 10°$. Fig. 1 provides a schematic view of the tested case.

α is defined as positive when the flap is above the axis of the airfoil, and as negative when the flap is below the axis of the airfoil; and ω is defined as positive when the flap rotates counterclockwise and as negative when the flap rotates clockwise.

The aerodynamic performance of the airfoil is directly related with

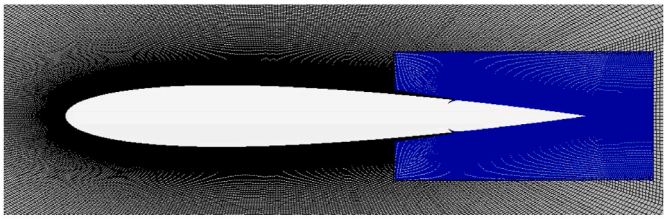


(a)

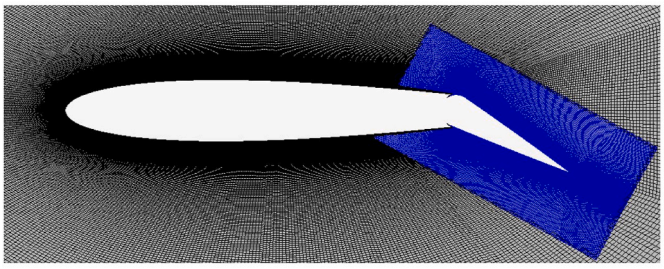


(b)

Fig. 2. Generated structured meshes. (**a**) Background Region; (**b**) Overset Region.



(a)



(b)

Fig. 3. Overset mesh. (**a**) $\alpha = 0^{\circ}$; (**b**) $\alpha = -30^{\circ}$.

the position of the flap and the *AoA* of the flow. Therefore, this paper aims to train a Reinforcement Learning (RL) agent that controls the flap depending on the *AoA* of the flow, in order to optimize the aerodynamic performance of the airfoil.

2.2. CFD setup

Different CFD simulations of the tested case are conducted, in order to obtain data for training the RL agent. For running the simulations StarCCM + v2019.1 commercial code was used.

As the analyzed case requires modeling the motion of the flap, the Overset Mesh technique has been considered for the meshing of the simulations. This technique consists of preparing two different regions, one with the stationary geometry (Background Region, fixed part of the airfoil) and another with the moving geometry (Overset Region, the flap). The Overset Mesh technique has been widely used for motion modelling on airfoils. Boudis et al. (2019) simulated a flapping airfoil by means of the Overset Mesh, Hu et al. (2021, 2022) successfully applied this meshing technique for active TE flap modelling and Chandar et al. (2018) proved the accuracy of the Overset Mesh on airfoils, showing good agreements of the results with experimental data.

For the Background Region an O-shaped domain is designed, with a radius equal to $R = 16{\cdot}c$. In this region, the outer contour is set as free stream, and no-slip wall conditions are assigned to the fixed part of the airfoil. A rectangular domain is designed with the Overset Region. In this region, an Overset interface is defined with the outer sides, in order to couple both regions. No-slip wall conditions are also assigned to the flap.

With these regions two different structured meshes are generated. The mesh of the Background Region consists of around 360000 cells and the mesh of the Overset Region contains around 80000 cells. Fig. 2 shows the generated meshes.

Once the meshes are generated, both regions are coupled by means of the Overset Mesh interface. Finally, a suitable mesh is generated, which combines both meshes. This mesh is updated in each time step, depending on the instantaneous position of the Overset Region. Fig. 3 provides an example of the mesh when $\alpha = 0^{\circ}$ and $= -30^{\circ}$.

Regarding the fluid physics, a constant free stream velocity of $u_{\infty} = 30\,m/s$ was set for all the cases. The air density was set at $\rho = 1.18415\,kg/m^3$, and the dynamic viscosity at $\mu = 1.85508{\cdot}10^{-5}\,Pa{\cdot}s$. Therefore, the Reynolds number (Re) for the tested cases is around $1.9{\cdot}10^6$, according to Equation (1).

$$Re = \frac{u_{\infty}{\cdot}c{\cdot}\rho}{\mu} \tag{1}$$

For modelling the turbulence, RANS-based k-ω Shear Stress Transport (SST) model introduced by Menter (1994) is selected. This model combines the k-ω model for the near-wall zones and k-ε for the regions far from walls.

Unsteady-state RANS (URANS) equations comprise the continuity (Equation (2)) and momentum (Equation (3)) equations.

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho U_i)}{\partial x_i} = 0 \tag{2}$$

$$\frac{\partial (\rho U_i)}{\partial t} + \frac{\partial \left(\rho U_i U_j\right)}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial \left(\tau_{ij} - \rho \overline{U_i U_j}\right)}{\partial x} \tag{3}$$

where the viscous tensor (τ_{ij}) is defined as detailed is Equation (4).

$$\tau_{ij} = \mu \left(\frac{\partial U_j}{\partial x_i} + \frac{\partial U_i}{\partial x_j} - \frac{2}{3}\delta_{ij}\frac{\partial U_i}{\partial x_i} \right) \tag{4}$$

Transport equation for turbulent kinetic energy (k) and specific dissipation rate (ω) are defined as shown in Equations (5) and (6), respectively.

$$\frac{\partial (\rho k)}{\partial t} + \frac{\partial (\rho \overline{u_i} k)}{\partial x_i} = P_k + D_k + \frac{\partial}{\partial x_i}\left[(\mu + \sigma_k \mu_t) \frac{\partial k}{\partial x_i} \right] \tag{5}$$

$$\frac{\partial (\rho \omega)}{\partial t} + \frac{\partial (\rho \overline{u_i} \omega)}{\partial x_i} = P_{\omega} + D_{\omega} + \frac{\partial}{\partial x_i}\left[(\mu + \sigma_{\omega} \mu_t) \frac{\partial \omega}{\partial x_i} \right] \tag{6}$$

where the production terms (P_k and P_{ω}) are defined in Equations (7) and (8), respectively, and destruction terms (D_k and D_{ω}) are defined in Equations (9) and (10), respectively.

$$P_k = \mu_t \overline{S_{ij}}^2 - \frac{2}{3}\rho k \frac{\partial (\overline{u_i})}{\partial x_i} - \frac{2}{3}\mu_t \left(\frac{\partial (\overline{u_i})}{\partial x_i} \right)^2 \tag{7}$$

$$P_{\omega} = \rho \gamma \overline{S_{ij}}^2 - \frac{2}{3}\rho \gamma \omega \frac{\partial (\overline{u_i})}{\partial x_i} - \frac{2}{3}\rho \gamma \left(\frac{\partial (\overline{u_i})}{\partial x_i} \right)^2 \tag{8}$$

Table 1
Mesh study for C_D.

AoA [°]	DRL Agent			Experimental	Richardson Extrapolation		
	Coarse	Medium	Fine		R	p	RE
−10	0.019	0.0177	0.0172	0.0182	0.3846	1.3785	0.0169
−8	0.0176	0.0161	0.0155	0.015	0.4	1.3219	0.0151
−6	0.015	0.0136	0.0127	0.0115	0.6429	0.6374	0.0111
−4	0.0119	0.0102	0.0096	0.009	0.3529	1.5025	0.0093
−2	0.0104	0.0087	0.0077	0.007	0.5882	0.7655	0.0063
0	0.0094	0.0079	0.007	0.006	0.6	0.737	0.0057
2	0.009	0.0076	0.0067	0.0058	0.6429	0.6374	0.0051
4	0.0097	0.0081	0.0072	0.0065	0.5625	0.8301	0.006
6	0.0113	0.0092	0.0088	0.008	0.1905	2.3923	0.0087
8	0.0132	0.0121	0.0116	0.011	0.4545	1.1375	0.0112
10	0.0151	0.0141	0.0138	0.0132	0.3	1.737	0.0137

Table 2
Mesh study for C_L.

AoA [°]	DRL Agent			Experimental	Richardson Extrapolation		
	Coarse	Medium	Fine		R	p	RE
−10	−0.931	−0.983	−1.005	−1.01	0.4231	1.241	−1.0211
−8	−0.919	−0.895	−0.885	−0.85	0.4167	1.263	−0.8779
−6	−0.878	−0.768	−0.701	−0.62	0.6091	0.7153	−0.5966
−4	−0.654	−0.541	−0.476	−0.41	0.5752	0.7978	−0.388
−2	−0.302	−0.251	−0.224	−0.21	0.5294	0.9175	−0.1936
0	0.049	0.018	0.005	0	0.4194	1.2538	−0.0044
2	0.487	0.338	0.271	0.22	0.4497	1.1531	0.2163
4	0.768	0.624	0.533	0.46	0.6319	0.6621	0.3768
6	0.929	0.81	0.73	0.65	0.6897	0.5361	0.5522
8	0.993	0.934	0.916	0.9	0.3051	1.7127	0.9081
10	1.077	1.048	1.03	1.05	0.6207	0.6881	1.0005

$$D_k = -\rho\beta^* k\omega \tag{9}$$

$$D_\omega = -\rho\beta\omega^2 \tag{10}$$

where the term $\overline{S_{ij}}$ is defined in Equation (11), and σ_k, σ_ω, β and β^* are model coefficients.

$$\overline{S_{ij}} = \frac{1}{2}\left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i}\right) \tag{11}$$

Each case was simulated for 10 s, with a time step of 10^{-3} s, which is considered to be small enough for obtaining a good residual convergence and properly capturing the movement of the flap and the changes in lift and drag coefficients. Pressure-velocity coupling was performed by UpWind algorithm, and mesh discretization by a linear upwind second order scheme.

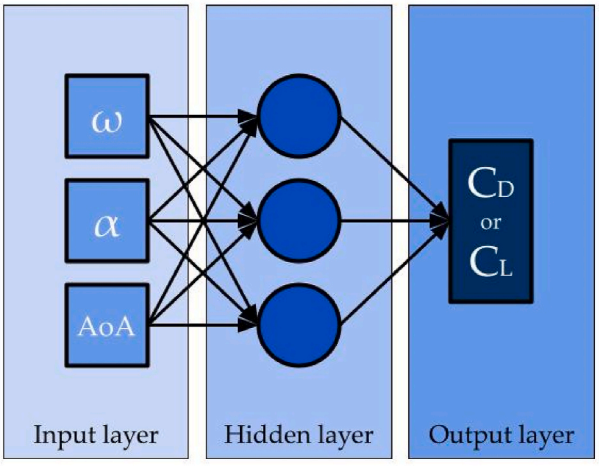


Fig. 4. Sketch of the designed ANN.

Ten different CFD simulations were conducted with the purpose of obtaining a diverse dataset. In all of them the *AoA* changes according to Equation (12). Regarding the flap, in 7 simulations the flap is stationary throughout the whole simulation, at $\alpha = -30°, -20°, -10°, 0°, 10°, 20°, and\ 30°$. In the other 3 simulations, the flap oscillates with a constant frequency (f) of 1 Hz, 0.5 Hz and 0.2 Hz, which is equivalent to a reduced frequency (k) of 0.1047, 0.0524 and 0.0209, respectively. Equation (13) specifies the instantaneous flap angle for cases with flap oscillation.

$$AoA(t) = 10 \cdot \sin(0.2\pi t) \tag{12}$$

$$\alpha(t) = \frac{\pi}{6} \cdot \sin(2\pi f t) \tag{13}$$

In order to verify sufficient mesh resolution, the Richardson Extrapolation method (Richardson and Gaunt, 1927) has been applied to C_D and C_L with $\alpha = 0°$. This method estimates the value of a parameter when the cell amount tends to infinity. For using this method, a minimum of three meshes are required. In this case, a coarse mesh (90000 cells in the Background Region and 20000 cells in the Overset Region), a medium mesh (180000 and 40000 cells) and a fine mesh (the explained mesh, 360000 and 80000 cells) are used, which means that the refinement between meshes is equal to 2. Additionally, in order to validate the simulations, the obtained results are compared with the experimental data from Sheldahl and Klimas (1981). In that study the NACA0012 airfoil is tested under different *AoA* and a Reynolds number $Re = 1.76 \cdot 10^6$. For making the comparison an upstroke cycle is considered, with the *AoA* going from $AoA = -10°$ to $AoA = 10°$. Table 1 shows the results of the mesh dependency study for C_D and Table 2 for C_L.

As detailed in Tables 1 and 2, the convergence condition (R), which should be between 0 and 1 to ensure monotonic convergence, is fulfilled for all the cases. Additionally, the results of the aerodynamic coefficients obtained with the fine mesh show very similar values in comparison

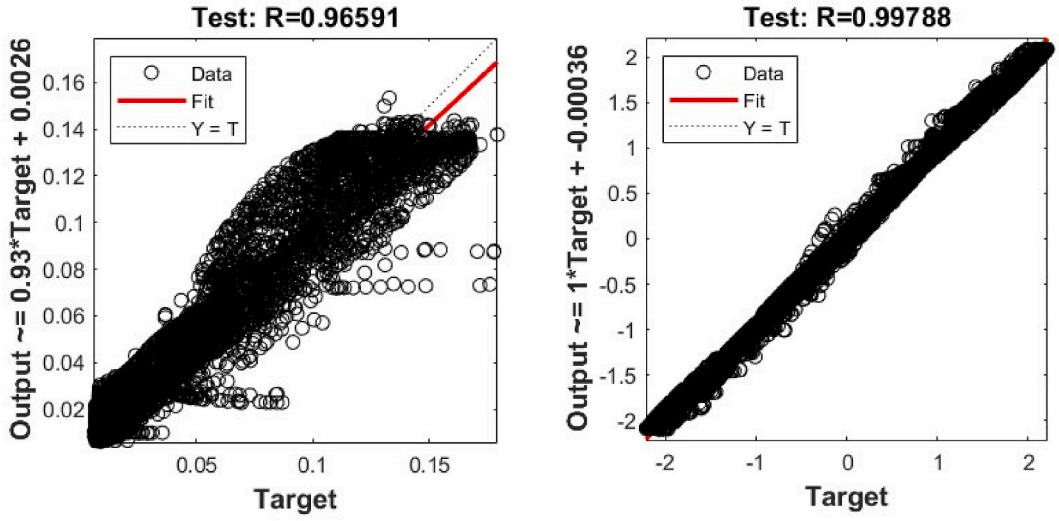


Fig. 5. ANN testing correlation coefficients. The vertical axis shows the predictions of the ANN and the horizontal axis the CFD data. (a) C_D; (b) C_L.

with experimental data, following a nearly equal trend. Therefore, the simulations are considered to be suitable for the tested case.

2.3. Case modelling

A numerical model of the tested case is created, in order to simplify and speed up the training process of the DRL agent. This numerical model consists of two ANN, one for the prediction of each aerodynamic coefficient. Due to their properties and advantages, this networks are exceptional tools for modelling a wide variety of systems, see Lopez--Guede et al. (2016; 2017)) and Ugarte-Anero et al. (2022). For designing and training the ANNs MATLAB 2023a (MATLAB) and its Deep Learning Toolbox (Deep Learning Toolbox Available online) were used.

The designed networks have three inputs, ω, α and AoA; and a single output, either C_D or C_L, depending on the network. Between input and output layers, the network contains a hidden layer. This network represents the typical configuration of Multilayer Perceptron with Back-propagation (BP-MLP). Fig. 4 provides a schematic view of the designed network.

The aerodynamic coefficient (C, either C_D or C_L), is calculated by Equation (14), and the hidden layers follow a sigmoid function, defined in Equation (15). The postsynaptic value h_i of each neuron i is calculated using the linear combination defined in Equation (16).

$$C = \sum_{i=1}^{i=N_{hidden}} \omega_i \cdot g_i(\vec{x}) + \theta_i \tag{14}$$

$$g_i(\vec{x}) = \frac{1}{1+e^{-h_i}} \tag{15}$$

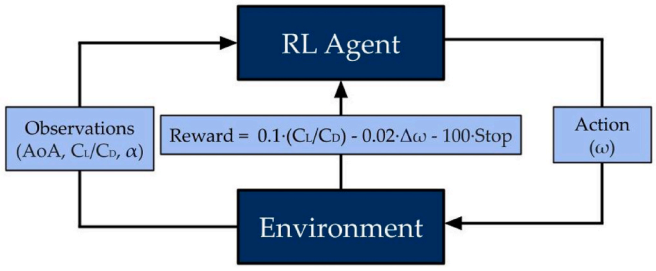


Fig. 6. Sketch of how RL works.

$$h_i(\vec{x}) = \sum_{j=1}^{j=N_{inputs}} \omega_{ij}^{'} \cdot x_j + \theta_j^{'} \tag{16}$$

where ω_i and ω_{ij} represent the weights of the output layer and input hidden layer, respectively; and θ the weighted biases.

For training the network, data has been split into 70% training, 20% validation and 10% testing. Matrices (17–20) contain the values of the input layer weights (ω_{ij}), output layer weights (ω_i) and weighted biases (θ_j and θ_i).

$$\omega_{ij}(C_D) = \begin{bmatrix} -20.4753 & 287.1933 & -0.0665 \\ -0.7966 & -2.1957 & 0.7091 \\ -0.8332 & -1.8335 & 0.7383 \end{bmatrix}$$
$$\omega_{ij}(C_L) = \begin{bmatrix} 0.1982 & -0.2339 & -0.4057 \\ 0.2067 & -0.2655 & -0.4024 \\ 1.2235 & -3.4269 & 1.254 \end{bmatrix} \tag{17}$$

$$\omega_i(C_D) = [-42.7014 \quad -1.0479 \quad 1.1116]$$
$$\omega_i(C_L) = [-20.0648 \quad 19.3885 \quad 0.1671] \tag{18}$$

$$\theta_j(C_D) = \begin{bmatrix} 290.6186 \\ 2.4582 \\ -2.3067 \end{bmatrix} \quad \theta_j(C_L) = \begin{bmatrix} 0.0059 \\ 0.0055 \\ 0.0083 \end{bmatrix} \tag{19}$$

$$\theta_i(C_D) = [43.8302] \quad \theta_i(C_L) = [0.0085] \tag{20}$$

To evaluate the accuracy of the networks, the correlation coefficient (R-value) is analyzed, which quantifies the relation between the ground-truth values and the predicted values. Hence, to ensure the reliability of the networks, this coefficient should be as close to 1 as possible. For this analysis the cases of the test-set are analyzed, shown in Fig. 5. In this case, the R-value is equal to 0.96591 for C_D, and 0.99788 for C_L. Therefore, the predictions of the ANNs are reliable. Additionally, larger errors appear on the peaks, but the trend is precisely followed, which means that the optimum operation conditions are correctly predicted.

2.4. RL agent

Reinforcement Learning (RL) is a Machine Learning (ML) paradigm for decision-making, see Sutton and Barto (2018) and Bertsekas (2012). In RL an agent finds its optimal behavior by exploring an unknown

Table 3
Comparison between the angle set by the DRL agent and the optimum angle for each *AoA*, calculated iteratively.

AoA [°]	DRL Agent		Optimum	
	α [°]	C_L/C_D	α [°]	C_L/C_D
−10	−28.0807	17.4741	−26.3363	17.8491
−8	−24.6085	26.4739	−22.4324	27.3077
−6	−21.188	35.833	−19.0691	36.8597
−4	−17.8591	45.0181	−16.006	45.9831
−2	−14.662	53.5791	−13.1231	54.3643
0	−11.5852	61.2766	−10.3604	61.8739
2	−8.6001	68.0942	−7.5375	68.5471
4	−5.6837	74.1854	−4.7748	74.5488
6	−2.8075	79.802	−2.012	80.1269
8	0.0422	85.2221	0.8709	85.5542
10	2.8591	90.6838	3.6937	91.0655

environment. The optimal behavior is learned by interacting with the environment and receiving observations of how it responds.

Firstly, the agent receives different observations from the environment. In this case, the observations are the instantaneous *AoA* of the air, lift-to-drag ratio (C_L/C_D) and flap angle (α). Depending on those observations, the agent performs an action, in this case, it sets the angular velocity (ω) of the flap. The decision taken by the agent causes different consequences in the environment. In the studied case, the flap velocity changes the flap angle; and therefore, the lift-to-drag ratio. These changes are sent back to the agent by means of the previously mentioned observations. Additionally, the agent gets a reward based on the environment. The agent, throughout the training process, learns the actions that maximize the reward. Fig. 6 provides a sketch of how this process works.

In this case, the reward is set according to Equation (21). As the main objective of flow control devices is to enhance the lift-to-drag ratio, the agent gets a positive reward proportional to this value. Nonetheless, in order to avoid sharp movements of the flap, specially under constant conditions, a negative reward is defined, proportional to the angular velocity difference between time steps ($\Delta\omega$). Since the flap should only move between $\alpha = -30°$ and $\alpha = 30°$, if the agent moves the flap out of this range, he receives a penalty (*Stop*, which is a boolean value that can only be equal to 0 or 1, depending on whether the flap is inside the mentioned range or not, respectively).

$$Reward = 0.1 \cdot \frac{C_L}{C_D} - 0.02 \cdot \Delta\omega - 100 \cdot Stop \tag{21}$$

The agent is designed and trained with MATLAB 2023a (MATLAB), Simulink (Simulink) and its Reinforcement Learning Toolbox (Reinforcement Learning Toolbox Available online). For this case a Deep Deterministic Policy Gradient (DDPG) agent (Lillicrap et al., 2019) is selected, since it is considered to be simple and accurate for continuous action spaces. This agent is an actor-critic RL agent that tries to maximize the expected cumulative long-term reward by searching for an optimal policy.

DDPG combines policy gradient methods and Q-learning (Watkins and Dayan, 1992). Q-learning agents consist of a critic that estimates the future reward of each possible decision for the given observations. These estimations are known as the Q-value. Considering the estimations of the critic, the agent selects the action that maximizes the reward. Since Q-learning algorithms consider all the possible actions for each decision, they rely on discrete and small dimension action selection, being infeasible to find the optimal action in continuous spaces, as the one of this study. In DDPG, since it is an actor-critic agent, two models are combined, the actor and the critic. The actor is a policy network that decides the continuous action that should be taken based on the observations; and the critic is a Q-value network, which informs the actor about how good the taken action has been and how it could be improved, taking the state and the decided action as inputs and providing the Q-value as output.

For the training process, a random and constant *AoA* is established. α is also set as random at $t = 0$ *s*, with the objective of training the agent under different conditions. The training is conducted until the average reward converges at a maximum value.

3. Results and discussion

With the objective of evaluating the capacity of the agent to operate under different conditions, various signals are considered for the *AoA*.

3.1. Constant values

Firstly, the agent is tested with constant *AoA* signals, with the objective of evaluating the agent at stationary conditions. For this case, as the flap remains stationary ($\omega = 0$) after reaching the optimum po-

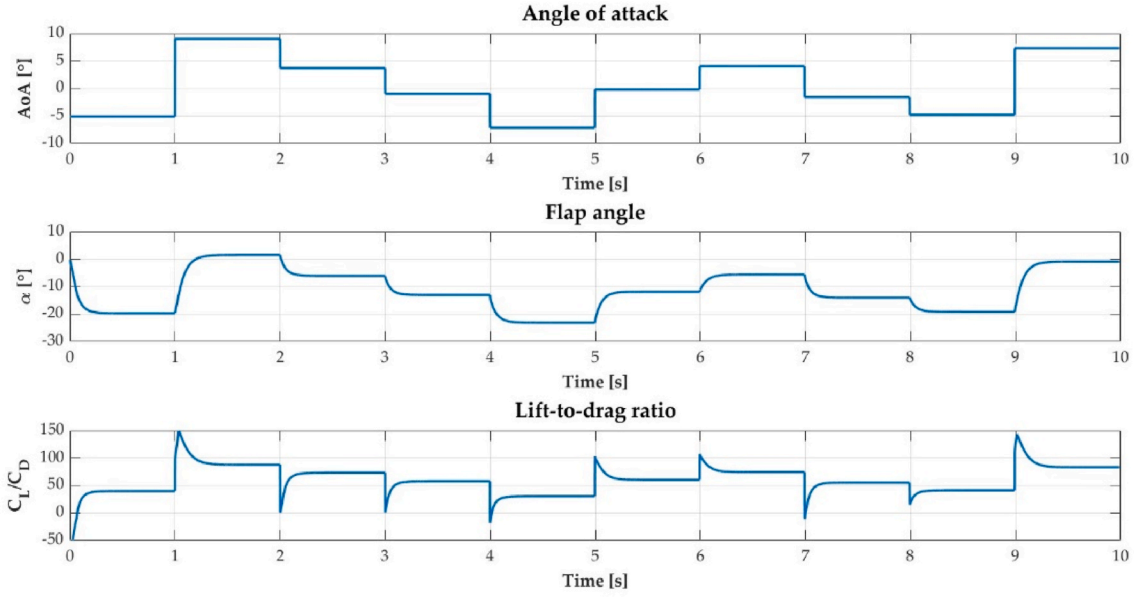


Fig. 7. Flap angle set by the agent for an *AoA* defined by a stair sequence.

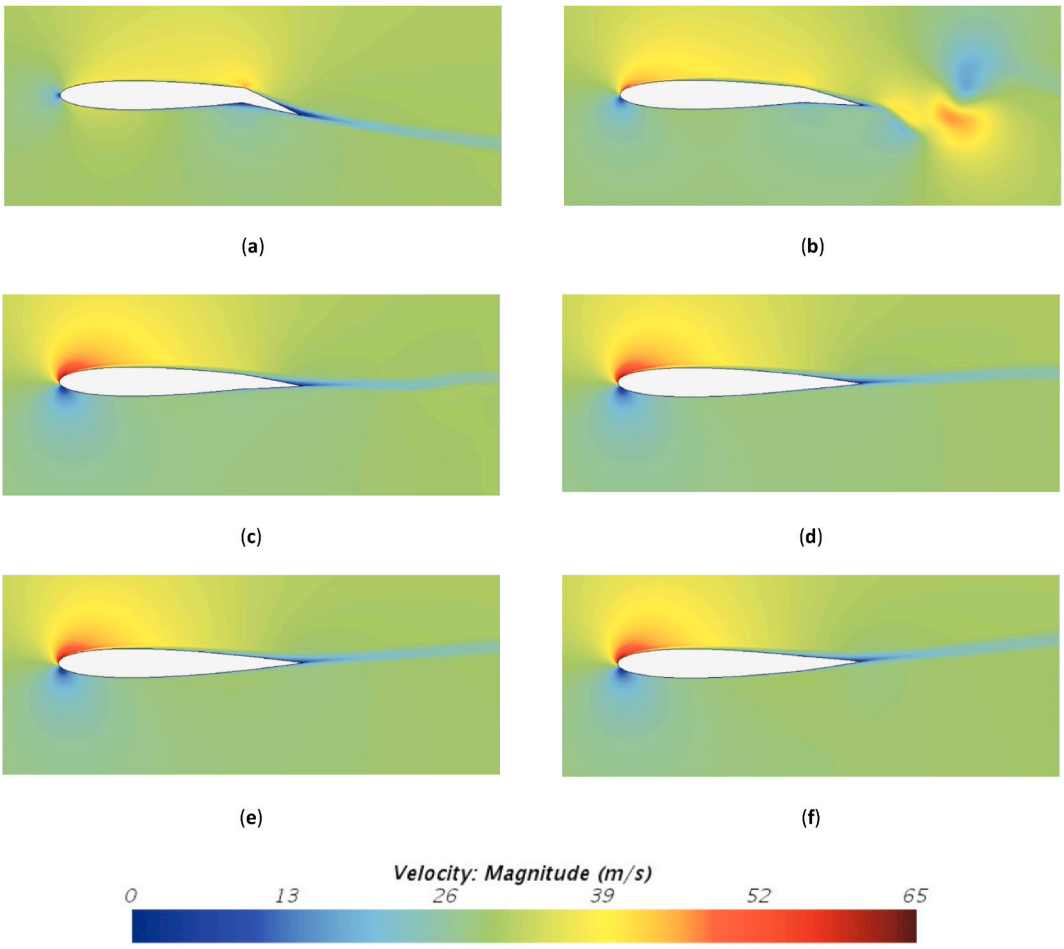


Fig. 8. Velocity profiles during a step. (a) $t = 1\,s$; (b) $t = 1.05\,s$; (c) $t = 1.1\,s$; (d) $t = 1.15\,s$; (e) $t = 1.2\,s$; (f) $t = 1.5\,s$.

sition, the decisions of the agent can be simply compared with the optimum values obtained with the ANN. These optimum values are obtained iteratively. Table 3 summarizes the α defined by the agent and the optimum α obtained iteratively with the ANNs.

The results show that the α set by the agent for all the *AoA* s is adequate, being very close to the optimum positions. The absolute α errors are between 0.7955 ° (for $AoA = 6°$) and 2.1761 ° (for $AoA = -8°$). In general, the largest errors appear for negative *AoA* s. Neverthe-

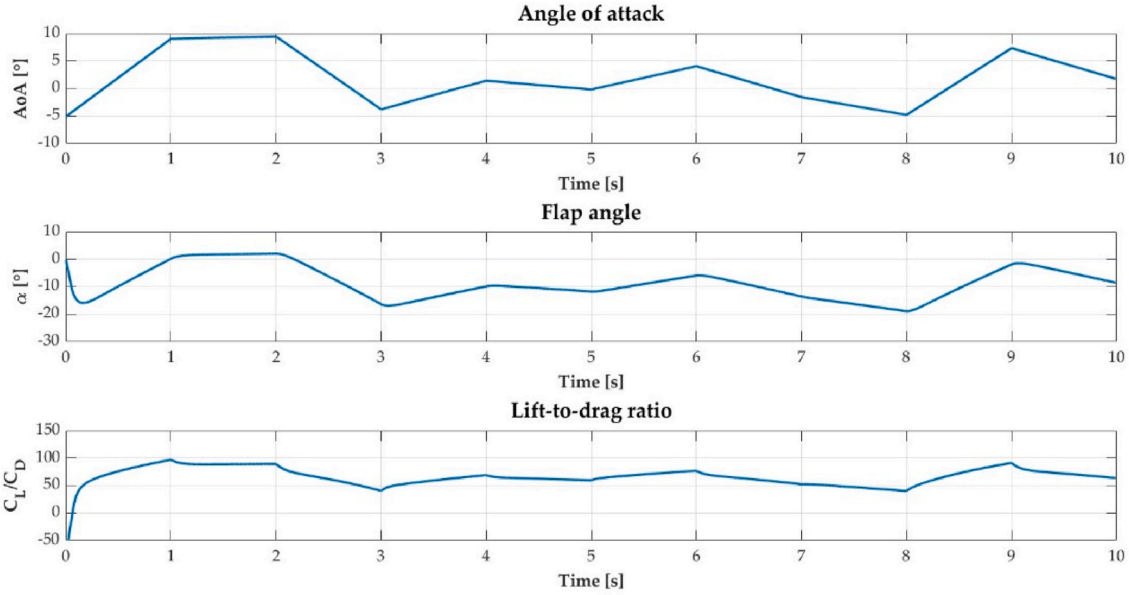


Fig. 9. Flap angle set by the agent for an *AoA* defined by a ramp sequence.

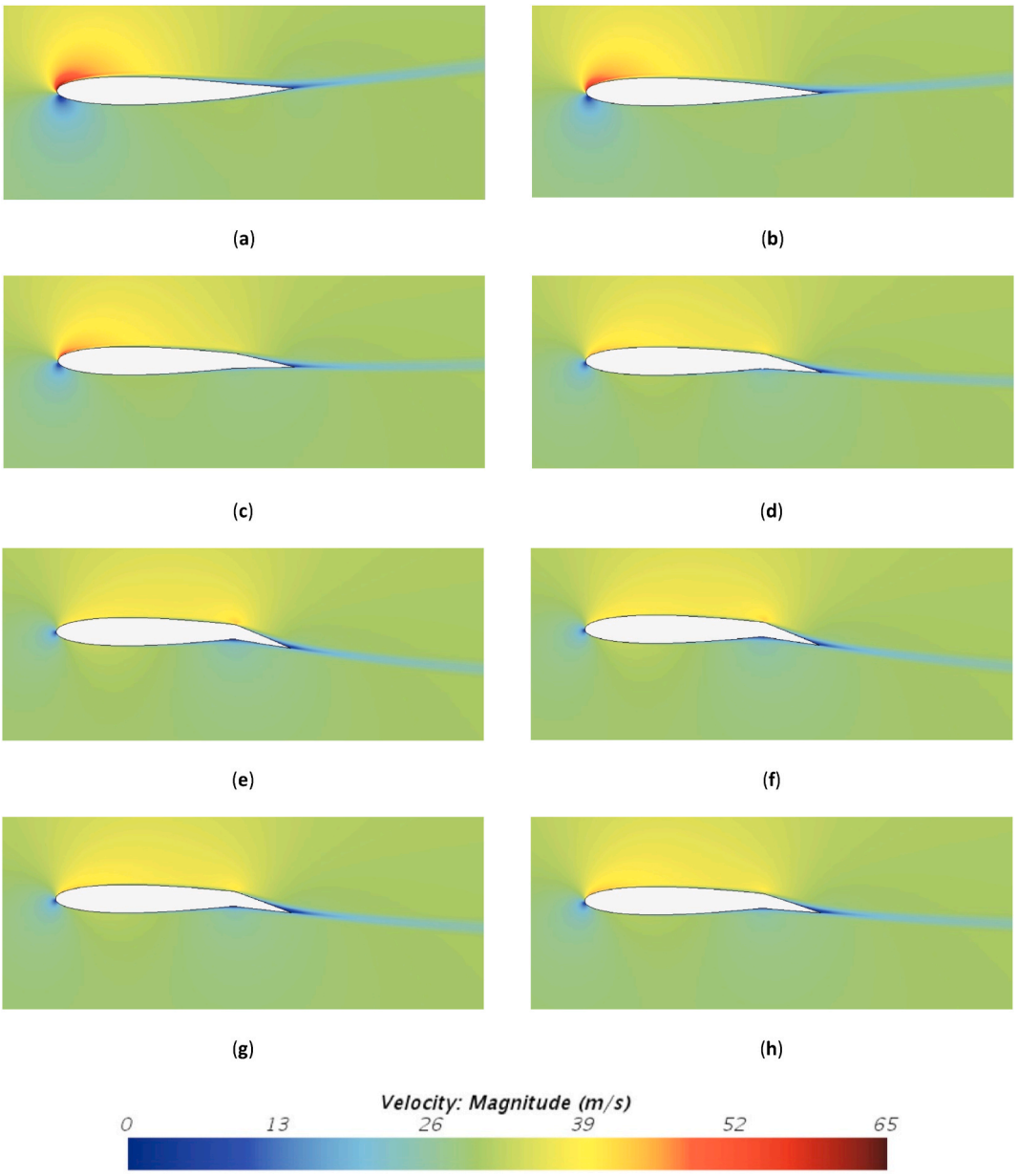


Fig. 10. Velocity profiles during the ramp sequence. (a) $t = 2\ s$; (b) $t = 2.25\ s$; (c) $t = 2.5\ s$; (d) $t = 2.75\ s$; (e) $t = 3\ s$; (f) $t = 3.25\ s$; (g) $t = 3.5\ s$; (h) $t = 3.75\ s$.

less, even if the α set by the agent is not the optimum one, these errors have almost no impact on the aerodynamic performance of the airfoil, as the relative C_L/C_D differences are between 0.39% (for $AoA = 8°$) and 3.05% (for $AoA = -8°$).

3.2. Stair sequence

In order to test the capability of the agent for adapting to sudden changes, a stair sequence of random values is considered for the *AoA*. Fig. 7 shows the response of the agent for this signal. As the results demonstrate, the agent is able to rapidly adapt to the *AoA* change, setting the optimum α for each case, and then, remaining static until the next *AoA* change. For all the tested cases, when the *AoA* changes, the agent sets a high ω; and then, as the flap approaches the optimum α, the agent decreases the ω. The C_L/C_D is kept high though the whole sequence, being the values close to the previously displayed optimum ones.

In order to provide a better understanding of flow phenomena during the step sequence, Fig. 8 shows the velocity profiles for the first step. Initially, at $t = 1\ s$ (Fig. 8a), the *AoA* is still $-5.11°$, and the flap remains stationary at $\alpha = -18.54°$. Immediately afterwards, the *AoA* changes to $9.06°$, causing the flap to move. In the initial moments after the change, the flap moves at high speed, leading to the generation of a vortex on the wake behind the airfoil, which can be seen at instant $t = 1.05\ s$ (Fig. 8b). As the flap approaches the desired angle, its velocity is gradually reduced, avoiding the generation of these vortexes. Finally, approximately at $t = 1.4\ s$, the flap reaches the desired angle, $\alpha = 1.53°$, and remains static until the next *AoA* step, as in $t = 1.5\ s$ (Fig. 8f).

3.3. Ramp sequence

The RL agent is tested with an *AoA* defined by a sequence of ramps to test the agent under continuously changing conditions. The values of the ramps are randomly defined. Fig. 9 shows the response of the agent for a ramp sequence. The results show that the agent continuously rotates the flap, adapting to the instantaneous *AoA*, and keeping the C_L/C_D high in all the cases. When the slope of the ramp changes, the agent changes the ω gradually, avoiding abrupt changes in α.

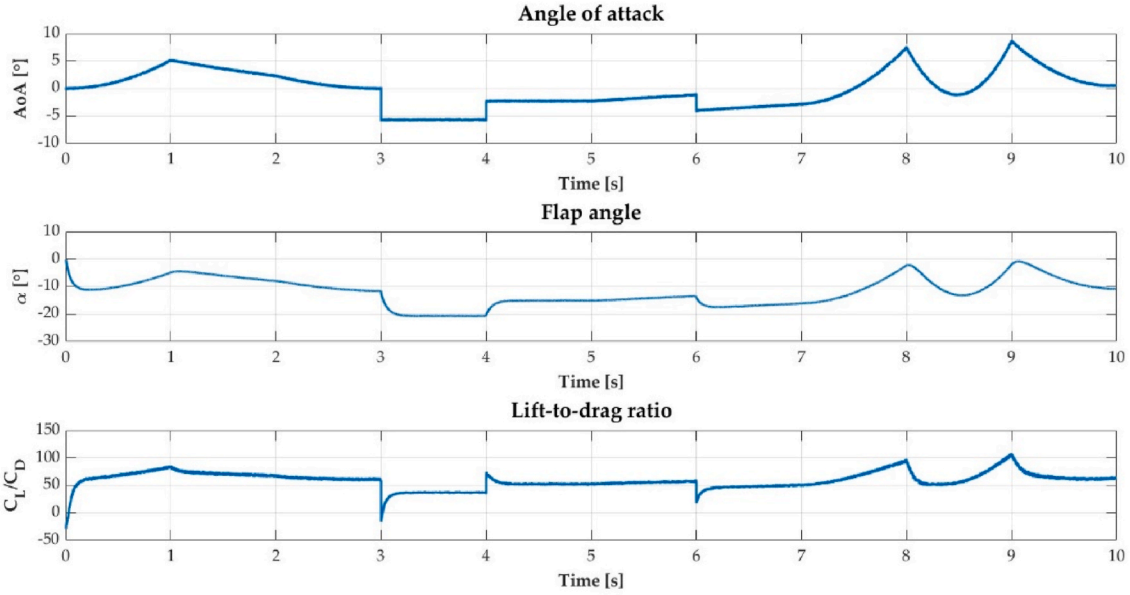


Fig. 11. Flap angle set by the agent for an *AoA* defined by a random sequence.

Fig. 10 shows the velocity profiles during the ramp sequence between $t = 2\,s$ and $t = 3.75\,s$, showing the results for an upward (from $AoA = 9.45°$ to $AoA = -3.81°$, between $t = 2\,s$ and $t = 3\,s$) and downward (from $AoA = -3.81°$ to $AoA = 0.11°$, between $t = 3\,s$ and $t = 3.75\,s$) *AoA* ramp. Velocity profiles show that, as the flap is continually rotating to adapt to each *AoA*, the changes on the wake behind the airfoil are slight. In contrast with the steps, as the rotation velocities are low, no vortex generation is observed on the wake.

3.4. Random sequence

A random *AoA* signal is defined to test the network under different conditions. This signal is composed of several different signals, such as ramps, steps, exponential signals, etc. Additionally, a slight noise signal is added through the whole simulation. Fig. 11 shows the response of the agent for this signal.

The results show that the agent is able to correctly respond to all the considered signals, keeping the C_L/C_D high. In addition, the agent ignores the noise, avoiding sharp changes of the flap angle. This occurs because of the penalty that has been added in the reward formula defined in Equation (21), which is proportional to the change in angular velocity.

4. Conclusions

The present paper aims to propose a reactive strategy based on a RL agent for controlling a rotating flap on a NACA0012 airfoil. To achieve this goal, firstly data about the evaluated case under various conditions is collected by means of CFD simulations. Then, two different ANNs are designed for aerodynamic coefficient prediction, in order to develop a fast and accurate model of the evaluated problem. Finally, a RL agent is trained, which determines the optimal angle of the rotating flap for each *AoA*.

The results demonstrate that the agent is able to adapt to all the *AoA* signals considered in this study. For constant *AoAs*, the α set by the agent is nearly equal to the optimum α, which is calculated iteratively with the designed ANNs. The main differences appear for negative *AoAs*. Nevertheless, the decisions made by the agent are considered adequate for all the cases; since, in the least favorable case, the C_L/C_D obtained by the agent, is only 3.05% below the optimum. With a step sequence, the agent rapidly changes the α, setting the optimum position; and with a ramp sequence, the agent constantly rotates the flap, establishing the optimum position for each instant. The agent is also tested with a random signal composed by various different signals, including a noisy signal. The agent is proven to be able to appropriately respond to all signals. With regards to the noisy signal, the agent is able to ignore it, avoiding abrupt changes in the angular velocity.

Therefore, the trained RL agent is able to provide a fast and accurate response for every type of *AoA* signal, rotating the flap to achieve a high C_L/C_D, and consequently, an appropriate aerodynamic performance of the airfoil in all conditions.

Although this study demonstrates the capability of RL-based methods for active flow control, only the *AoA* of the flow is being considered. For that reason, further research should focus on more complex systems, with more parameters that difficult the decision making of the agent. In addition, the vast majority of studies using RL for fluid mechanics focus on active flow control. However, fluid mechanics, and especially CFD, has many more areas where RL could be applicable, such as iterative parameter setting, and many others.

Funding

The authors were supported by the government of the Basque Country through the research grant ELKARTEK KK-2021/00014 BASQNET (Estudio de nuevas técnicas de inteligencia artificial basadas en Deep Learning dirigidas a la optimización de procesos industriales) and IT1514-22. K. P.-P. was supported by INVESTIGO program of the Basque Country 2022.

CRediT authorship contribution statement

Koldo Portal-Porras: Conceptualization, Validation, Investigation, Data curation, Writing – original draft, preparation. **Unai Fernandez-Gamiz:** Conceptualization, Validation, Formal analysis, Resources, Supervision, Funding acquisition. **Ekaitz Zulueta:** Methodology, Software, Writing – review & editing, Visualization, Project administration. **Roberto Garcia-Fernandez:** Software, Visualization, Project administration. **Saioa Etxebarria Berrizbeitia:** Formal analysis, Writing – review & editing, All authors have read and agreed to the published version of the manuscript.

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.

Data availability

Data will be made available on request.

Nomenclature and abbreviations

ANN	Artificial Neural Network
CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
DDPG	Deep Deterministic Policy Gradient
DL	Deep Learning
DRL	Deep Reinforcement Learning
GP	Genetic Programming
LE	Leading Edge
ML	Machine Learning
MLP	Multi-Layer Perceptron
RANS	Reynolds-Averaged Navier-Stokes
RL	Reinforcement Learning
SST	Shear Stress Transport
TE	Trailing Edge
AoA	Angle of Attack
α	Position of the flap
c	Chord length
C_D	Drag Coefficient
C_L	Lift Coefficient
C_L/C_D	Lift-to-drag ratio
f	Flap oscillation frequency
R	Radius of the mesh
R-value	Correlation coefficient
Re	Reynolds number
ρ	Density
u_∞	Freestream velocity
μ	Dynamic viscosity
ω	Angular velocity of the flap

Acknowledgments

The authors are grateful for the support provided by the SGIker of UPV/EHU.

References

Anılır, B.; Funda Kurtuluş, D.; Platzer, M.F. Flow control by oscillating trailing-edge flaps. AIAA J. 0, 1–4, doi:10.2514/1.J062947.

Anzalotta, C., Joshi, K., Fernandez, E., Bhattacharya, S., 2020. Effect of forcing the tip-gap of a NACA0065 airfoil using plasma actuators: a proof-of-concept study. Aero. Sci. Technol. 107, 106268 https://doi.org/10.1016/j.ast.2020.106268.

Aramendia, I., Fernandez-Gamiz, U., Ramos-Hernanz, J.A., Sancho, J., Lopez-Guede, J. M., Zulueta, E., 2017. Flow control devices for wind turbines. In: Bizon, N., Mahdavi Tabatabaei, N., Blaabjerg, F., Kurt, E. (Eds.), Energy Harvesting and Energy Efficiency, Lecture Notes in Energy, 37. Springer International Publishing, Cham, pp. 629–655, 978-3-319-49874-4.

Aramendia, I., Fernandez-Gamiz, U., Zulueta, E., Saenz-Aguirre, A., Teso-Fz-Betoño, D., 2019. Parametric study of a Gurney flap implementation in a DU91W(2)250 airfoil. Energies 12, 294. https://doi.org/10.3390/en12020294.

Aramendia-Iradi, I., Fernandez-Gamiz, U., Sancho-Saiz, J., Zulueta-Guerrero, E., 2016. State of the art of active and passive flow control devices for wind turbines. Dyna 91, 512–516. https://doi.org/10.6036/7807.

Bertsekas, D., 2012. Dynamic Programming and Optimal Control, I. Athena Scientific, 978-1-886529-43-4.

Boudis, A., Bayeul-Lainé, A.C., Benzaoui, A., Oualli, H., Guerri, O., Coutier-Delgosha, O., 2019. Numerical investigation of the effects of nonsinusoidal motion trajectory on the propulsion mechanisms of a flapping airfoil. J. Fluid Eng. 141, 041106 https://doi.org/10.1115/1.4042175.

Chandar, D.D., Boppana, B., Kumar, V., 2018. A comparative study of different Overset grid solvers between OpenFOAM, StarCCM+ and ansys-fluent. In: Proceedings of the 2018 AIAA Aerospace Sciences Meeting. American Institute of Aeronautics and Astronautics, Kissimmee, Florida.

Chen, H., He, L., Qian, W., Wang, S., 2020. Multiple aerodynamic coefficient prediction of airfoils using a convolutional neural network. Symmetry 12, 544. https://doi.org/10.3390/sym12040544.

Deep learning Toolbox. Available online. https://www.mathworks.com/products/deep-learning.html. (Accessed 22 May 2023).

Fan, D., Yang, L., Wang, Z., Triantafyllou, M.S., Karniadakis, G.E., 2020. Reinforcement learning for bluff body active flow control in experiments and simulations. Proc. Natl. Acad. Sci. USA 117, 26091–26098. https://doi.org/10.1073/pnas.2004939117.

Fernandez-Gamiz, U., Zulueta, E., Boyano, A., Ramos-Hernanz, J.A., Lopez-Guede, J.M., 2017. Microtab design and implementation on a 5 MW wind turbine. Appl. Sci. 7, 536. https://doi.org/10.3390/app7060536.

Gad-el-Hak, M., 1996. Modern developments in flow control. Appl. Mech. Rev. 49, 365–379. https://doi.org/10.1115/1.3101931.

Han, B.-Z., Huang, W.-X., Xu, C.-X., 2022. Deep reinforcement learning for active control of flow over a circular cylinder with rotational oscillations. Int. J. Heat Fluid Flow 96, 109008. https://doi.org/10.1016/j.ijheatfluidflow.2022.109008.

Hu, Z., Xu, G., Shi, Y., 2021. A new study on the gap effect of an airfoil with active flap control based on the Overset grid method. Int. J. Aeronaut. Space Sci. 22, 779–801. https://doi.org/10.1007/s42405-021-00364-0.

Hu, Z., Xu, G., Shi, Y., Xia, R., 2022. Airfoil–vortex interaction noise control mechanism based on active flap control. J. Aero. Eng. 35, 04021111 https://doi.org/10.1061/(ASCE)AS.1943-5525.0001356.

Hui, X., Bai, J., Wang, H., Zhang, Y., 2020. Fast pressure distribution prediction of airfoils using Deep learning. Aero. Sci. Technol. 105, 105949 https://doi.org/10.1016/j.ast.2020.105949.

IRENA, 2023. CPI Global Landscape of Renewable Energy Finance 2023. International Renewable Energy Agency, Abu Dhabi, 978-92-9260-523-0.

Julian, J., Iskandar, W., Wahyuni, F., 2022. Aerodynamics improvement of NACA 0015 by using Co-flow jet. International Journal of Marine Engineering Innovation and Research 7. https://doi.org/10.12962/j25481479.v7i4.14898, 2548–1479.
Kim, M.I., Yoon, H.S., 2022. Geometric modification for the enhancement of an airfoil performance using Deep CNN. Ocean. Eng. 266, 113000 https://doi.org/10.1016/j.oceaneng.2022.113000.
Koza, J.R., 1992. Genetic Programming. On the Programming of Computers by Means of Natural Selection.
Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D., 2019. Continuous Control with Deep Reinforcement Learning.
Lopez-Guede, J.M., Ramos-Hernanz, J.A., Zulueta, E., Fernadez-Gamiz, U., Oterino, F., 2016. Systematic modeling of photovoltaic modules based on artificial neural networks. Int. J. Hydrogen Energy 41, 12672–12687. https://doi.org/10.1016/j.ijhydene.2016.04.175.
Lopez-Guede, J.M., Ramos-Hernanz, J.A., Zulueta, E., Fernandez-Gamiz, U., Azkune, G., 2017. Dual model oriented modeling of monocrystalline PV modules based on artificial neuronal networks. Int. J. Hydrogen Energy 42, 18103–18120. https://doi.org/10.1016/j.ijhydene.2017.02.062.
MATLAB - MathWorks. Available online. https://www.mathworks.com/products/matlab.html. (Accessed 22 May 2023).
Menter, F.R., 1994. Two-equation eddy-viscosity turbulence models for engineering applications. AIAA J. 32, 1598–1605. https://doi.org/10.2514/3.12149.
Portal-Porras, K., Fernandez-Gamiz, U., Zulueta, E., Ballesteros-Coll, A., Zulueta, A., 2022. CNN-based flow control device modelling on aerodynamic airfoils. Sci. Rep. 12, 8205. https://doi.org/10.1038/s41598-022-12157-w.
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. J. Fluid Mech. 865, 281–302. https://doi.org/10.1017/jfm.2019.62.
Rasmussen, C.E., 2004. Gaussian processes in machine learning. February 2 - *14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures*. In: Bousquet, O., von Luxburg, U., Rätsch, G. (Eds.), Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 63–71, 978-3-540-28650-9.
Reinforcement learning Toolbox. Available online. https://www.mathworks.com/products/reinforcement-learning.html. (Accessed 22 May 2023).
Ren, F., Hu, H., Tang, H., 2020. Active flow control using machine learning: a brief review. J. Hydrodyn. 32, 247–253. https://doi.org/10.1007/s42241-020-0026-0.
Ren, F., Rabault, J., Tang, H., 2021. Applying Deep reinforcement learning to active flow control in weakly turbulent conditions. Phys. Fluids 33, 037121. https://doi.org/10.1063/5.0037371.
Richardson, L.F., Gaunt, J.A., 1927. The deferred approach to the limit. Philos. Trans. R. Soc. Lond. - Ser. A Contain. Pap. a Math. or Phys. Character 226, 299–361. https://doi.org/10.1098/rsta.1927.0008.
Rodriguez-Eguia, I., Errasti, I., Fernandez-Gamiz, U., Blanco, J.M., Zulueta, E., Saenz-Aguirre, A., 2020. A parametric study of trailing edge flap implementation on three different airfoils through an artificial neuronal network. Symmetry 12, 828. https://doi.org/10.3390/sym12050828.
Sekar, V., Jiang, Q., Shu, C., Khoo, B.C., 2019. Fast flow field prediction over airfoils using Deep learning approach. Phys. Fluids 31, 057103. https://doi.org/10.1063/1.5094943.
Sheldahl, R.E., Klimas, P.C., 1981. Aerodynamic Characteristics of Seven Symmetrical Airfoil Sections through 180-Degree Angle of Attack for Use in Aerodynamic Analysis of Vertical Axis Wind Turbines. Sandia National Lab. (SNL-NM), Albuquerque, NM (United States).
Simulink - Simulation and Model-Based Design Available online: https://www.mathworks.com/products/simulink.html (accessed on 24 May 2023).
Spalart, P., Allmaras, S., 1992. A one-equation turbulence model for aerodynamic flows. In: Proceedings of the 30th Aerospace Sciences Meeting and Exhibit. American Institute of Aeronautics and Astronautics, Reno, NV, U.S.A.
Spens, A., Pisano, A.P., Bons, J.P., 2023. Leading-edge active flow control enabled by curved fluidic oscillators. AIAA J. 61, 1675–1686. https://doi.org/10.2514/1.J062329.
STAR-CCM+ V2019.1 Available online: https://www.plm.automation.siemens.com/(accessed on 2 June 2020).
Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning. An Introduction, second ed. MIT Press. 978-0-262-35270-3.
Thuerey, N., Weißenow, K., Prantl, L., Hu, X., 2020. Deep learning methods for Reynolds-averaged Navier–Stokes simulations of airfoil flows. AIAA J. 58, 25–36. https://doi.org/10.2514/1.J058291.
Tunio, I.A., Kumar, D., Hussain, T., Jatoi, M., 2020. Safiullah investigation of variable spanwise waviness wavelength effect on wing aerodynamic performance. Fluid Dynam. 55, 657–669. https://doi.org/10.1134/S0015462820040102.
Ugarte-Anero, A., Fernandez-Gamiz, U., Portal-Porras, K., Zulueta, E., Urbina-Garcia, O., 2022. Computational characterization of the behavior of a saliva droplet in a social environment. Sci. Rep. 12, 6405. https://doi.org/10.1038/s41598-022-10180-5.
Vinuesa, R., Lehmkuhl, O., Lozano-Durán, A., Rabault, J., 2022. Flow control in wings and discovery of novel approaches via Deep reinforcement learning. Fluid 7, 62. https://doi.org/10.3390/fluids7020062.
Wang, Z., Zeng, T., Chu, X., Xue, D., 2023. Multi-objective Deep reinforcement learning for optimal design of wind turbine blade. Renew. Energy 203, 854–869. https://doi.org/10.1016/j.renene.2023.01.003.
Watkins, C.J.C.H., Dayan, P., 1992. Q-Learning. Mach Learn 8, 279–292. https://doi.org/10.1007/BF00992698.
Zhang, Y., Sung, W.J., Mavris, D.N., 2018. Application of convolutional neural network to predict airfoil lift coefficient. In: AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. American Institute of Aeronautics and Astronautics.