Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 34(2): 1290–1314.
DOI: 10.3934/era.2026059
Received: 16 December 2025
Revised: 28 January 2026
Accepted: 02 February 2026
Published: 09 February 2026

Research article

Optimization of compound performance of two tandem pitching and heaving aerofoils in Martian environment

Muhammad Khan, Li Wang and Fang-Bao Tian*

School of Engineering and Technology, University of New South Wales, Canberra, ACT 2600, Australia

* **Correspondence:** Email: fangbao.tian@unsw.edu.au.

Abstract: Mars exploration has increased the interest in bio-inspired flapping-wing micro aerial vehicles due to their high maneuverability and efficient performance in low Reynolds number environments. This work presents an optimization of the compound performance of two tandem pitching and heaving aerofoils in the Martian environment by using the hybrid method of the reinforcement learning (RL) and the immersed boundary-finite difference method (IB-FDM). The searching parameters include the horizontal and vertical spacing, phase shift, and mean angle of attack (AoA) of the hindwing at a Reynolds number of 1100. It is found that the optimal phase shifts lie close to in-phase and out-of-phase motions, while intermediate phase shifts tend to produce suboptimal results. For the in-phase cases, an increase in AoA and a decrease in horizontal spacing led to an increase in the compound performance. For the out-of-phase cases, a similar trend is observed until the optimal value is reached, after which the performance begins to decline. The effects of vertical spacing vary vastly on a case-to-case basis, depending on other motion and positional parameters.

Keywords: pitching and heaving aerofoils; bio-inspired flapping wing; Martian environment; immersed boundary-finite difference method; low Reynolds number

1. Introduction

Aerial vehicles, such as aeroplanes and helicopters, have attracted increasing attention in Mars exploration as they have better manoeuvrability than land rovers and provide higher resolution imaging than satellites [1]. The development and successful deployment of the Ingenuity [2] made a significant advancement in Mars exploration with aerial vehicles. However, there is still considerable area for improvement. At the low Reynolds number regime, flapping-wing aerial vehicles (FWAVs) inspired by insects and birds have better aerodynamic efficiency compared with the fixed-wing and rotary-wing designs [3]. This is particularly important, as the Reynolds numbers on Mars are much lower

compared to the scenario on the Earth [4]. As insects have an excellent ability to hover and steer during forward flight by changing the flapping motion of their wings, the FWAVs could also provide better maneuverability [5]. Therefore, it is desired to study the flapping wings in Martian environment.

The aerodynamics of flapping wings has been an area under significant research since the early 1960s. The National Aeronautics and Space Administration (NASA) conducted a series of tests over the tandem configurations to improve the performance of helicopter blades [6,7], aircraft flaps [8], and boundary layer control mid chord flaps [9]. Their results led to further investigation of the tandem wing configuration. The most notable ones were in the 2000s when Broering and Lian [10] investigated the effect of airfoil spacing on the aerodynamics of pitching airfoils. They concluded that for this configuration, the overall lift increases with increase in spacing between the airfoils. Afterwards, they performed extensive research on the effect of phase between the pitching and heaving tandem airfoils and the effects that it has on the lift coefficient, both for the 2-dimensional [11] and 3-dimensional [12] cases. They found out that the vortex interactions between the fore and hind wings can be characterized as either constructive or destructive depending on the phase. For a 0° phase difference, they found out that constructive vortex interference happens which results in increased coefficients of lift and thrust, whereas destructive vortex interference happens for the 180° case. These results show that the in-phase motion produces the highest lift and thrust; however, it also requires the highest power. This means that the compound performance may not be maximum for the in-phase case. Seet [13] and Lee [14] separately showed this by concluding that the maximum aerodynamic efficiency is achieved for a 180° phase shift for pitching airfoils at minimum spacing. They also showed that increasing the phase shift forces the trend of increasing spacing enhances lift to reverse.

Since 2010, there has been a significant amount of research in the field of flapping-wing-drone systems for low-density environments like Mars. Jaroszewicz et al. [15] presented a working model of a flapping wing micromechanical flying insect (MFI) called Entomopter. They concluded that this Entomopter can be used in environments like Mars where the Reynolds number is low. Afterwards, Mannam et al. [16] also explored the potential for an insect-like flapping propulsion system for Mars and concluded this option to be highly efficient given the low gravity on Mars. Bluman et al. [17] and Bezard et al. [18] separately numerically researched the performance of their bio-inspired flapping wing robots in the Martian environment using dynamic scaling. They suggested that the wing span of a flapping robot can be easily increased on Mars to produce higher lifts at the cost of a slightly higher power required. As it is easier to ensure structural integrity on Mars, longer wings can be created, thereby requiring a smaller lift coefficient for the same lift required. Furthermore, Mccain and Lee [19] and Singh et al. [20] have demonstrated the potential of using FWAVs that generate reasonable lift for easy Mars exploration. The effects of compressibility and rarefaction, as well as scaling laws of flapping wings have been studied numerically by Widdup et al. [4,21,22]. In addition, researchers have developed vacuum Martian environment chamber and low-pressure wind tunnel to study the aerodynamics of aerial vehicles in Martian environment [23,24]. Some bio-inspired research projects have explored the feasibility, advantages, and scaling of FWAVs for Martian exploration [25] whereas, others focused on the navigation and control of these bio-inspired machines [26,27].

As this is an emerging field, the advancement of computational capabilities opens up multiple avenues for the implementation of machine learning techniques like reinforcement learning (RL) for parameter optimization and active control. As Mars is remote, the autonomous capability that RL provides can significantly improve the performance of these machines thereby improving the

overall mission [28], showing tremendous potential to promote the next generation of FWAVs in the Martian environment.

Recently, Solar et al. [29] published research on the maximisation of wind turbine generation using RL. They used a discrete action space and defined seven actions that the agent could take by controlling the value of yaw, pitch, and rotor speed. The model that they used was a double deep Q-network (DDQN) [30] consisting of two networks, a main and a target network. The target network has fixed parameters that are updated periodically to match the main network parameters. This enables the model to learn slowly but in the right direction as the main model constantly has the target model to compare with. However, a significant drawback of this algorithm is that it is not suitable for continuous action spaces. Koldo et al. [31] successfully trained an active control model of an airfoil flap whose aim was to maximize the overall lift-to-drag ratio with changing AoA. They achieved this by first training two supervized learning models for lift and drag coefficients on a known dataset and then training an RL model based on the outputs from these supervized learning models. However, the downside of this technique is that it requires a huge dataset, which is computationally expensive and time-consuming. Additionally, Bao et al. [32] trained an RL model whose aim was to maximize the propulsive performance of a pitching and heaving airfoil using a policy gradient method by controlling the pitching and heaving amplitudes and their offset. An important technique used by them was the implementation of a multi-environment [33] for agent training. This allows the agent to run multiple simulations at a time, thereby reducing the overall training time significantly. Furthermore, Lee et al. [34] developed a novel RL technique to optimize the topology of a heat exchanger for better performance. This resulted in a 14.8% improved heat exchanger performance. Another closely related work was conducted by Ji et al. [35] who used Bayesian Optimization [36] to optimize the efficiency of tandem pitching and heaving airfoils by finding the best combination of the motion parameters. This is a powerful strategy, however, a significant drawback to it is that it requires an initial dataset with simulation results and does not have the ability to generalize and adapt to different environments as RL does, which could be used to explore configurations that have different airfoils, *Re* or Mach numbers.

This research aims to implement the use of RL for two tandem pitching and heaving airfoils to optimize the lift and thrust compound performance by controlling the motion and positional parameters of the hindwing in the Martian environment. The tandem airfoil configuration was chosen as it results in a much more maneuverable system and the hindwing consumes less power as compared to a single airfoil [37], making it more efficient. According to Nagai et al. [38], the effect of phase difference is still not understood properly due to differences in the results of different studies. This means that other parameters also affect this relation; hence, using RL could lead to a better understanding of this phenomenon.

2. Physical description and numerical method

This section will discuss the methodology adopted. Initially, the simulation setup will be discussed, whereas the RL architecture will be discussed later.

2.1. Physical problem

A tandem airfoil arrangement of two NACA 0012 airfoils of chord c in a uniform flow, as shown in Figure 1, is considered in this work. The pivot point is fixed at $c/3$, and the horizontal spacing between

the two airfoils is d. Here the NACA 0012 airfoil is part of the 4-digit NACA airfoil series and is chosen as it is the best airfoil for efficiency and thrust production for flapping motion [39].

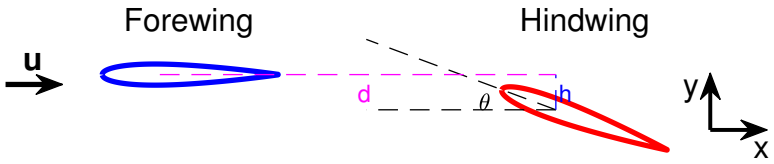


Figure 1. Sketch of two airfoils in tandem arrangement.

The pitching and heaving motion of the airfoils can be described by

$$\alpha = \alpha_{\mathrm{m}} \sin(2\pi f t + \phi) + \theta, \tag{2.1}$$

$$y = A_m \sin(2\pi f t + \phi + \psi) + h, \tag{2.2}$$

where α is the pitching angle; y is the heaving distance; ψ is the phase difference between the pitching and the heaving motion; α_{m} is the pitching amplitude; A_m is the heaving amplitude; t is the instantaneous time; f is the frequency of both pitching and heaving; and ϕ, θ, and h are the phase shift, the initial angle of attack, and the initial vertical spacing for the hindwing from the forewing, respectively. The three ϕ, θ, and h are zero for the forewing equations. Each motion will be described as a set of these hindwing parameters: (ϕ, h, d, θ), whereas they are zero for the forewing equations. The motion over a single oscillation can be seen in Appendix A.

The aerodynamics is governed by the two-dimensional compressible Navier-Stokes equations, which are detailed in our previous publications [40, 41]. The Mach number based on the incoming flow u_0 is fixed at 0.1, based on which the maximum local Mach number approaches 0.3. This number is used because it is in the range where the NASA's Ingenuity was operated, while the formation of vortex structures is not significantly inhibited [4, 21, 22]. The Reynolds number (Re) given by

$$Re = \frac{\rho u_o c}{\mu}, \tag{2.3}$$

where ρ and μ the density and the viscosity, respectively. $Re = 1100$ is used which is based on the atmospheric conditions on Mars given in Appendix 4 and the average chord size of $5cm$. The pitching and heaving amplitudes are respectively fixed at 30° and $1c$, based on [32]. In addition, $St = fc/u_o = 0.3$ at which the maximum thrust is achieved for $25° < \alpha_m < 30°$ and $A_m = 1c$ [32]. Finally, ψ is 90° as Isogai et al. [42] demonstrated that this value results in maximum thrust efficiency.

2.2. Numerical methods

The aerodynamics due to the flapping foils is solved by using an immersed boundary-finite difference method, as detailed in [43, 44]. This method employs the fifth-order weighted/targeted non-oscillatory scheme [45, 46] for the convective term. The viscous term is discretized by the fourth-order central difference method. The unsteady term is handled by the third-order Runge–Kutta scheme. The boundary condition at the fluid-foil surface is implemented by an immersed boundary method due to its convenience in handling flows involving complex geometries and large-amplitude movement of the bodies [47–49]. The CFD solver used in this work is an in-house solver that has been well validated

in our previous work [43, 44], and recently has been adapted to model the aerodynamics of flapping wings in the Martian environment [4, 21, 22, 44, 50]. Furthermore, the optimal mesh spacing of $0.02c$ was chosen based upon the comparison between the results of multiple simulations with different mesh spacings. The non-reflection characteristic boundary conditions are applied on the boundaries. The CFD simulations are being initiated by the RL algorithm based on its decisions, aimed at improving the performance of the RL agent.

2.3. RL model setup

The RL algorithm used for the agent is the proximal policy optimization (PPO) algorithm [51]. It is a policy gradient method based on an actor which takes an action in an environment and a critic which provides feedback on the quality of action taken [52]. It is a very popular and vastly applicable method which has been successfully implemented for a similar aerodynamics application [53].

The training of the RL agent is performed using TensorFlow [54] in Python. The actor model consists of three dense layers with the first two layers consisting of 256 neurons and an activation function of noisy rectified linear unit (NReLU) which is used to ensure that the actor decently explores the environment before finalising the policy and also to limit any bias that initial training results could introduce [55]. On the other hand, the last dense layer which is the output layer, consisted of 4 neurons based on the amount of parameters that the actor was controlling. The tanh (hyperbolic tangent) activation function is chosen for this layer to avoid NaN values and restrict the actions within -1 and 1 [56]. The critic model is also very similar to the actor but with minor changes. Firstly, its output layer which is the feedback to the actor had only 1 neuron and no activation function. Secondly, unlike the actor which only focuses on the actions, the critic is set up to draw one-to-one correspondence between the actions and the states which could lead to a better training outcome [57].

The environment is created using the Gymnasium library [58] with the 4 motion parameters as variables. The continuous action and observation spaces are set up in a way such that each action represents a certain combination of parameters, regardless of the previous state. This helps the agent to easily navigate within the observation space and also removes the dependency of the model on the previous state, which is the case as future simulation results are not dependent on past simulation results.

The variable bounds for θ are chosen based on the results of Bao et al. [32] showing that 30° pitching amplitude produces the best thrust for the chosen St. On the other hand, the bounds for h and d are fixed to ensure that the vortex interaction occurs effectively, as it results in better performance [37]. These values can be seen in Table 1.

Table 1. Environment variables.

Variables	ϕ	h	d	θ
Bounds	(0°, 180°)	(–0.4c, 0.4c)	(1.05c, 3.45c)	(0°, 30°)

The Adam optimizer with a small learning rate of 1×10^{-4} is used for both neural networks. The actor tries to learn the policy $\pi(a_t, s_t)$ as can be seen in Figure 2 by updating it after each learning step which consisted of 32 different simulations. The reward from each step and the advantage from the

critic are then used to calculate the clipped loss which is the actor loss and is given by

$$\mathcal{L}^{\text{CLIP}}(\theta) = \hat{E}[\min(r(\theta)\hat{A}_t, \text{clip}\,(r(\theta), 1-\epsilon, 1+\epsilon)\,\hat{A}_t)], \tag{2.4}$$

with the critic loss given by,

$$\mathcal{L}_{\text{critic}}(w) = \left(V_w(a_t, s_t) - V_t^{\text{target}}\right)^2, \tag{2.5}$$

where $\hat{A}_t$ is the advantage at time t, $\hat{A}_t = \delta_t + (\gamma\lambda)\delta_{t+1} + \ldots + (\gamma\lambda)^{T-t+1}\delta_{T-1}$ with δ_t calculated based on the reward; θ is the vector containing the weights for the policy; V is the value of the state; w is the vector containing weights for the value network; γ is known as the discount factor and a smaller value for it is used to degrade the impact of future state values on current state values, thereby focusing on the current value [53]; and λ is the smoothing factor and was kept lower to ensure high dependence on immediate step rewards rather than the long-term ones [53]. The batch size, policy clip, and number of epochs are all chosen based on recommendations from Andrychowicz et al. [59] and are also fine-tuned throughout the training procedure with the final values shown in Table 2.

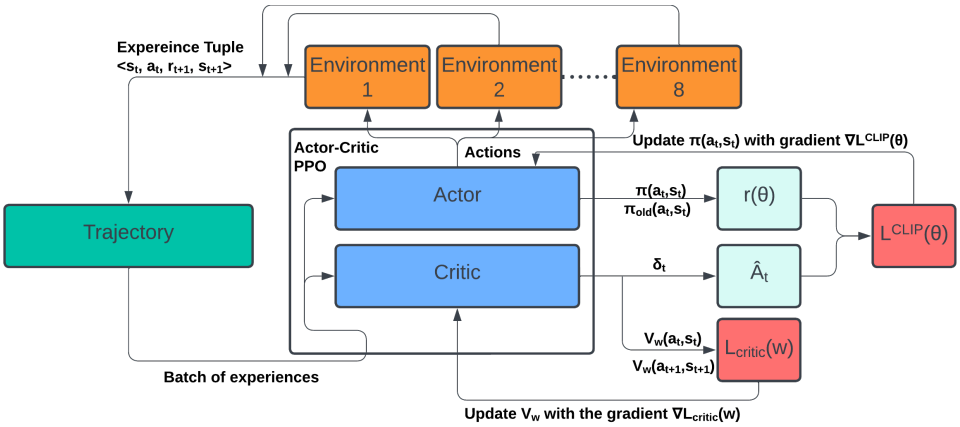


Figure 2. Flowchart of the PPO algorithm.

Table 2. Hyperparameters for RL model.

Hyper-parameters	γ	λ(GAE)	Policy clip	Batch size	Epochs
Value	0.01	0.05	0.2	32	8

The multi-environment technique is also implemented, similar to the implementation of Bao et al. [32]. This allowed the agent to interact with 8 different environments at once significantly reducing the training time by 8 times as well. The number of environments is chosen based on the availability of computational resources and considerations, such as the quality of training, which could decrease with excessive environments [33].

After initialisation of the parameters, the agent executes a step that constitutes an action on each of the 8 CFD solvers by submitting jobs for a particular combination of parameters. The environment then outputs the trajectory which consists of the observation, action taken and reward. As the aim of the agent is to find the best lift and thrust compound performance, the reward function is formulated to account for it. It also heavily penalizes the agent for a negative thrust or lift coefficient as can be seen in Eq (2.6).

$$Reward = \frac{\langle C_T \rangle}{\langle C_P \rangle} + \frac{\langle C_L \rangle}{\langle C_P \rangle} - 5ifneg(\langle C_T \rangle) - 5ifneg(\langle C_L \rangle), \tag{2.6}$$

where $ifneg()$ returns 1 for a negative value and 0 for a positive value and $\langle \cdot \rangle$ indicates the average values of C_T, C_L, and C_P over a single oscillation which are the coefficients of thrust, lift and power respectively. Their formulas are as follows:

$$C_T = \frac{Thrust}{\frac{1}{2}\rho u_o^2 c}, \quad C_L = \frac{Lift}{\frac{1}{2}\rho u_o^2 c}, \quad C_P = \frac{Power}{\frac{1}{2}\rho u_o^3 c}. \tag{2.7}$$

The performance parameter accounts for the compound performance and is defined as the reward function without the $ifneg()$ conditions.

In order to ensure that the correct mesh size has been employed for the simulations, mesh independence tests were conducted to find the optimal mesh size that provides decent accuracy and consumes reasonable time. Figure 3 contains the results for 3 different mesh sizes. Doubling the mesh size was leading to a four times increase in the simulation time hence a mesh size of 0.02c was chosen resulting in a single simulation run time of 80 minutes on an 8 core CPU utilising 4 GB of RAM. The mesh grid comprises of multiple squares that are increasing in size with increasing distance from the center.

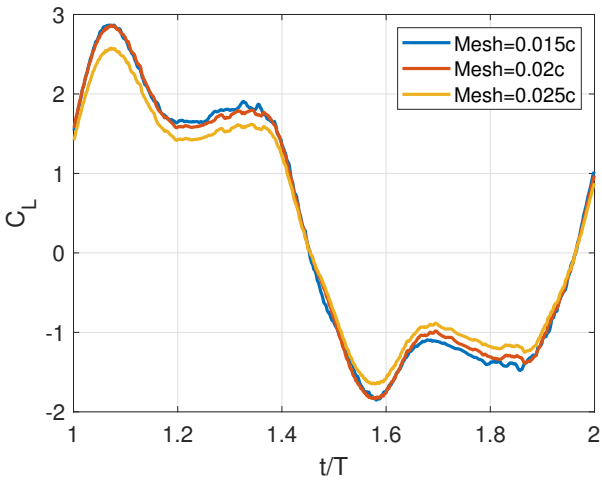


Figure 3. C_L vs t/T plot for mesh independence test.

3. Results and discussion

3.1. RL training

The model is trained for over 3500 steps which took about 640 hours of CPU time to finish. The average reward along with the standard deviation can be seen in Figure 4. The figure shows that the reward initially goes up as the agent finds the first case in Table 3 with a compound performance as in Eq (2.6) of 0.67 Afterwards, the agent goes on to cleverly explore intermediate parameter combinations which leads to a higher standard deviation and a decrease in the average rewards. However, as can be seen on the 17th policy update, the agent discovers the out-of-phase Case 3. Though the performance

was not the best, it was still decent with a smaller C_T and a reasonable C_L, hence could be used for hovering. It then went on further and discovered Case 2 in Table 3 after 43 policy updates. These optimal cases were found within just 1300 steps and the remaining training was done to ensure that no other case with better performance was left undiscovered and that the most optimum policy was achieved with reasonable standard deviation. The final agent can be seen to have the highest average reward and lowest standard deviation and further training would lead to overfitting, leading to a decrease in the overall performance [60]. Due to noise in the agent, the average rewards would not converge however will maintain a reasonably high value [61].

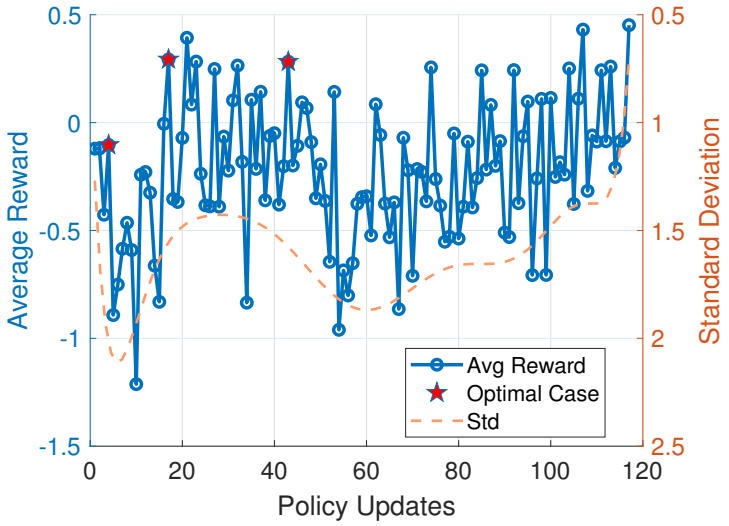


Figure 4. Average reward and standard deviation during training.

Table 3. Optimal cases.

Parameters	C_T	C_L	C_P	**Per**
$(0°, 0c, 1.05c, 30°)$	1.88	9.26	16.62	0.67
$(0°, 0.15c, 1.05c, 30°)$	1.93	10.15	17.30	0.70
(177.2°,0.39c,1.26c,23.5°)	0.60	6.57	11.20	0.64

Overall, as can be seen in Figure 5, it was observed that the in-phase and out-of-phase motions were producing the best performance. In line with [11], the in-phase motions produced higher values of C_L, C_T and C_P, mainly due to constructive interference between the forewing and hindwing whereas, out-of-phase motions produced a relatively low value for these coefficients due to destructive interference however, maintained a similar value for the overall performance as can be seen in Table 3. In general, in line with [62], an increase in θ led to a decrease in C_T, an increase in C_L, an increase in performance, and vice versa. Whereas a decrease in d mostly led to an increase in the overall performance; however, for certain cases, mostly for the out-of-phase cases like Case 3, the optimal value for d lay somewhere in between. The effect of h on the performance can be seen to be varying significantly with other parameters, especially with ϕ. This is because each value for h affects the vortex interaction differently for each parameter combination. However, the positive values for h can be seen to produce better performance, with the optimal value lying somewhere in between. For better understanding, the scatter

plots for compound C_T, C_L, C_P and performance can be seen in Appendix C.

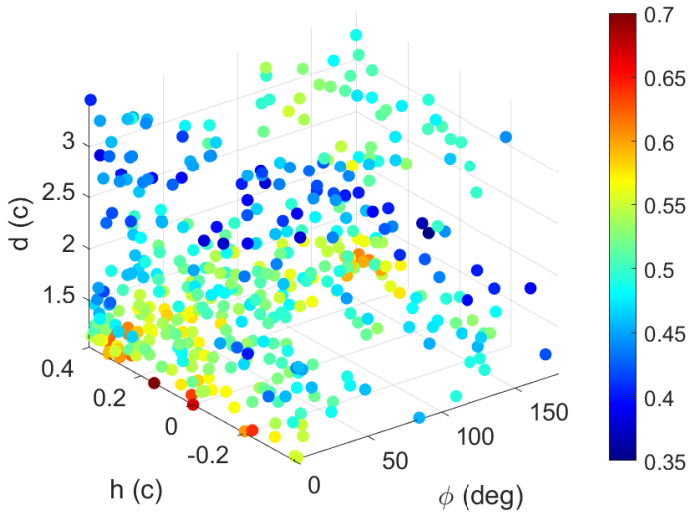


Figure 5. Performance scatter plot for $\alpha = 30°$.

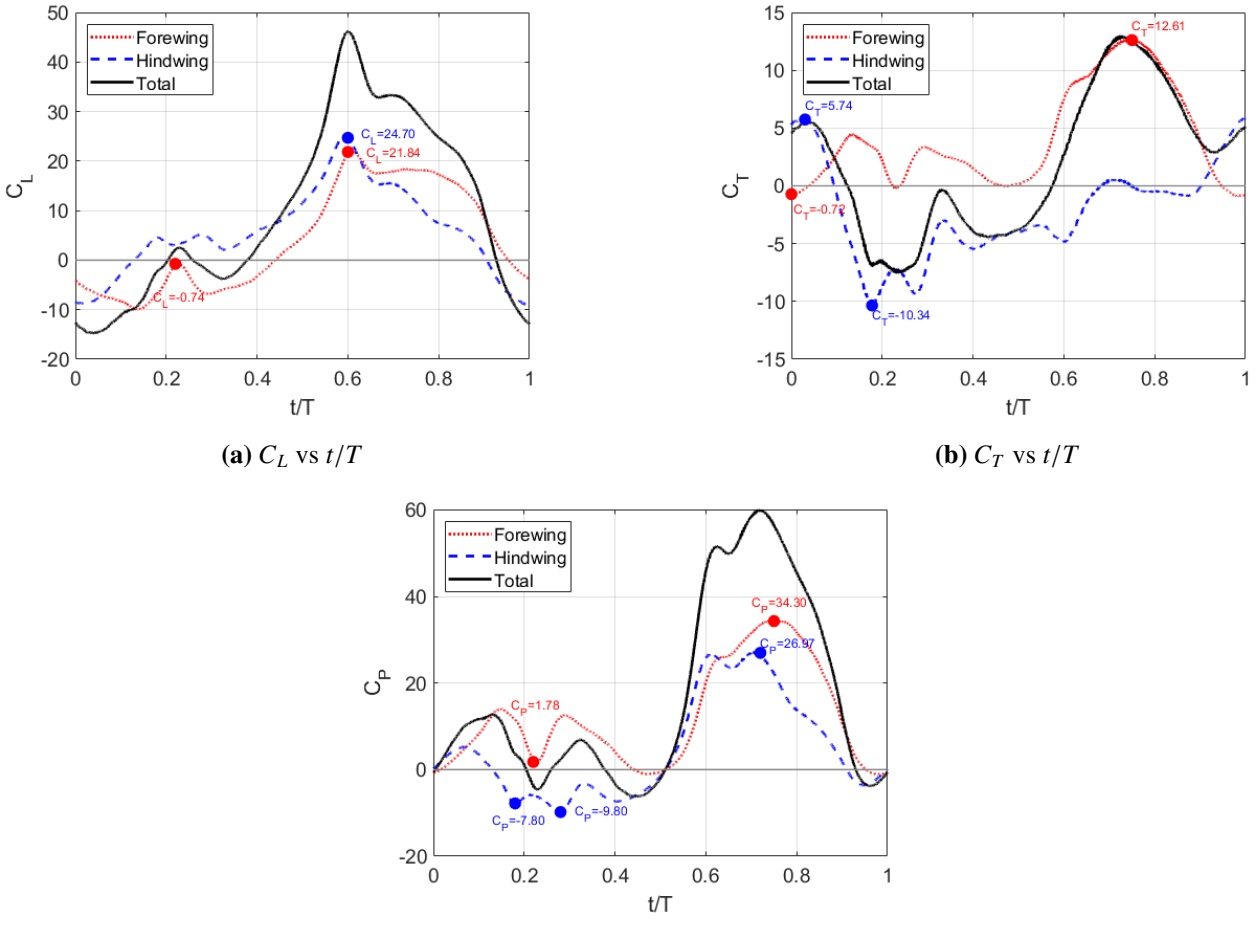
3.2. In-phase Optimal Case-$(0°, 0.15c, 1.05c, 30°)$

This case is in line with [63], which suggests that the forewing and hindwing at minimum d produce a higher overall performance as compared to an isolated wing. The C_L, C_T, and C_P for this case can be seen in Figure 6, respectively. Whereas, the coefficient of pressure and vorticity flow fields for the important points can be seen in Figure 7. For the forewing, the leading edge vortices (LEVs) generated during the downstroke and upstroke phases are referred to as D-LEV and U-LEV respectively, with the trailing edge vortices (TEVs) following a similar notation. Furthermore, for the hindwing, the corresponding vortices are denoted by lowercase letters, like u-LEV or d-TEV. Another notation C_{xfor} and C_{xhin} will be used for the forewing and hindwing C_T, C_L, and C_P, respectively.

As can be seen at $t/T = 0.03$, maximum C_{Thin} is produced by the strong low-pressure region at the hindwing LE caused by the D-LEV. The C_{Tfor} is minimum at this point as the forewing has a very low AoA and only the pressure regions near the LE and TE could produce a horizontal motion, which can be seen are very small and weak. Afterwards, C_{Thin} drops steeply to its minima at $t/T = 0.18$ which is caused by the detachment of the D-LEV from the hindwing LE and the strong shear layer on top of it producing a strong low-pressure region behind it, sucking it back. Additionally, as this suction is happening in the direction of the motion, the C_{Phin} can be seen is negative, indicating that the fluid is causing the motion, effectively lifting the hindwing. Afterwards, the shear layer separates resulting in a gradual increase in C_{Thin}.

Both the C_{Lfor} and C_{Lhin} stay relatively moderate at the start with a minor fluctuation for C_{Lfor} leading to a local maxima at $t/T = 0.22$, caused by the interaction of the D-LEV with the end part of the forewing. This resulted in the production of a limited low-pressure region above the forewing which caused C_{Lfor} to rize. This interaction allowed the fluid to work on the forewing in the direction of motion which caused a dip in C_{Pfor}. Both C_{Lfor} and C_{Lhin} then start to increase rapidly after $t/T = 0.31$ which is because the D-LEV and U-LEV which were producing a low-pressure region below the hindwing and forewing respectively till $t/T = 0.31$ escaped, which allowed for a rapid increase in

C_{Lfor} and C_{Lhin} till they reach their maximum value of 21.84 and 24.70 at $t/T = 0.60$ respectively. At $t/T = 0.6$, both the airfoils have started the downstroke and the hindwing being close to the forewing shaped like an inverted cup is producing a high-pressure region underneath both the airfoils, thereby producing lift. The u-LEV is producing low pressure over the hindwing which is also contributing to C_{Lhin}. Afterwards, C_{Lfor} and C_{Lhin} start to decrease with a positive bump from $t/T = 0.69$ till $t/T = 0.75$ which is caused by the interaction between the U-LEV, D-LEV, and D-TEV, in line with [64] suggesting that this blockage of D-TEV by the hindwing would lead to this bump. This resulted in a momentary low-pressure region above the LE of the hindwing, thereby increasing C_{Lhin}. On the other hand, maximum C_{Tfor} is achieved at $t/T = 0.75$ where the forewing has a strong low-pressure region above it caused by the downstroke and D-LEV which is sucking the forewing forward, producing thrust. Similar to the minimum C_{Tfor} case, C_{Thin} is almost zero at this point as the hindwing has a very low AoA with a pressure balance between LE and TE. The maximum C_{Pfor} and C_{Phin} are achieved close to $t/T = 0.75$ as the airfoils are continuing the pitching and heaving downstroke motion against the net pressure distribution which is producing a decent C_T and C_L. This forces the airfoils to work against the fluid; hence large C_{Pfor} and C_{Phin} are required.



(a) C_L vs t/T

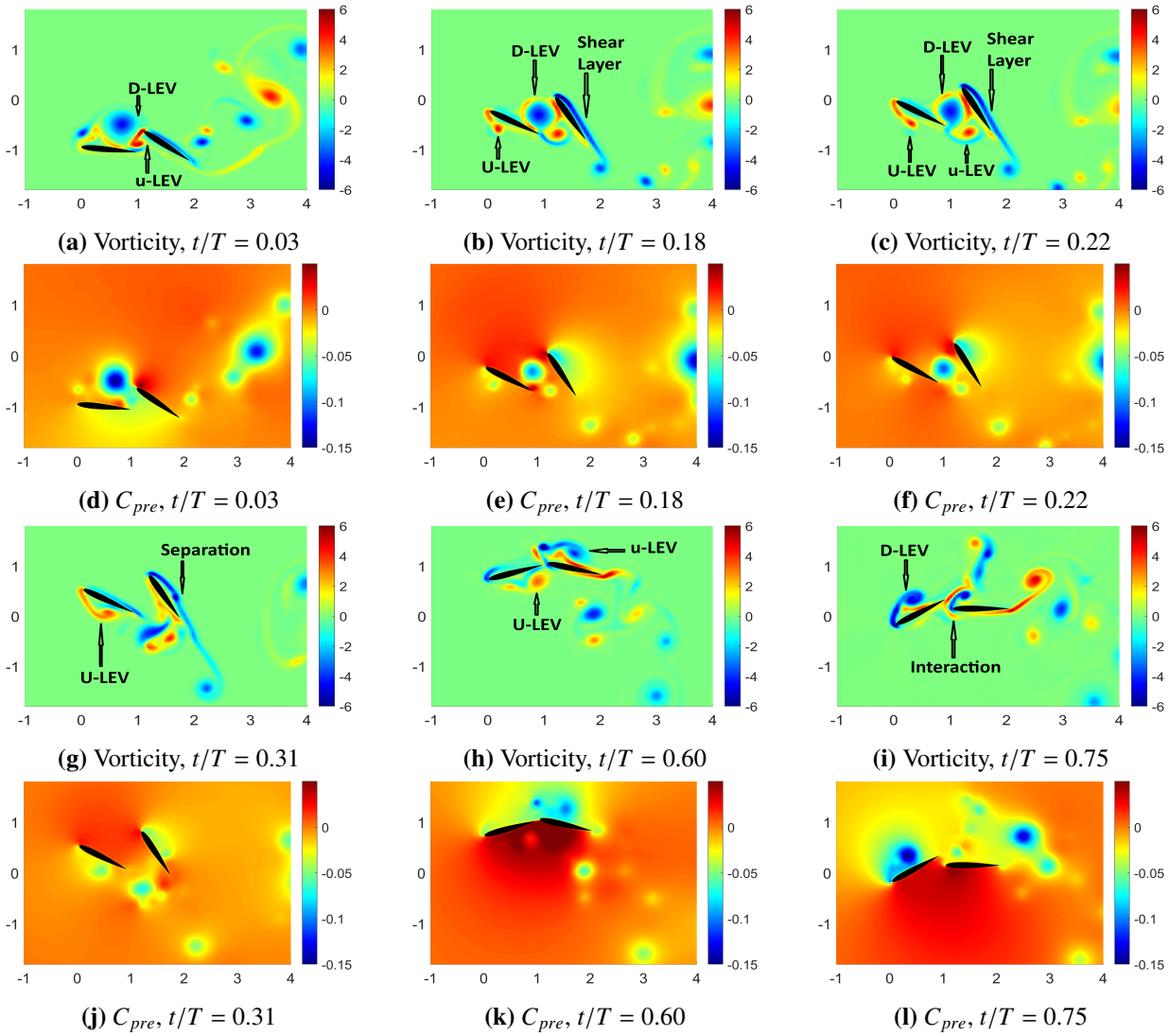
(b) C_T vs t/T

(c) C_P vs t/T

Figure 6. Performance coefficients for $(0°, 0.15c, 1.05c, 30°)$: (a) C_L, (b) C_T, (c) C_P.

An increase in h was observed to delay an reduce the maximum compound C_L due to the delay in

the inverted cup shape because of additional spacing. This is because, by that time, the u-LEV would have escaped from the top surface of the hindwing, thereby increasing the pressure over it, resulting in a comparatively smaller C_{Lhind}. Hence, $h = 0.15c$ provides the best timing between the inverted cup shape and u-LEV, which along with D-LEV majorly dictate the coefficients for the in-phase case.



(a) Vorticity, $t/T = 0.03$ **(b)** Vorticity, $t/T = 0.18$ **(c)** Vorticity, $t/T = 0.22$

(d) C_{pre}, $t/T = 0.03$ **(e)** C_{pre}, $t/T = 0.18$ **(f)** C_{pre}, $t/T = 0.22$

(g) Vorticity, $t/T = 0.31$ **(h)** Vorticity, $t/T = 0.60$ **(i)** Vorticity, $t/T = 0.75$

(j) C_{pre}, $t/T = 0.31$ **(k)** C_{pre}, $t/T = 0.60$ **(l)** C_{pre}, $t/T = 0.75$

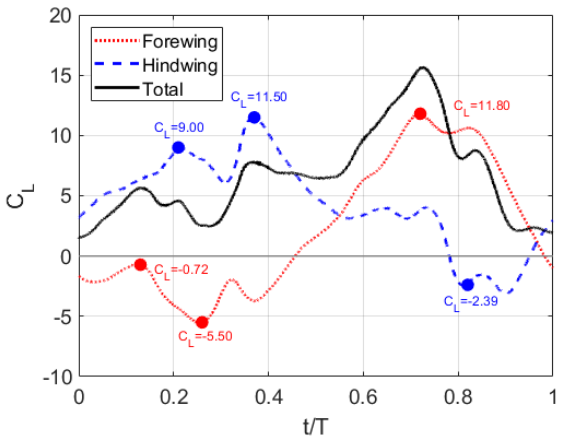
Figure 7. Vorticity and C_{pre} fields (0°,0.15c,1.05c,30°).

3.3. Out-of-phase optimal case-(177.2°,0.39c,1.26c,23.5°)

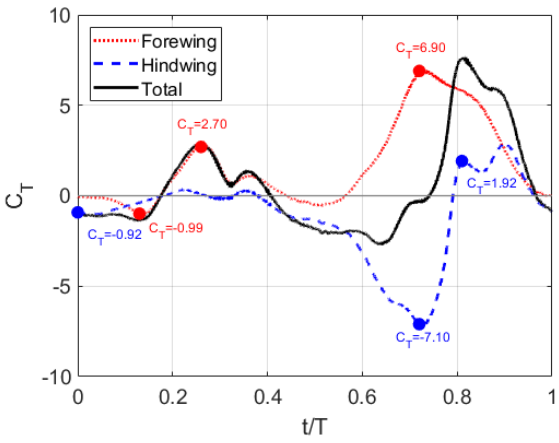
For the out-of-phase case, the C_L, C_T, and C_P can be seen in Figure 8, respectively. Whereas, the coefficient of pressure and vorticity flow fields for the important points can be seen in Figure 9.

At $t/T = 0$, both the forewing and hindwing are present at opposite extremes. The D-LEV separates and escapes much earlier than the in-phase case resulting in a pressure increase on top of the forewing reducing C_{Lfor}. On the other hand, C_{Tfor} and C_{Pfor} are both zero as the forewing is at a 0° AoA with no vortex interaction. The same goes for C_{Thin} and C_{Phin}; however, contrary to C_{Lfor}, C_{Lhin} gradually increases as the downstroke starts. At $t/T = 0.13$, a small bump can be seen for C_{Lfor}, C_{Tfor}, and C_{Pfor}

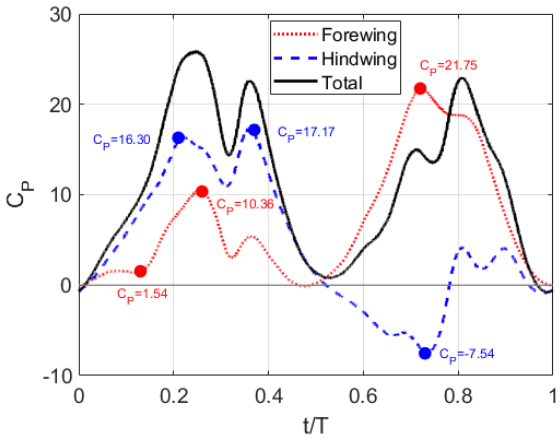
which is caused by the interaction of the previously separated D-LEV due to the forewing upstroke. Afterwards, C_{Lfor} and C_{Lhin} decrease and increase respectively till $t/T = 0.25$ where the strong low-pressure region below the LE of the forewing produces a reasonable C_{Tfor}. Another bump for all the parameters can be seen between $t/T = 0.31$ and $t/T = 0.37$ which is caused by the interaction between the d-LEV and the U-TEV while the airfoils crossed each other. This intensified the low-pressure region below the forewing and on top of the hindwing which led to a decrease and an increase in C_{Lfor} and C_{Lhin} respectively. Both C_{Tfor} and C_{Thin} reached their maximum and minimum value at $t/T = 0.72$. This is because the D-LEV produces a strong low-pressure region at the LE of the forewing as it reaches its minimum AoA. On the other hand, the hindwing approaches its maximum AoA where along with the u-LEV, a strong low-pressure region is created on top and back of it, sucking it back and producing the minimum C_{Thin}. Furthermore, this low-pressure region is sucking the hindwing along the direction of both pitching and heaving which results in C_{Phin} being minimum and negative. This means that the fluid is causing the motion. The second airfoil crossing allows for an interaction between the D-TEV and u-LEV at $t/T = 0.81$, however, it is not as strong as the first interaction as the vortices are farther apart.



(a) C_L vs t/T



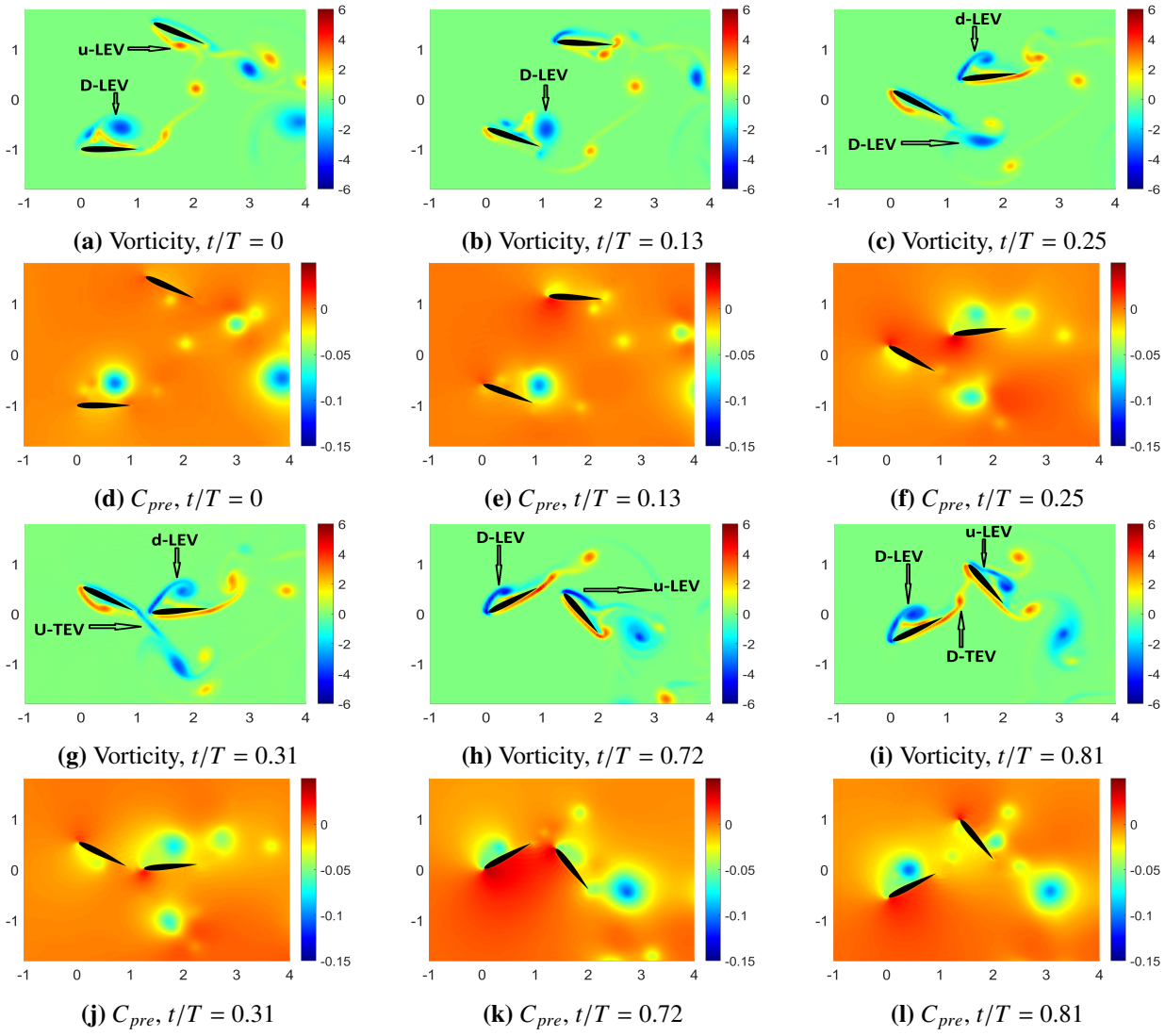
(b) C_T vs t/T



(c) C_P vs t/T

Figure 8. Performance coefficients for (177.2°, 0.39c, 1.26c, 23.5°): (a) C_L, (b) C_T, (c) C_P.

For this out-of-phase case, slight changes in parameters would lead to a decrease in performance. Increasing θ would lead to a negative C_T, whereas, increasing ϕ to 180° would lead to a decrease in C_L as the destructive interference increases. A decrease in d which mostly increases performance, would lead to a significant drop in it to 0.50 for $d = 1.05c$. The main reason for this drop is that C_{Phin} increases significantly because the vortex interaction between D-TEV and u-LEV at $t/T = 0.81$ gets much stronger which creates a stronger low-pressure region below the hindwing. This also leads to a slightly higher C_{Thin} but a lower C_{Lhin}. Hence, the D-TEV and u-LEV interaction majorly dictates the coefficients for the out-of-phase case.



(a) Vorticity, $t/T = 0$ **(b)** Vorticity, $t/T = 0.13$ **(c)** Vorticity, $t/T = 0.25$

(d) C_{pre}, $t/T = 0$ **(e)** C_{pre}, $t/T = 0.13$ **(f)** C_{pre}, $t/T = 0.25$

(g) Vorticity, $t/T = 0.31$ **(h)** Vorticity, $t/T = 0.72$ **(i)** Vorticity, $t/T = 0.81$

(j) C_{pre}, $t/T = 0.31$ **(k)** C_{pre}, $t/T = 0.72$ **(l)** C_{pre}, $t/T = 0.81$

Figure 9. Vorticity and C_{pre} fields (177.2°,0.39c,1.26c,23.5°).

3.4. Model testing on S-1020

The trained model was tested on a non-symmetrical airfoil Selig S-1020, which produces reasonably high C_T and C_L with decent efficiency [39]. The agent was able to find both the locally optimal in-phase and out-of-phase cases with a performance of 0.71 and 0.72, respectively, within 16 steps which

took around 2.5 hours as can be seen in Figure 10. If allowed to explore more, the agent has the ability to find cases with slightly higher performance, similar to NACA 0012. This shows the ability of the agent to generalize as it can be used for different environments with different airfoils which other optimization methods lack.

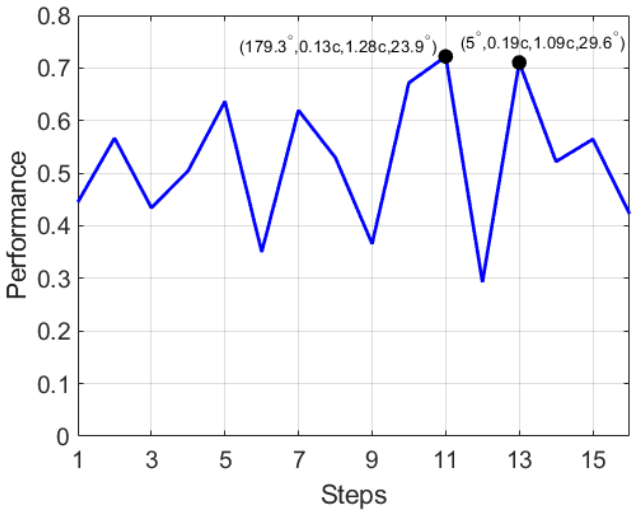


Figure 10. Agent exploration for S1020.

3.5. Flapping wing for Ingenuity

With the optimized parameters provided by RL, it is possible to implement them on the Ingenuity helicopter. However, it has to be kept in mind that the imminent increase in the aerodynamic parameters due to wing-wake interaction in the 2D case significantly decreases for the 3D case unless the aspect ratio (AR) and Rosby number (R_o) employed are greater than 4 [65]. To quantify these effects, some calculations on the optimal Case 2 in Table 3 were conducted using the low Re flow results from [65] and can be seen in Appendix B. The calculations resulted in a 3D C_L of about 5.83, sufficient for hover and take-off to 10 m height in 7.5 s on Mars. Additionally, C_T comes down to 1.11 which allows the FAV to travel about 10 m in 5 s. A speed of about 4 m/s could also be achieved within 5 s, which is comparable to Ingenuity's average speed [66]. On the other hand, the out-of-phase Case 3 in Table 3 produces negligible thrust and can be used for hovering.

4. Conclusions

This report successfully presents and implements the methodology aimed at optimising the compound lift and thrust performance of two tandem pitching and heaving NACA 0012 airfoils in the Martian environment using RL by controlling the horizontal and vertical spacing, phase shift and mean AoA for the hindwing. Furthermore, the trained RL agent has demonstrated applicability for different airfoils which shows the ability of the agent to generalize, which the other methods lack. The results imply that the best performance is for the in-phase and out-of-phase motions with the in-phase motions producing higher C_T, C_L, and C_P. The forewing downstroke LEV and the timing between the inverted cup shape and hindwing upstroke LEV were observed to be mostly dictating the aerodynamic coefficients for the in-phase motions, whereas, the interaction between the forewing and hindwing

vortices as they crossed each other was mostly dictating the aerodynamic coefficients for the out-of-phase motions. For most cases, a decrease in horizontal spacing and an increase in mean hindwing AoA led to an increase in performance and vice versa. However, mostly for the out-of-phase cases, the maximums are reached for their intermediate values, after which, further change leads to a decrease in performance. The effect of vertical spacing depends significantly on other parameters; however, its positive values resulted in better performance compared to the negative values. The most optimal case obtained is the in-phase case with the airfoils closest together and the hindwing's mean AoA equal to 30° with a vertical spacing of 0.15c. An out-of-phase case, ideal for hovering with lesser but similar performance producing negligible C_T and adequate C_L was also found with intermediate values for mean AoA, horizontal, and vertical spacing.

Appendix

A. Motion

Figure A1 shows a typical motion of airfoils over one oscillation for $\phi = 180°$, $h = 0c$, $d = 2.25c$, $\theta = 20°$.

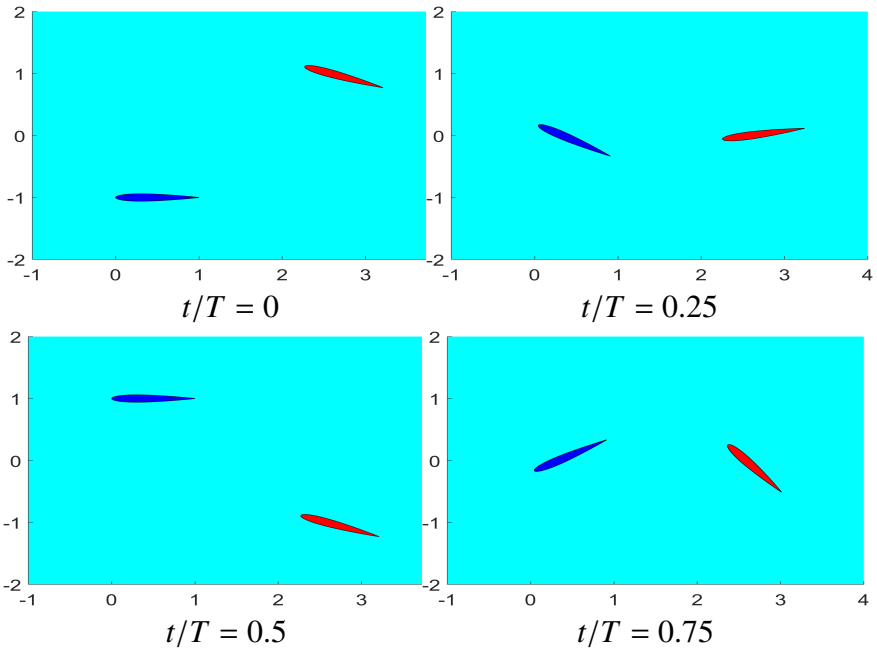


Figure A1. Airfoil motion over one oscillation for $\phi = 180°$, $h = 0c$, $d = 2.25c$, $\theta = 20°$.

B. Mars atmosphere and Ingenuity wing change calculations

In this section, reasonable approximations will be made to assess if the obtained results could potentially be used to design a Mars FAV. Hence, the Ingenuity helicopter wing will be replaced with a pair of tandem flapping wings. The atmospheric and blade properties used are mentioned in Table B1. It is worth mentioning that the 3D wing shape used is a simple rectangular one which may not lead to the best results and can be optimized significantly but provides a reasonable start. For hovering, the lift is equal to the weight, and hence the $C_{L_{hov}}$ required is as follows:

$$C_{L_{hov}} = \frac{Weight}{\frac{1}{2}\rho u_o^2 S} = \frac{1.8 \times 3.72}{\frac{1}{2} \times 1.55 \times 10^{-2} \times 22^2 \times 0.17 \times 2} = 5.25. \tag{B1}$$

The AR of the wing according to the convention followed by Lee and Lua [65] can be calculated using the following formula:

$$AR = \frac{b^2}{S} = \frac{(0.605)^2}{(0.17/2)} = 4.3 > 4. \tag{B2}$$

On the other hand, the 3D effect is quantified by the Rossby number (R_o). R_o for a rectangular flapping wing is calculated by the following equation [65]:

$$R_o = \sqrt{\frac{\left(\frac{\Delta R}{c} + AR\right)^3 - \left(\frac{\Delta R}{c}\right)^3}{3AR}} = \sqrt{\frac{(-0.333 + 4.3)^3 - (-0.333)^3}{3 \times 4.3}} = 2.2 < 4, \tag{B3}$$

where ΔR is the distance between the wing pivot axis and LE which is set to $c/3$.

From Figure 3 in Lee and Lua [65], for a R_o of ∞ and AR of 4.3, the C_L decreases by about 50% due to the effects of the aspect ratio caused by the vortex shedding at the tips. On the other hand, from Figure 8 in Lee and Lua [65], the C_L also increases by about 15% once the effect of R_o which is 2.2 is added back due to the effects of vortex interactions. This results in a final equation for C_L from the 2D case to 3D with the current AR and R_o to be,

$$C_{L3D} = 0.5 \times C_{L2D} \times 1.15 = 0.575 \times C_{L2D}. \tag{B4}$$

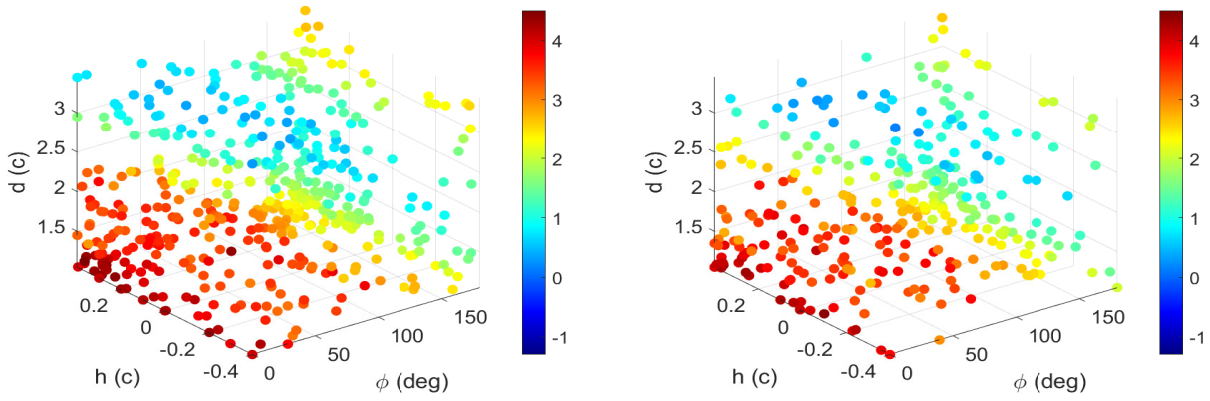
The most optimal case for the 2D flapping wings for Case 2 in Table 3 produces a C_L of 10.15 and C_T of 1.93. Hence, using Eq (B4) results in a 3D C_L of about 5.83 which is sufficient for hover and take-off to 10 m height in 7.5 s on Mars. Additionally, if C_T is also assumed to follow the same equation as C_L, it comes down to 1.11 which allows the FAV to travel about 10 m in 5 s. A speed of about 4 m/s could also be achieved within 5 s which is comparable to Ingenuity's average speed [66]. It is also worth mentioning that the R_o for this case is less than 4 and there is significant room for improvement for designing a 3D wing with better performance.

Table B1. Properties of Martian atmosphere and ingenuity measurements.

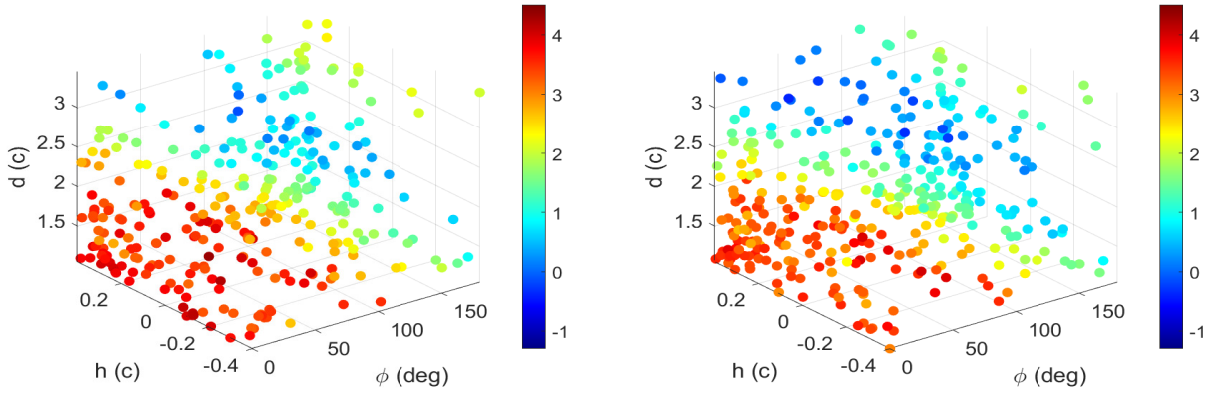
Description	Value
Gravitational acceleration (m/s^2) [67]	3.72
Atmospheric density (kg/m^3) [67]	1.55×10^{-2}
Dynamic viscosity (kg/ms) [67]	1.5×10^{-5}
Blade area (m^2) [2]	0.17
Blade radius (m) [2]	0.605
Mass (kg) [2]	1.8

C. Scatter plots for compound C_T, C_L, C_P, and Performance

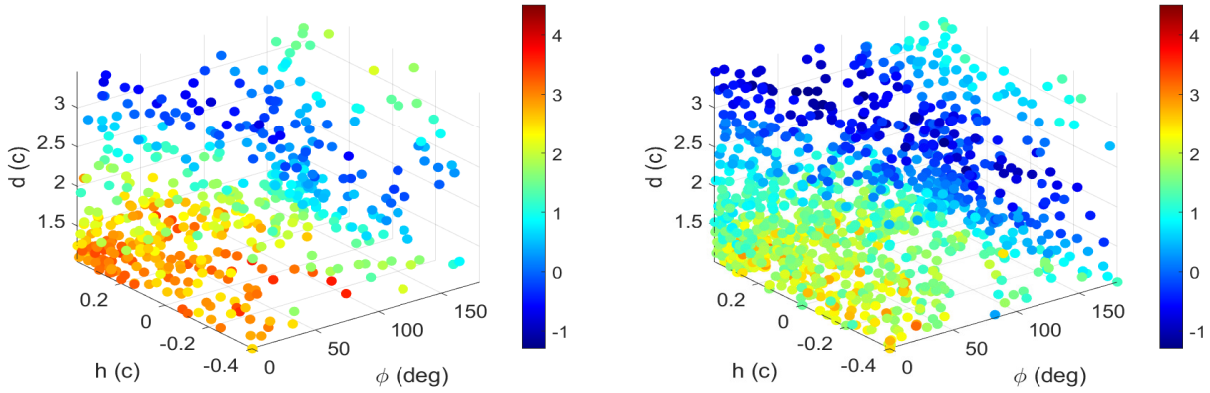
Figures C1–C4 show C_T, C_L, C_P, and performance of the foils.



(a) $0° < \alpha < 5°$ **(b)** $5° < \alpha < 10°$

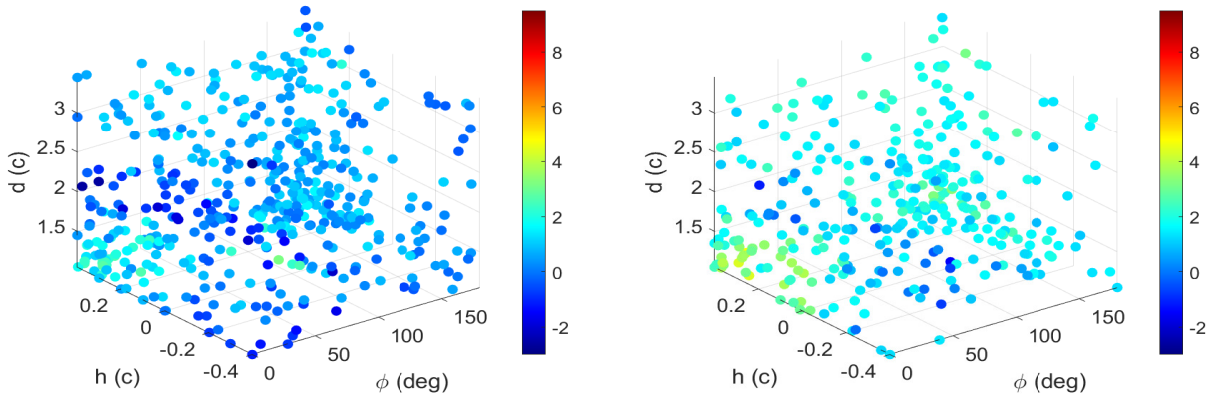


(c) $10° < \alpha < 15°$ **(d)** $15° < \alpha < 20°$

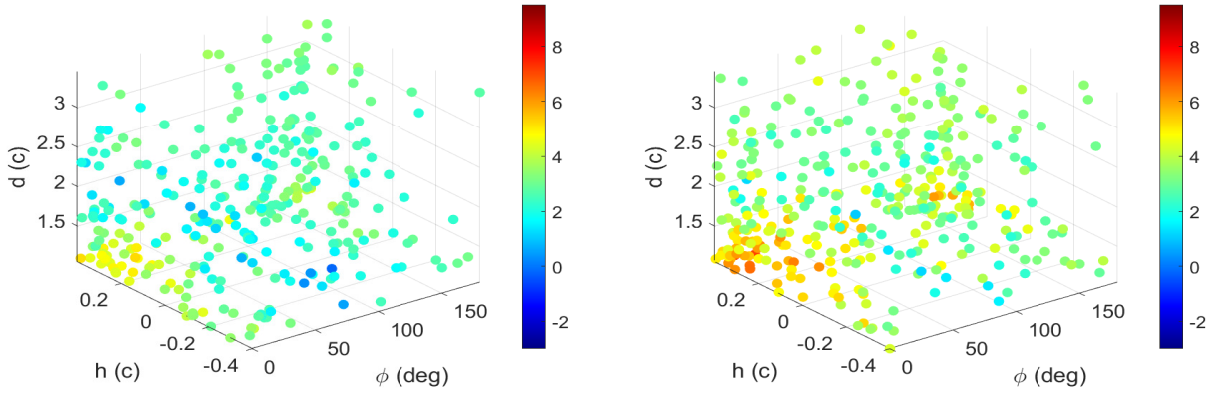


(e) $20° < \alpha < 25°$ **(f)** $25° < \alpha < 30°$

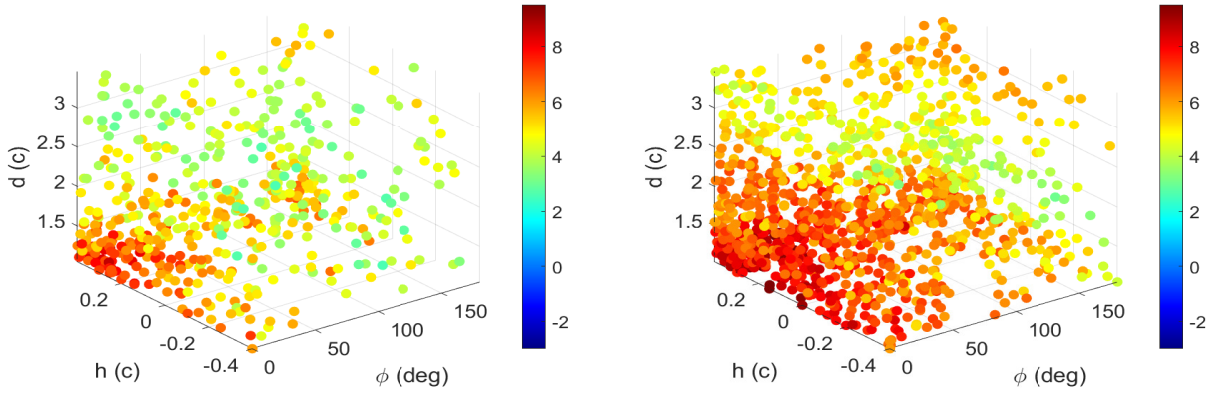
Figure C1. C_T scatter plots for different α ranges.



(a) $0° < \alpha < 5°$ **(b)** $5° < \alpha < 10°$

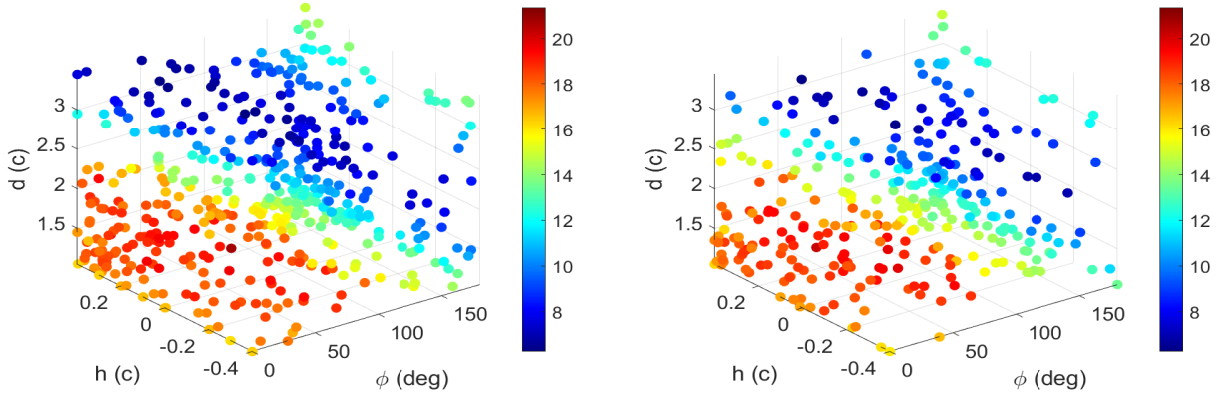


(c) $10° < \alpha < 15°$ **(d)** $15° < \alpha < 20°$

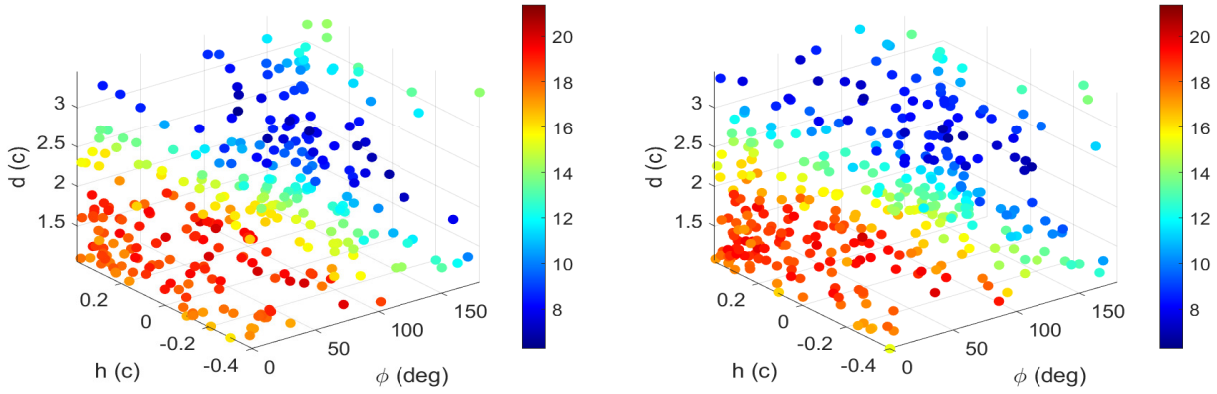


(e) $20° < \alpha < 25°$ **(f)** $25° < \alpha < 30°$

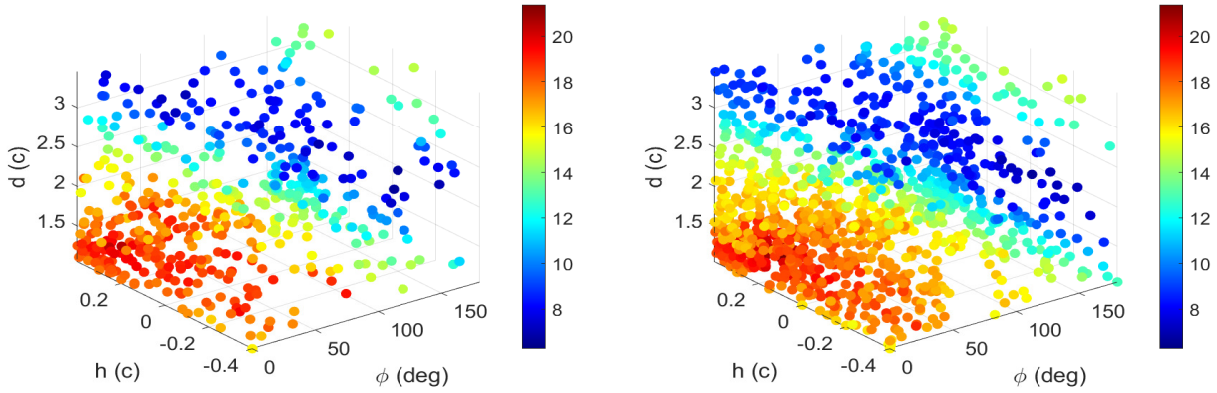
Figure C2. C_L scatter plots for different α ranges.



(a) $0° < \alpha < 5°$ **(b)** $5° < \alpha < 10°$

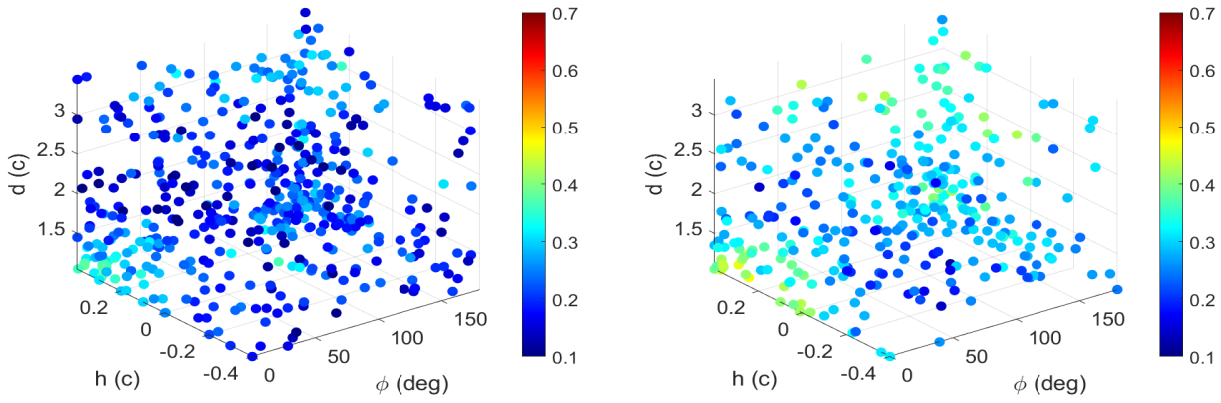


(c) $10° < \alpha < 15°$ **(d)** $15° < \alpha < 20°$

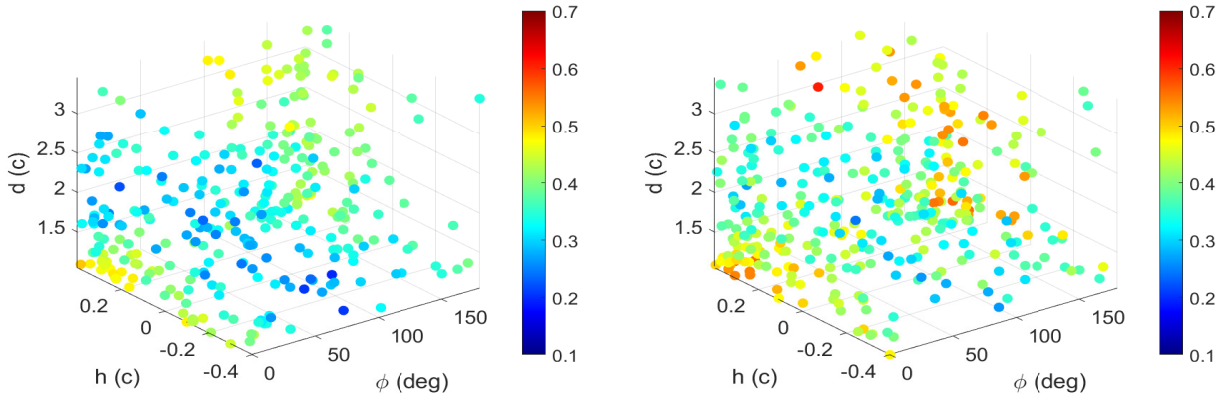


(e) $20° < \alpha < 25°$ **(f)** $25° < \alpha < 30°$

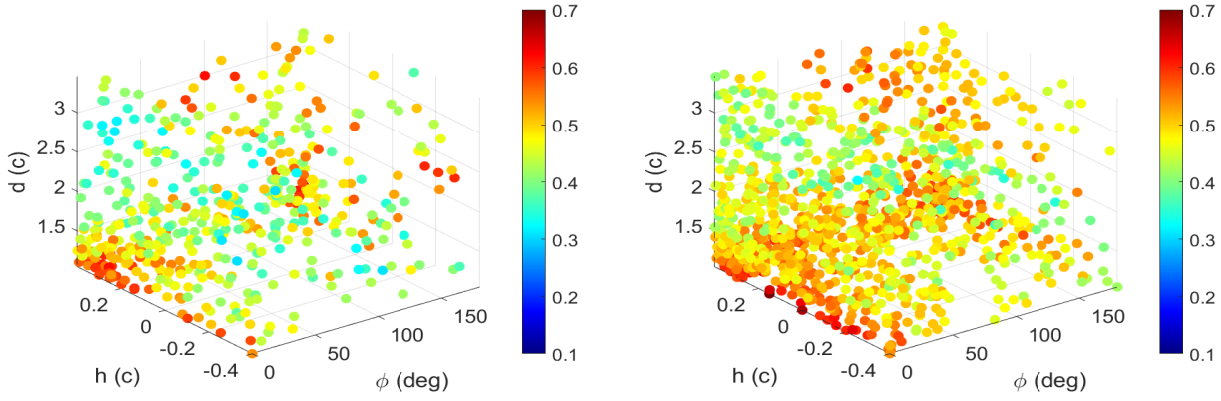
Figure C3. C_P scatter plots for different α ranges.



(a) $0° < \alpha < 5°$ **(b)** $5° < \alpha < 10°$



(c) $10° < \alpha < 15°$ **(d)** $15° < \alpha < 20°$



(e) $20° < \alpha < 25°$ **(f)** $25° < \alpha < 30°$

Figure C4. Performance scatter plots for different α ranges.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgements

This work was supported by the Australian Research Council (grant number DP240100294) and Australian Composites Manufacturing Cooperative Research Centre (APN009), and was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

Conflict of interest

The authors declare there are no conflicts of interest.

References

1. T. Suwa, K. Nose, D. Numata, H. Nagai, K. Asai, Compressibility effects on airfoil aerodynamics at low Reynolds number, in *30th AIAA Applied Aerodynamics Conference*, (2012), 3029. https://doi.org/10.2514/6.2012-3029
2. W. J. Koning, M. Dominguez, *Mars Helicopter Ingenuity Rotor Geometry*, Technical report, Ames Research Center, Moffett Field, California, 2024.
3. J. D. Eldredge, A. R. Jones, Leading-edge vortices: mechanics and modeling, *Annu. Rev. Fluid Mech.*, **51** (2019), 75–104. https://doi.org/10.1146/annurev-fluid-010518-040334
4. N. Widdup, L. Wang, J. Young, V. Daria, H. Liu, F. B. Tian, A scaling law for the lift of a bio-inspired wing hovering in low-density compressible flows, *J. Fluid Mech.*, **1003** (2025), A10. https://doi.org/10.1017/jfm.2024.1216
5. U. Pesavento, Z. J. Wang, Flapping wing flight can save aerodynamic power compared to steady flight, *Phys. Rev. Lett.*, **103** (2009), 118102. https://doi.org/10.1103/PhysRevLett.103.118102
6. R. J. Huston, *Wind-tunnel measurements of performance, blade motions, and blade air loads for tandem-rotor configurations with and without overlap*, Technical report, 1963.
7. P. J. Arcidiacono, A method for computation of the induced velocity field of a rotor in forward flight, suitable for application to tandem rotor configurations, *J. Am. Helicopter Soc.*, **9** (1964), 34–45. https://doi.org/10.4050/JAHS.9.34
8. D. R. Croom, H. G. Wiley, *Investigation at transonic speeds of the hinge-moment and lift-effectiveness characteristics of a single flap and a tandem flap on a 60 degree delta wing*, Technical report, 1953.
9. V. R. Corsiglia, D. Koenig, *Large-scale wind-tunnel tests on an aspect ratio 2.17 delta-wing model equipped with midchord boundary-layer-control flaps*, Technical report, 1964.
10. T. Broering, Y. Lian, The effect of wing spacing on tandem wing aerodynamics, in *28th AIAA Applied Aerodynamics Conference*, (2010), 4385. https://doi.org/10.2514/6.2010-4385
11. T. M. Broering, Y. S. Lian, The effect of phase angle and wing spacing on tandem flapping wings, *Acta Mech. Sin.*, **28** (2012), 1557–1571. https://doi.org/10.1007/s10409-012-0210-8

12. T. M. Broering, Y. Lian, Numerical study of tandem flapping wing aerodynamics in both two and three dimensions, *Comput. Fluids*, **115** (2015), 124–139. https://doi.org/10.1016/j.compfluid.2015.04.003

13. A. S. A. Seet, *Investigation of Flow Over Oscillating NACA 4421 Airfoils in Tandem Configuration at Low Reynolds Number*, PhD thesis, Nanyang Technological University in Singapore, 2019.

14. T. Lee, Flow past two in-tandem airfoils undergoing sinusoidal oscillations, *Exp. Fluids*, **51** (2011), 1605–1621. https://doi.org/10.1007/s00348-011-1173-4

15. A. Jaroszewicz, J. Sąsiadek, K. Sibilski, Modeling and simulation of flapping wings entomopter in Martian atmosphere, in *Aerospace Robotics*, Springer, (2013), 143–162. https://doi.org/10.1007/978-3-642-34020-8_12

16. N. P. B. Mannam, M. M. Alam, P. Krishnankutty, Review of biomimetic flexible flapping foil propulsion systems on different planetary bodies, *Results Eng.*, **8** (2020), 100183. https://doi.org/10.1016/j.rineng.2020.100183

17. J. E. Bluman, J. A. Pohly, M. K. Sridhar, C. K. Kang, D. B. Landrum, F. Fahimi, et al., Achieving bioinspired flapping wing hovering flight solutions on Mars via wing scaling, *Bioinspir. Biomim.*, **13** (2018), 046010. https://doi.org/10.1088/1748-3190/aac876

18. H. Bézard, T. Désert, T. Jardin, J. M. Moschetta, Numerical and experimental aerodynamic investigation of a micro-uav for flying on mars, in *76th Annual Forum & Technology Display*, 2020.

19. J. L. McCain, Experimental force and wing motion measurements of a bioinspired flapping wing in a martian density condition, 2019. Available from: https://louis.uah.edu/uah-theses/306

20. B. Singh, N. Yidris, A. A. Basri, R. Pai, K. A. Ahmad, Study of mosquito aerodynamics for imitation as a small robot and flight in a low-density environment, *Micromachines*, **12** (2021), 511. https://doi.org/10.3390/mi12050511

21. N. Widdup, L. Wang, J. Young, F. B. Tian, A numerical study of flexible flapping wings in highly compressible flows, *Phys. Fluids*, **37** (2025), 071912. https://doi.org/10.1063/5.0274245

22. N. Widdup, L. Wang, J. Young, F. B. Tian, The aerodynamic and acoustic performance of a bumblebee-inspired micro-air vehicle in Martian atmosphere, *Phys. Fluids*, **37** (2025), 091916. https://doi.org/10.1063/5.0288043

23. S. S. Bhat, T. O'Shaughnessy, J. T. Hrynuk, J. Young, F. B. Tian, S. Ravi, Propeller blade performance in low-pressure environments, *Phys. Fluids*, **38** (2026), In press. https://doi.org/10.1063/5.0311082

24. M. Anyoji, M. Okamoto, H. Hidaka, T. Nonomura, A. Oyama, K. Fujii, Planetary atmosphere wind tunnel tests on aerodynamic characteristics of a mars airplane scale model, *Trans. JSASS Aerospace Tech. Japan*, **12** (2014), 7–12. https://doi.org/10.2322/tastj.12.Pk_7

25. J. A. Pohly, C. K. Kang, D. B. Landrum, J. E. Bluman, H. Aono, Data-driven CFD scaling of bioinspired Mars flight vehicles for hover, *Acta Astronaut.*, **180** (2021), 545–559. https://doi.org/10.1016/j.actaastro.2020.12.037

26. H. Mahjoubi, K. Byl, Steering and horizontal motion control in insect-inspired flapping-wing MAVs: the tunable impedance approach, in *2012 American Control Conference (ACC)*, (2012), 901–908. https://doi.org/10.1109/ACC.2012.6314655

27. H. Mahjoubi, K. Byl, Analysis of a tunable impedance method for practical control of insect-inspired flapping-wing MAVs, in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, (2011), 3539–3546. https://doi.org/10.1109/CDC.2011.6160297

28. P. A. Oche, G. A. Ewa, N. Ibekwe, Applications and challenges of artificial intelligence in space missions, *IEEE Access*, **12** (2024), 44481–44509. https://doi.org/10.1109/ACCESS.2021.3132500

29. D. Soler, O. Mariño, D. Huergo, M. de Frutos, E. Ferrer, Reinforcement learning to maximize wind turbine energy generation, *Expert Syst. Appl.*, **249** (2024), 123502. https://doi.org/10.1016/j.eswa.2024.123502

30. H. Van Hasselt, A. Guez, D. Silver, Deep reinforcement learning with double q-learning, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **30** (2016), 2094–2100. https://doi.org/10.1609/aaai.v30i1.10295

31. K. Portal-Porras, U. Fernandez-Gamiz, E. Zulueta, R. Garcia-Fernandez, S. E. Berrizbeitia, Active flow control on airfoils by reinforcement learning, *Ocean Eng.*, **287** (2023), 115775. https://doi.org/10.1016/j.oceaneng.2023.115775

32. Y. Bao, X. Shi, Z. Wang, H. Zhu, N. Srinil, A. Li, et al., Deep reinforcement learning for propulsive performance of a flapping foil, *Phys. Fluids*, **35** (2023), 103610. https://doi.org/10.1063/5.0169982

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

34. G. Lee, Y. Joo, S. U. Lee, T. Kim, Y. Yu, H. G. Kim, Design optimization of heat exchanger using deep reinforcement learning, *Int. Commun. Heat Mass Transfer*, **159** (2024), 107991. https://doi.org/10.1016/j.icheatmasstransfer.2024.107991

35. T. Ji, F. Jin, F. Xie, H. Zheng, X. Zhang, Y. Zheng, Active learning of tandem flapping wings at optimizing propulsion performance, *Phys. Fluids*, **34** (2022), 047117. https://doi.org/10.1063/5.0084160

36. P. I. Frazier, A tutorial on bayesian optimization, preprint, arXiv:1807.02811.

37. N. P. B. Mannam, V. R. Avula, Propulsive performance of tandem flapping wings for autonomous underwater vehicles (AUVs) and surface ships, in *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)*, (2021), 1–7. https://doi.org/10.1109/GUCON50781.2021.9573893

38. H. Nagai, K. Fujita, M. Murozono, Experimental study on forewing–hindwing phasing in hovering and forward flapping flight, *AIAA J.*, **57** (2019), 3779–3790. https://doi.org/10.2514/1.J058335

39. W. B. Tay, K. B. Lim, Analysis of non-symmetrical flapping airfoils, *Acta Mech. Sin.*, **25** (2009), 433–450. https://doi.org/10.1007/s10409-009-0259-1

40. L. Wang, F. B. Tian, Numerical study of sound generation by three-dimensional flexible flapping wings during hovering flight, *J. Fluids Struct.*, **99** (2020), 103165. https://doi.org/10.1016/j.jfluidstructs.2020.103165
41. L. Wang, F. B. Tian, Sound generated by the flow around an airfoil with an attached flap: From passive fluid–structure interaction to active control, *J. Fluids Struct.*, **111** (2022), 103571. https://doi.org/10.1016/j.jfluidstructs.2022.103571
42. K. Isogai, Y. Shinmoto, Y. Watanabe, Effects of dynamic stall on propulsive efficiency and thrust of flapping airfoil, *AIAAJ*, **37** (1999), 1145–1151. https://doi.org/10.2514/2.589
43. L. Wang, F. B. Tian, J. C. Lai, An immersed boundary method for fluid–structure–acoustics interactions involving large deformations and complex geometries, *J. Fluids Struct.*, **95** (2020), 102993. https://doi.org/10.1016/j.jfluidstructs.2020.102993
44. L. Wang, J. Young, F. B. Tian, An immersed boundary method for the thermo–fluid–structure interaction in rarefied gas flows, *Phys. Fluids*, **36** (2024), 013616. https://doi.org/10.1063/5.0181397
45. X. D. Liu, S. Osher, T. Chan, Weighted essentially non-oscillatory schemes, *J. Comput. Phys.*, **115** (1994), 200–212. https://doi.org/10.1006/jcph.1994.1187
46. L. Fu, X. Y. Hu, N. A. Adams, A family of high-order targeted ENO schemes for compressible-fluid simulations, *J. Comput. Phys.*, **305** (2016), 333–359. https://doi.org/10.1016/j.jcp.2015.10.037
47. F. B. Tian, X. Lu, H. Luo, Onset of instability of a flag in uniform flow, *Theor. Appl. Mech. Lett.*, **2** (2012), 022005. https://doi.org/10.1063/2.1202205
48. F. B. Tian, W. Wang, J. Wu, Y. Sui, Swimming performance and vorticity structures of a mother-calf pair of fish, *Comput. Fluids*, **124** (2016), 1–11. https://doi.org/10.1016/j.compfluid.2015.10.006
49. W. X. Huang, F. B. Tian, Recent trends and progress in the immersed boundary method, *Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci.*, **233** (2019), 7617–7636. https://doi.org/10.1177/0954406219842606
50. L. Wang, F. B. Tian, H. Liu, Numerical study of three-dimensional flapping wings hovering in ultra-low-density atmosphere, *Phys. Fluids*, **34** (2022), 041903. https://doi.org/10.1063/5.0085021
51. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms, preprint, arXiv:1707.06347.
52. H. K. Lim, J. B. Kim, J. S. Heo, Y. H. Han, Federated reinforcement learning for training control policies on multiple IoT devices, *Sensors*, **20** (2020), 1359. https://doi.org/10.3390/s20051359
53. K. P. Haughn, L. L. Gamble, D. J. Inman, Deep reinforcement learning achieves multifunctional morphing airfoil control, *J. Compos. Mater.*, **57** (2023), 721–736. https://doi.org/10.1177/00219983221137644
54. S. Ravichandiran, *Deep Reinforcement Learning with Python: Master Classic RL, Deep RL, Distributional RL, Inverse RL, and more with OpenAI Gym and TensorFlow*, Packt Publishing Ltd, 2020.

55. V. Nair, G. E. Hinton, Rectified linear units improve restricted Boltzmann machines, in *Proceedings of the 27th International Conference on Machine Learning*, (2010), 807–814.
56. N. Fuengfusin, H. Tamukoh, Nan attacks: Bit-flipping deep neural network parameters to nan or infinity, in *024 1st International Conference on Robotics, Engineering, Science, and Technology (RESTCON)*, (2024), 33–37. https://doi.org/10.1109/RESTCON60981.2024.10463548
57. D. Liu, X. Xiong, Y. Zhang, Action-dependent adaptive critic designs, in *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, **2** (2001), 990–995. https://doi.org/10.1109/IJCNN.2001.939495
58. G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, et al., Openai gym, preprint, arXiv:1606.01540.
59. M. Andrychowicz, A. Raichuk, P. Stańczyk, M. Orsini, S. Girgin, R. Marinier, et al., What matters in on-policy reinforcement learning? a large-scale empirical study, preprint, arXiv:2006.05990.
60. M. Schilling, Avoid overfitting in deep reinforcement learning: Increasing robustness through decentralized control, in *Artificial Neural Networks and Machine Learning – ICANN 2021*, (2021), 638–649. https://doi.org/10.1007/978-3-030-86380-7_52
61. D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, M. Riedmiller, Deterministic policy gradient algorithms, in *International Conference on Machine Learning*, (2014), 387–395.
62. M. Moriche, O. Flores, M. García-Villalba, Generation of thrust and lift with airfoils in plunging and pitching motion, *J. Phys. Conf. Ser.*, **574** (2015), 012163. https://doi.org/10.1088/1742-6596/574/1/012163
63. J. Zhang, X. Y. Lu, Aerodynamic performance due to forewing and hindwing interaction in gliding dragonfly flight, *Phys. Rev. E*, **80** (2009), 017302. https://doi.org/10.1103/PhysRevE.80.017302
64. D. Bie, D. Li, Numerical analysis of the wing–wake interaction of tandem flapping wings in forward flight, *Aerosp. Sci. Technol.*, **121** (2022), 107389. https://doi.org/10.1016/j.ast.2022.107389
65. Y. Lee, K. B. Lua, Wing–wake interaction: comparison of 2D and 3D flapping wings in hover flight, *Bioinspir. Biomim.*, **13** (2018), 066003. https://doi.org/10.1088/1748-3190/aadc31
66. O. W. Ata, Aerodynamic performance of advanced ingenuity and dragonfly drones for future space missions to mars and titan, in *2021 4th International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, (2021), 01–06. https://doi.org/10.1109/ISAECT53699.2021.9668399
67. Z. Terze, V. Pandža, M. Kasalo, D. Zlatar, Discrete mechanics and optimal control optimization of flapping wing dynamics for Mars exploration, *Aerosp. Sci. Technol.*, **106** (2020), 106131. https://doi.org/10.1016/j.ast.2020.106131