**33RD CONGRESS
OF THE INTERNATIONAL COUNCIL
OF THE AERONAUTICAL SCIENCES
STOCKHOLM, SWEDEN, 4–9 SEPTEMBER, 2022**

ICAS
2022
SWEDEN

MULTI-OBJECTIVE AERODYNAMIC OPTIMIZATION OF 2D HIGH-LIFT DEVICE BASED ON DISTRIBUTED DEEP REINFORCEMENT LEARNING

Jiahua Dai[1], Peiqing Liu[2], Ling Li[3]

[1]School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, People's Republic of China
[2]School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, People's Republic of China
[3]School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, People's Republic of China

Abstract

The design of high-lift device (HLD) is of great significance for large aircraft, which can effectively improve the take-off, landing performance and safety. The aerodynamic design of HLD is a multi-objective optimization problem which consumes a lot of computing resources. In our previous work, we proposed a method based on deep reinforcement learning (DRL) and transfer learning (TL), which transfers the experience of two-dimensional (2D) optimization to three-dimensional (3D) optimization. However, for the multi-objective optimization problem, this method still merges several objectives into a single objective through linear weighted superposition, which is mainly due to the fact that the loss function of neural network and the reward function of Reinforcement Learning (RL) must be scalar. In order to obtain the ability to find the complete Pareto solution similar to heuristic algorithms, we further improve the algorithm by combining the idea of distributed reinforcement learning which is still based on the Deep Deterministic Policy Gradient. The new algorithm changes the original single actor into multiple actors, and each actor trials and errors in the environment. A Critic is used to score all the actors' actions and give the correct value gradient. Three reward functions are compared and the function based on Hypervolume is better. The trailing-edge high lift device of a commercial aircraft is optimized and the agent can find a set of relatively complete Pareto solutions.

Keywords: high-lift device, distributed deep reinforcement learning, multi-objective aerodynamic optimization, aircraft design

1. Introduction

As one of the components of commercial aircraft, the high-lift device is of great significance for the safe take-off and landing [1-2]. The flow mechanism of the high-lift device is complex [3], so it is difficult to design the optimal aerodynamic layout by the theoretical methods. In contrast, Computational Fluid Dynamics (CFD) can simulate the complex flow more accurately. In the traditional design method of high-lift device, there are two basic approaches. First, the multi-element airfoil is optimized based on CFD and heuristic algorithm [4-6], and then it is extended to the three-dimensional (3D) layout [7]. This approach requires the designers to have sufficient experience, and the designed 3D layout is often not the best. The 3D optimization of the high-lift device is a more direct approach, but the numerical simulation will cost a large computational cost [8-9]. Due to the limitation of heuristic algorithm, only one of the two approaches can be selected. The agents of Deep Reinforcement Learning (DRL) for decision-making can be pre-trained with moderately inaccurate samples, so as to get around the limitation. This coupling method of DRL and Transfer Learning (TL) can improve the search efficiency of the optimization algorithm and the performance of the optimal solution at the same time. Yan, et al. [10] used this method to improve the performance of a missile. In our previous work [11], a 2D-to-3D method based on DRL and TL is proposed to optimize a 3D high-lift configuration. The results show that the experience of 2D optimization can be extracted by neural network to improve the efficiency of 3D optimization and the performance of the optimal solution.

The optimization of high-lift device is a complex multi-objective problem. Commercial aircrafts have different aerodynamic performance requirements in take-off, climb, approach, landing and other stages [2,12]. For the similar multi-objective optimization problems, instead of finding an optimal solution, the optimization target is to find a set of solutions that are not dominated by each other, which is called Pareto front [13]. The current DRL algorithm is difficult to solve this kind of multi-objective optimization problem due to both the loss function of Deep Learning (DL) and the reward function of RL are scalar. Therefore, it is a common method to convert a multi-objective problem into a single objective problem. Gabor, et al. [14], Manor, et al. [15], Vampler et al. [16] have proposed such methods. However, those kinds of method can only find the partial solutions of the complete Pareto front, compared with the heuristic algorithm[17-18]. The super parameter setting in the algorithm also needs to be given by experienced designers, which improves the difficulty of solving the optimization problem. Moffaert [19] proposed a multi-objective RL method based on Q-Learning. However, the algorithm based on Q-Learning cannot solve the problem with continuous state space and action space.

In this work, a multi-objective optimization algorithm based on DRL is proposed. The trailing-edge high-lift device of a commercial aircraft is used for testing to explore the feasibility of the algorithm being used for optimization.

2. Optimization Algorithm

2.1 Deep Reinforcement Learning

RL is a method in which agents try to change their strategies and make actions through the feedback they receive in order to maximize the benefits [20]. The basic process in a step t of RL can be expressed as equation (1). S_t is often used to represent the current state of the agent in the environment. The action A_t taken by the agent in the state S_t can be expressed as $A_t = \pi(S_t)$. The feedback given to the agent is expressed by the reward function R_t. A value function $Q_t(S_t, A_t)$ is used to evaluate the action A_t, which is influenced by R_t. The strategy π will be modified according to the gradient of the value function $Q_t(S_t, A_t)$.

$$S_t \rightarrow A_t \rightarrow R_t \rightarrow S_{t+1} \tag{1}$$

DRL uses neural networks to represent strategy $\pi(S_t | \boldsymbol{\omega})$ and value function $Q_t(S_t, A_t | \boldsymbol{\theta})$, so as to conduct tasks including continuous state space and action space, where $\boldsymbol{\omega}$ and $\boldsymbol{\theta}$ are the parameters of the neural networks. In the commonly used Actor-Critic (A-C) framework, the strategy network is called the Actor and the value function is called the Critic. Our algorithm is modified from Depth Deterministic Policy Gradient (DDPG) [21], which is based on the A-C framework. DDPG adopts double networks (eval network and target network) and replay buffer to make the convergence process more stable. For multi-objective optimization problem we design the reward function R_t as shown in equation (2).

$$R_t = ND(S_{t+1}, S_t) \times D(S_{t+1}, S_t) \tag{2}$$

The first item, ND operation, judges the dominance of the new state S_{t+1}, which represents the quality of the solution as a coefficient of the reward function, as shown in equation (3). If each target of the new solution is better than the old one (Dominating Solution), we set a larger coefficient of 3; If some of the performances are better than the old solution's, called Non-dominated Solution, a coefficient of 1 is set; If the new solution is not as good as the old solution in all aspects, called Dominated Solution, the coefficient of -3 is set.

$$ND(S_{t+1}, S_t) = \begin{cases} 3, & \text{Dominating Solution} \\ 1, & \text{Non-dominated Solution} \\ -3, & \text{Dominated Solution} \end{cases} \tag{3}$$

The second item, D operation, is used to evaluate the good or bad degree of the new solution, so as to improve the search efficiency of the optimization algorithm. We have adopted three evaluation methods, namely, Merge into Single Target (MS), Euclidean Distance (ED) and Hypervolume (HV),

as shown in equation (4).

$$D(S_{t+1}, S_t) = \begin{cases} MS(S_{t+1}, S_t), & \text{MS-reward} \\ ED(S_{t+1}, S_t), & \text{ED-reward} \\ HV(S_{t+1}, S_t), & \text{HV-reward} \end{cases} \tag{4}$$

Weighted average is chosen for MS operation. For n objects, the specific operation is shown in equation (5), where the ω is the weight.

$$MS(S_{t+1}, S_t) = \sum_{i=1}^{n} \omega_i Object_i^{t+1} - \sum_{i=1}^{n} \omega_i Object_i^t \tag{5}$$

ED operation uses Euclidean distance to describe the similarity between the new solution and the old solution, in order to find a more complete Pareto front.

$$ED(S_{t+1}, S_t) = \sqrt{\sum_{i=1}^{n} \left(Object_i^{t+1} - Object_i^t\right)^2} \tag{6}$$

HV operation is to calculate the increasing effect of the new solution on the Hypervolume. For a double-objective optimization problem, the meaning of Hypervolume is shown in the gray area of the Figure 1, where J_{ref} is the reference point, and J_1 to J_3 represent the solutions on the Pareto front. The method of Fonseca, et al. [22] to calculate the Hypervolume is referenced.
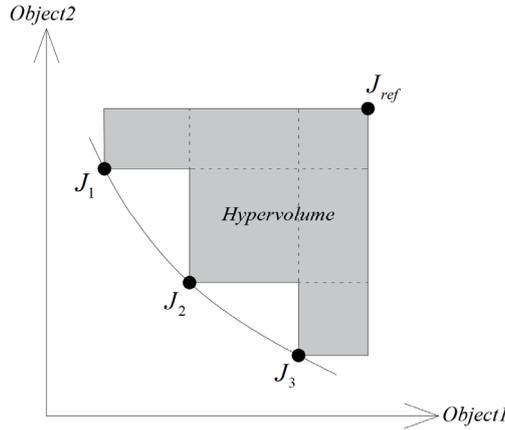
Figure 1 – Illustration of the Hypervolume calculator.

2.2 Distributed Reinforcement Learning

In DDPG algorithm, although the Actor represented by neural network can fit a strategy for searching multiple targets, it will only sample one at a time. In order to further enrich the richness and uniformity of Pareto front, distributed reinforcement learning with multiple actor networks in parallel is adopted. Single Critic is set because the rules of value evaluation for all Actors are the same. A single critic can use the samples of all actors to adjust their strategies and improve the optimization efficiency.

In step t, the flowchart of the optimization algorithm is shown in the Figure 2. m Actors sample in turn, and the strategy is $\pi_t = \left[\pi_t^1, \pi_t^2, ..., \pi_t^m\right]$. The state of each actor is $S_t = \left[S_t^1, S_t^2, ..., S_t^m\right]$ and the action is $A_t = \left[\pi_t^1(S_t^1), \pi_t^2(S_t^2), ..., \pi_t^m(S_t^m)\right]$.
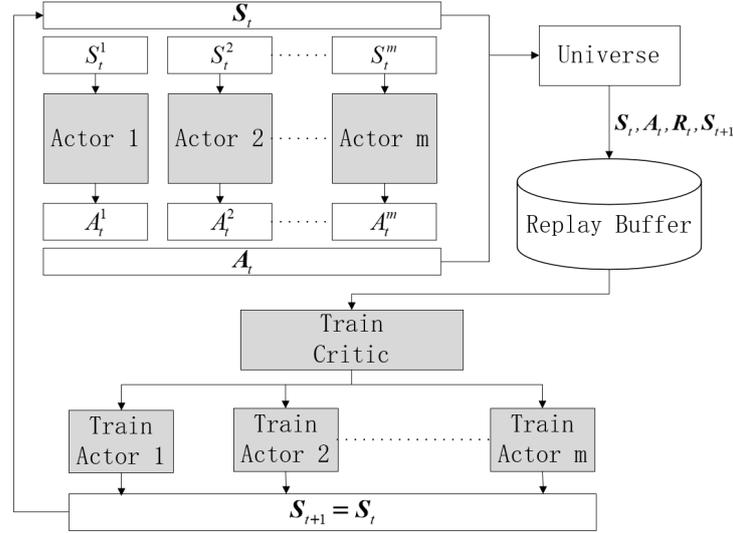
Figure 2 – Flowchart of the optimization algorithm in a step.

N groups of samples are extracted from the Replay Buffer by Mini-batch Gradient Descent method for training. For eval critic, the loss function L is shown in equation (7), in which π_t' is the target actor and Q_t' is the target critic network respectively.

$$L = \frac{1}{N}\sum_{i=1}^{N}\left(R_i + \gamma Q_t'\left(S_{i+1}, \pi_t'\left(S_{i+1} \mid \boldsymbol{\omega}'\right) \mid \boldsymbol{\theta}'\right) - Q_t\left(S_i, A_i \mid \boldsymbol{\theta}\right)\right)^2 \tag{7}$$

Each actor network is updated by Gradient Descent method:

$$\nabla_{\omega}J \approx \frac{1}{N}\sum_{i}\nabla_{A}Q\left(S, A \mid \boldsymbol{\theta}\right)\big|_{S=S_i, A=\pi(S_i)} \cdot \nabla_{\boldsymbol{\theta}}\pi\left(S \mid \boldsymbol{\omega}\right)\big|_{S=S_i} \tag{8}$$

When all the actors have finished training, go to the next step $t+1$. After completing all steps of sampling, return to their respective starting points and restart sampling.

2.3 Algorithm Test

In order to compare the advantages and disadvantages of the three reward functions, Fonseca function [23] is used for testing. Fonseca function contains two independent variables (x, y) and two dependent variables (f_1, f_2), as shown in equation (9).

$$\begin{cases} f_1\left(x, y\right) = 1 - e^{\left(-(x-1)^2 - (y+1)^2\right)} \\ f_2\left(x, y\right) = 1 - e^{\left(-(x+1)^2 - (y-1)^2\right)} \end{cases}, x \in \left(-2, 2\right), y \in \left(-2, 2\right) \tag{9}$$

Taking the minimum of two dependent variables as the optimization objectives, as shown in equation (10), the optimal solution set is a nonconvex solution set and there is no expression in analytical form.

$$\begin{cases} min\, f_1\left(x, y\right) \\ min\, f_2\left(x, y\right) \end{cases} \tag{10}$$

Figure 3 is the test result of the three reward functions. For MS-reward, it searches more in the direction of minimizing f_1, and there is very little search for the target f_2. The search of ED-reward method is more uniform, but a large number of samples are gathered in the middle of the Pareto front, and the search for both sides is less. HV-reward method can search the whole Pareto front more completely, and the distribution of non-dominated solutions is more uniform. In comparison, HV reward is more suitable as the reward function, so we use this method for the aerodynamic optimization of the high-lift device.

a) MS-reward   b) ED-reward   c) HV-reward
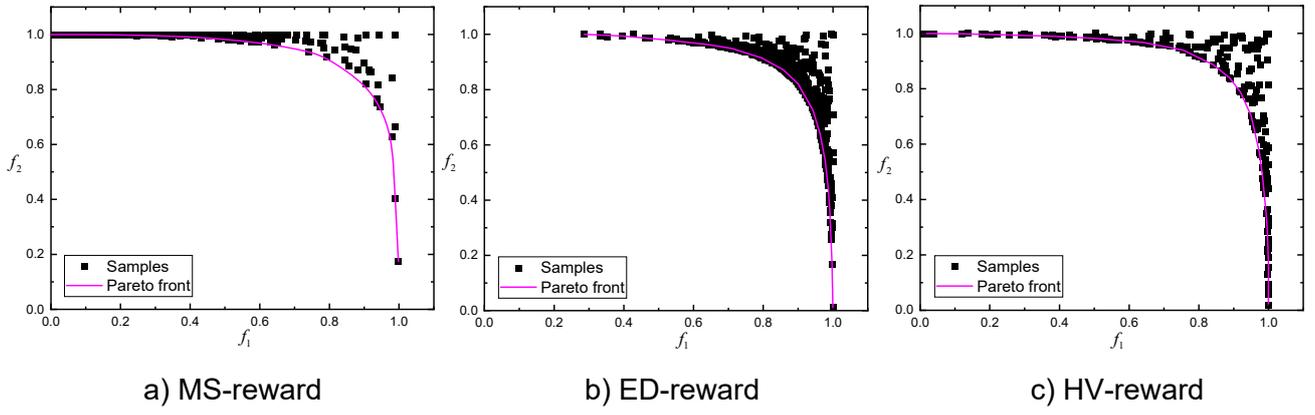
Figure 3 – Comparison of the different reward functions.

3. Optimization Problem

The trailing-edge high-lift device of a wide body aircraft is optimized, which is equipped with two flaps and eight spoilers that can deflect downward, as shown in the Figure 4. The spoiler with active downward deflection can change the Gap width between the spoiler and the flap, which is a technology that can significantly improve the aerodynamic performance. The origin of the Cartesian coordinate system is located at the nose of the aircraft. X direction is the horizontal inflow direction, Y is the height direction, and Z is the spanwise direction of the wing.
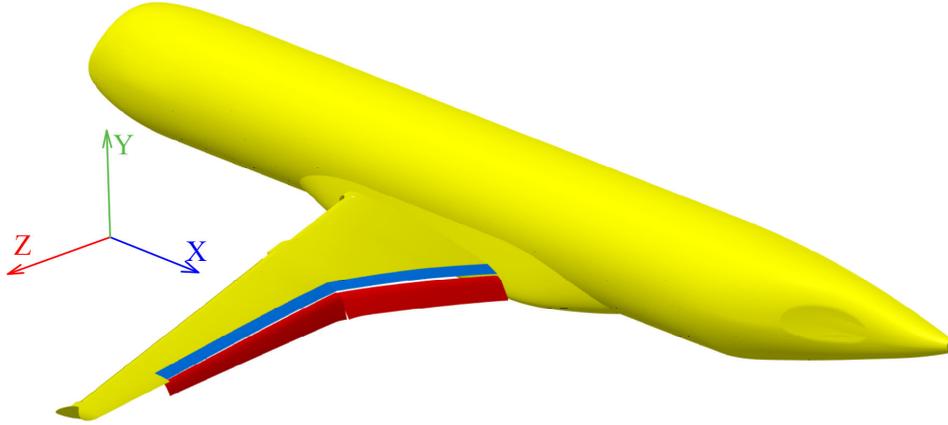
Figure 4 – The optimization model.

3.1 Optimization Variable

The deployment process of flap can be regarded as the spatial motion of rigid body. According to the kinematic principle, the flap from the stowed position to the deployed position can be converted into the flap rotation around an axis and the translation along the axis. The process can be described by the finite helix matrix $\left[R_H \right]$ in the form of 4×4 as shown in equation (11),

$$\left[R_H \right] = \begin{bmatrix} \left[R_\theta \right]_u & P + S\boldsymbol{u} - \left[R_\theta \right]_u P \\ 0 \quad 0 \quad 0 & 1 \end{bmatrix}$$

$$\left[R_\theta \right]_u = \begin{bmatrix} u_x^2 V_\theta + \cos\theta & u_x u_y V_\theta - u_z \sin\theta & u_x u_z V_\theta + u_y \sin\theta \\ u_x u_y V_\theta + u_z \sin\theta & u_y^2 V_\theta + \cos\theta & u_y u_z V_\theta - u_x \sin\theta \\ u_x u_z V_\theta - u_y \sin\theta & u_y u_z V_\theta + u_x \sin\theta & u_z^2 V_\theta + \cos\theta \end{bmatrix}$$

(11)

$$V_\theta = 1 - \cos\theta$$

where $P = (X, Y, Z)$ is a reference point of the axis, $\boldsymbol{u} = (u_x, u_y, u_z)$ is the direction vector of the axis.

In our work two reference points P^1 and P^2 are used to describe the axis which contains four independent variables X^1, Y^1, X^2 and Y^2. The deployed position coordinates (X_d, Y_d, Z_d) of a point on the flap can be obtained from the stowed position coordinates (X_s, Y_s, Z_s) and $\left[R_H \right]$, as shown in

equation (12).

$$\begin{bmatrix} X_d & Y_d & Z_d & 1 \end{bmatrix}^T = \begin{bmatrix} R_H \end{bmatrix}\begin{bmatrix} X_s & Y_s & Z_s & 1 \end{bmatrix}^T \tag{12}$$

In our work, there are 18 optimization variables, which are as follows.

For the inboard flap, there is only rotational motion during its deployment due to its small trailing-edge sweep angle. In addition, the axes of rotation used to describe the take-off position and landing position are the same. Its variables of the inboard flap contain an axis (reference point 1 $\boldsymbol{P}_{in}^1\left(X_{in}^1, Y_{in}^1, Z_{in}^1\right)$ and reference point 2 $\boldsymbol{P}_{in}^2\left(X_{in}^2, Y_{in}^2, Z_{in}^2\right)$), takeoff deflection angle $\theta_{in}^{takeoff}$ and landing deflection angle θ_{in}^{land}, a total of 6 variables.

For the outboard flap, the deployed position has a spanwise displacement, and the rotation axes can be different. Its variables contain take-off axis (reference point 1 $\boldsymbol{P}_{out}^{takeoff,1}\left(X_{out}^{takeoff,1}, Y_{out}^{takeoff,1}, Z_{out}^{takeoff,1}\right)$, reference point 2 $\boldsymbol{P}_{out}^{takeoff,2}\left(X_{out}^{takeoff,2}, Y_{out}^{takeoff,2}, Z_{out}^{takeoff,2}\right)$), take-off deflection angle $\theta_{out}^{takeoff}$, landing axis (reference point 1 $\boldsymbol{P}_{out}^{land,1}\left(X_{out}^{land,1}, Y_{out}^{land,1}, Z_{out}^{land,1}\right)$, reference point 2 $\boldsymbol{P}_{out}^{land,2}\left(X_{out}^{land,2}, Y_{out}^{land,2}, Z_{out}^{land,2}\right)$), landing deflection angle θ_{out}^{land}, a total of 10 variables. The spanwise displacement of the outboard flap is used to adjust the clearance between the two flaps, so it is not used as an optimization variable.

For the inboard spoilers, they contain takeoff deflection angle $\phi_{in}^{takeoff}$ and landing deflection angle ϕ_{in}^{land} as optimization variables. The deflection angles of the outboard spoilers change with the inboard spoilers to ensure that the kink position can be aligned.

3.2 Optimization Constraint and Target

Part of the constraints in the optimization comes from the misalignment degree of the two flaps at kink position, which is described by four parameters, as shown in the Figure 5. Where ΔX is the X coordinates difference between the inboard and outboard flaps, ΔY is the difference between the Y coordinates, $\Delta\beta_{XZ}$ is the included angle of the chord of the inboard and outboard flaps under the projection of the XZ plane, $\Delta\beta_{XY}$ is the included angle under the XY plane.
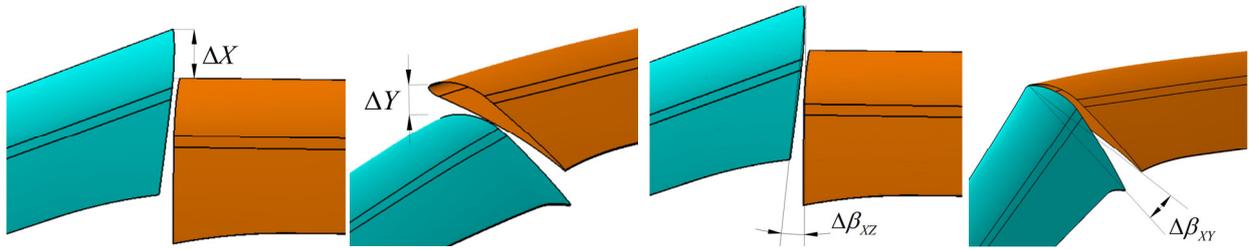
Figure 5 – Illustration of the optimization constraint.

The other part of the constraints is from Gap, which prevents the interference between the spoilers and the flaps and invalid aerodynamic numerical simulation. The specific constraints are shown in equation (13) where C represents the reference chord length.

$$Constraints: \begin{cases} \Delta X < 4.5\%C \\ \Delta Y < 4.5\%C \\ \Delta\beta_{XZ} < 7° \\ \Delta\beta_{XY} < 2° \\ 0\%C < Gap_{takeoff} < 2\%C \\ 0\%C < Gap_{land} < 4\%C \end{cases} \tag{13}$$

According to the common takeoff and landing conditions of commercial aircrafts, the Design Point (DP) is set as equation (14).

$$DP: \quad M_\infty = 0.2; \quad Re_\infty = 3\times10^7; \quad \alpha = 8° \tag{14}$$

The takeoff configuration and landing configuration are optimized. According to the Federal Aviation Regulations (FAR) Part 25 [2,12] and related research, the takeoff configuration requires a large lift-

to-drag ratio $K_{takeoff}$ to climb quickly, and a large lift coefficient $Cl_{takeoff}$ to reduce the taxiing distance; The landing configuration generally needs a large lift coefficient Cl_{land} to reduce the landing speed. The optimization objective can be expressed as:

$$Target : \begin{cases} K_{takeoff}\big|_{\alpha=8°} \\ Cl_{takeoff}\big|_{\alpha=8°} \\ Cl_{land}\big|_{\alpha=8°} \end{cases} \quad (15)$$

3.3 Two-Dimensional Approximate Simulation

The method of 2D CFD is adopted to evaluate the aerodynamic performance approximately. In general, the approximated aerodynamic performance has the same trend as the accurate performance. Three sections are selected, because for a flap, two sections can basically reflect the change of flap position. It is impossible to distinguish whether the flap moves or rotates in 3D space if only a single section is selected. The inboard and outboard flaps sampled by the optimization algorithm should be aligned at kink. Therefore, the three sections of the two flaps can basically describe the 3D position change of the flaps. The Section 1 and Section 3 are located in the middle of the inboard flap and the outboard flap respectively, while the Section 2 is selected near the kink position. The leading edge of the Section 1 is equipped with droop nose, while Section 2 and 3 are both equipped with slat.
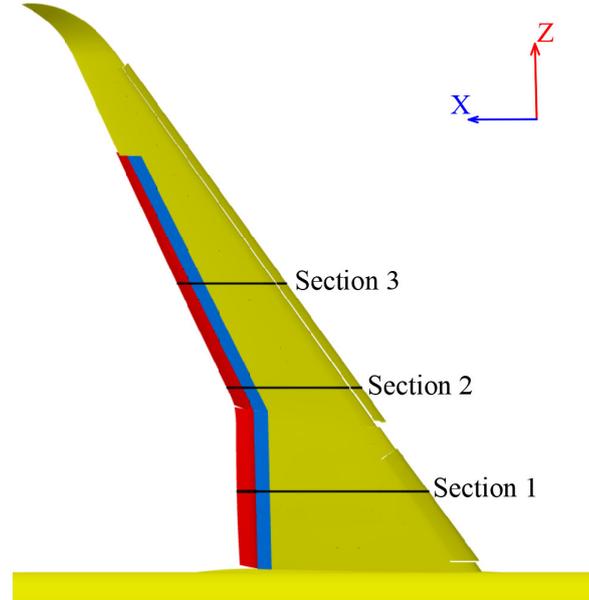
Figure 6 – Illustration of the three sections selected.

The 2D evaluation of each section is based on the theory of an infinite swept wing [11,24]. The incoming flow through the leading edge will be divided into two parts, one is perpendicular to the leading edge while the other is parallel, and the former is the main source of aerodynamic force. The aerodynamic coefficient of each section ($Cl_{\text{Section i}}$ and $Cd_{\text{Section i}}$) can be obtained from the equation (16) and equation (17),

$$Cl_{\text{Section i}} = Cl_{\text{Section i}}^{n} \times cos^2 \chi \quad (16)$$

$$Cd_{\text{Section i}} = Cd_{\text{Section i}}^{n} \times cos^3 \chi \quad (17)$$

where χ is the leading-edge sweep angle, $Cl_{\text{Section i}}^{n}$ and $Cd_{\text{Section i}}^{n}$ are the lift coefficient and the drag coefficient of the section projected onto the plane perpendicular to the leading edge respectively. The lift coefficient Cl and drag coefficient Cd of the wing are obtained from the equation (18). $C_{\text{Section i}}$ is the chord length of the section to balance the area effects of the inboard and outboard wing.

$$Cl = \frac{\sum_{i=1}^{3} Cl_{\text{Section i}} \times C_{\text{Section i}}}{\sum_{i=1}^{3} C_{\text{Section i}}}, \ Cd = \frac{\sum_{i=1}^{3} Cd_{\text{Section i}} \times C_{\text{Section i}}}{\sum_{i=1}^{3} C_{\text{Section i}}}, \ K = Cl \ / \ Cd \qquad (18)$$

4. Mesh and CFD Validation

The numerical method is verified based on the 30P30N model [25]. The incoming Mach number $M_\infty = 0.2$ and Reynolds number $Re_\infty = 9 \times 10^6$. The cube dimensions are considered based on the Mean Aerodynamic Chord (MAC) of the model. More specifically, the downstream side of this cube is 80 times greater than the MAC and the others are 40 times larger than the MAC. The structural grid is used to discretize the space, $Y^+ \approx 1$, and the growth rate is 1.1. The total number of nodes is about 0.13 million.

The commercial software ANSYS CFX was used to solve Reynolds Averaged Navier–Stokes (RANS) equations for aerodynamic performance calculation. The symmetrical surface of the calculation domain is set as the symmetry condition, the wall is set as no-slip condition, and the far field is set as the opening boundary. $SST \ k - \omega$ two equation model is used to describe turbulence. The residual target is 10^{-6}, or the number of iterations reaches 400.

The comparison between the calculated lift coefficient and the experimental one is shown in the Figure 7, and the comparison between the pressure coefficient distribution of the airfoil at $\alpha = 8°$ and the experimental one is shown in the figure. The results show that the calculated results are in good agreement with the experimental values in the whole range of angle of attack.
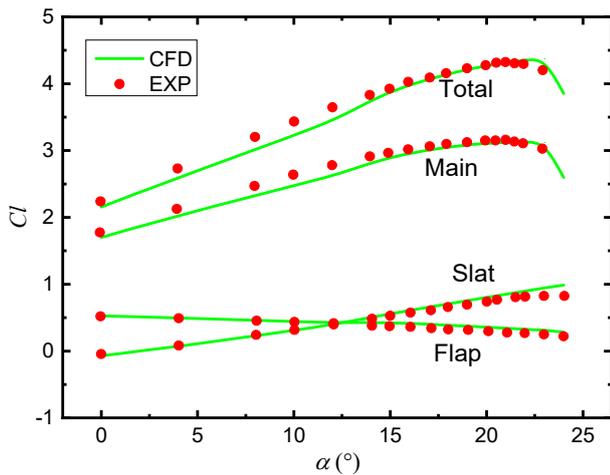
Figure 7 – Variation of Cl with α.
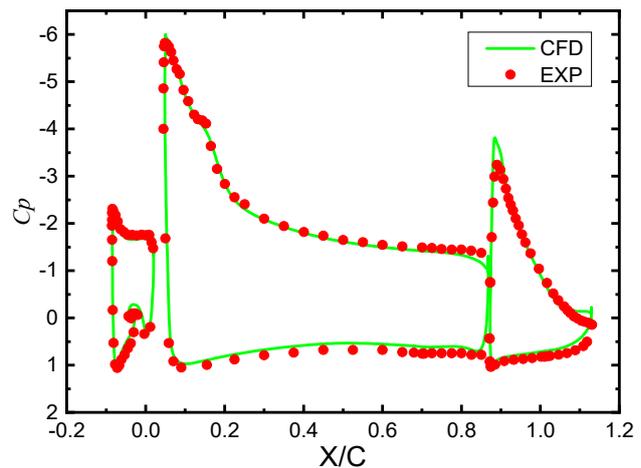
Figure 8 – Distributed of Cp.

ANSYS ICEM script [26] is used to deform the mesh in the optimization process. This process is conducted by keeping the mesh topology unchanged and moving vertices. The topology of the blocks and mesh are shown in Figure 9.
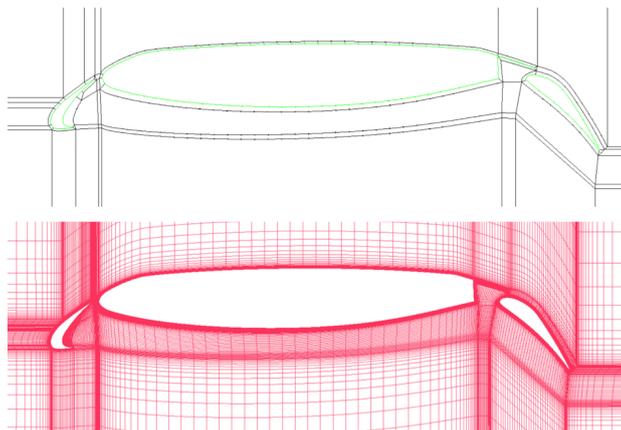
Figure 9 – Grid topology and mesh overview.

5. Results and Discussion

In this section, the convergence process, samples distribution and optimized configuration are studied in detail.

15 Actors are set to search in parallel, and 8 episodes are iterated. Each episode contains 6 steps, with a total of 720 iterations. The Hypervolume change in the convergence process is shown in the Figure 10. The reference point is set on (0, 0, 0), and the three objects are normalized during calculation. With the optimization, the growth rate of the Hypervolume slows down gradually.
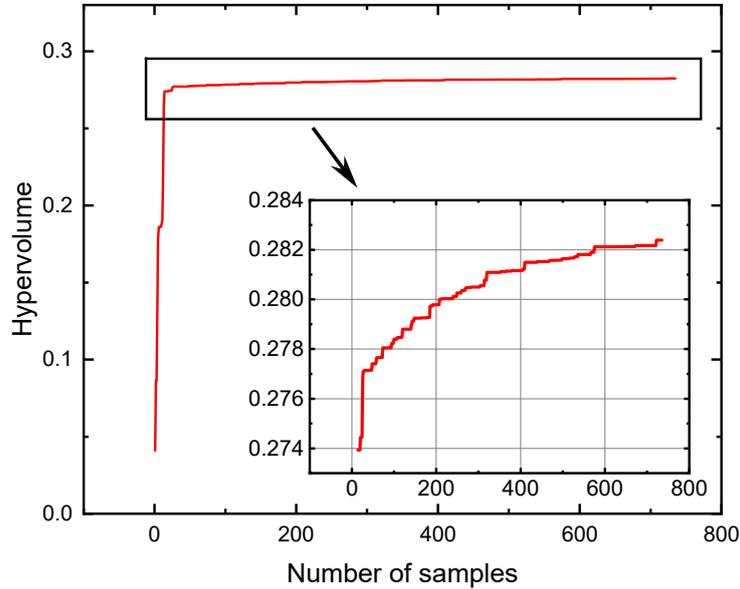
Figure 10 – Optimization variables

The $K_{takeoff} - Cl_{takeoff}$, $Cl_{land} - K_{takeoff}$ and $Cl_{land} - Cl_{takeoff}$ chart of the samples are shown in the Figure 11, 12, 13 respectively. Among them, the negative correlation between $K_{takeoff}$ and $Cl_{takeoff}$ is obvious. That is, the increase of the lift-to-drag ratio is generally accompanied by the decrease of the lift coefficient. This relationship also makes it easy for the optimization algorithm to find high quality Pareto solutions of takeoff configurations. Figure 11 shows that the Pareto solution is basically distributed in the leading edge of the all samples. In contrast, it is difficult to find the non-dominated solutions with better performance considering both the performances of take-off and landing. More Pareto solutions are not located at the leading edge. Three optimized configurations A, B and C are selected for detailed analysis.
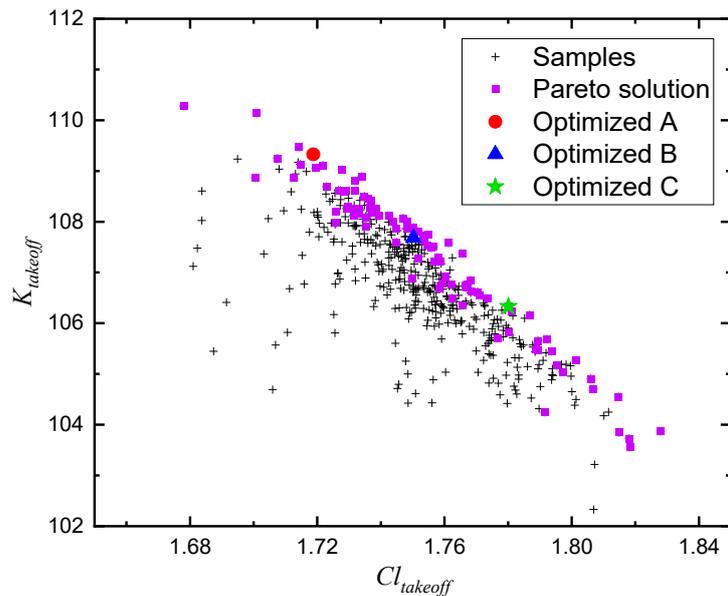
Figure 11 – $K_{takeoff} - Cl_{takeoff}$ distribution of the samples.

9
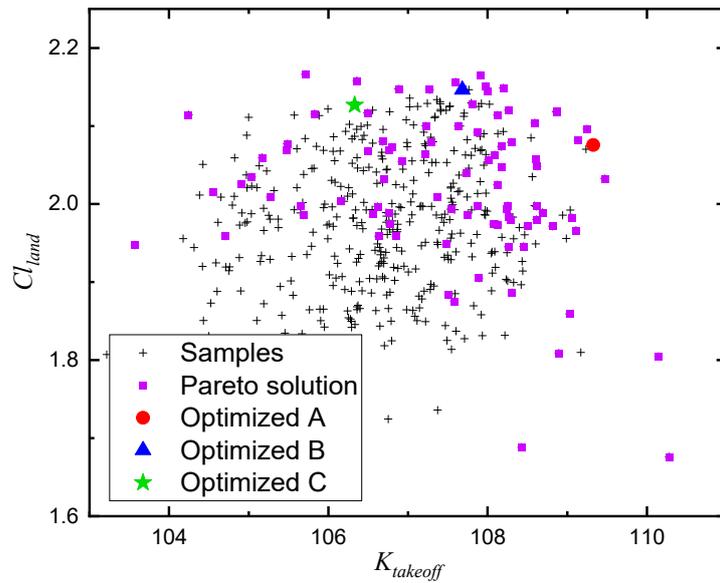
Figure 12 – $Cl_{land} - K_{takeoff}$ distribution of the samples.
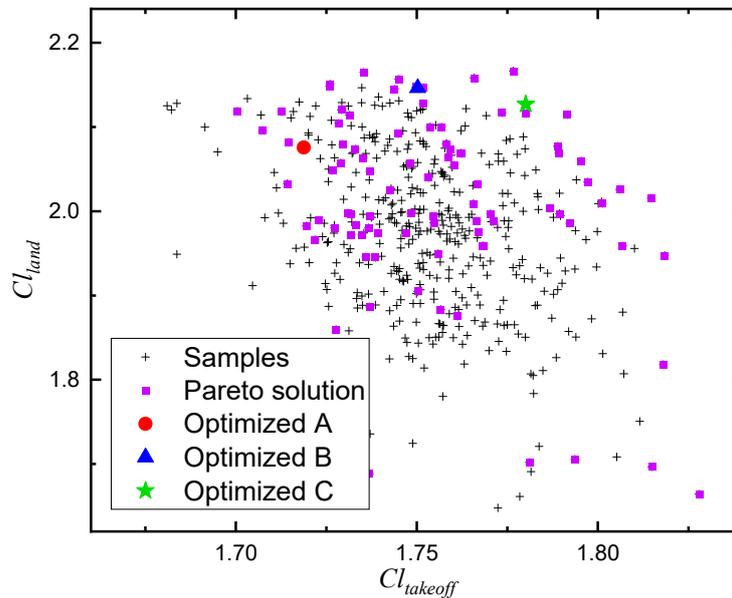
Figure 13 – $Cl_{land} - Cl_{takeoff}$ distribution of the samples.

The lift coefficient curves and lift-to-drag ratio curves of Sections 1, 2 and 3 of the three takeoff configurations (Optimized A, B and C) are compared in Figure 14 and 15 respectively. In the full range, the lift coefficient of configuration C is significantly greater than that of the other two configurations. Although the lift-to-drag ratio of C is relatively small at an angle of attack of 8 ° (as can be seen from Figure 11), if the three configurations are compared with a constant large lift coefficient, the lift drag ratio of C is better. The change of flap position basically does not change the stall angle of attack. The lift coefficient curves of sections 1, 2 and 3 of the three landing configurations are compared in Figure 16. Although the lift coefficient of configuration B is large at medium and small angles of attack, its stall angle of attack is small.

a) Section 1 b) Section 2 c) Section 3
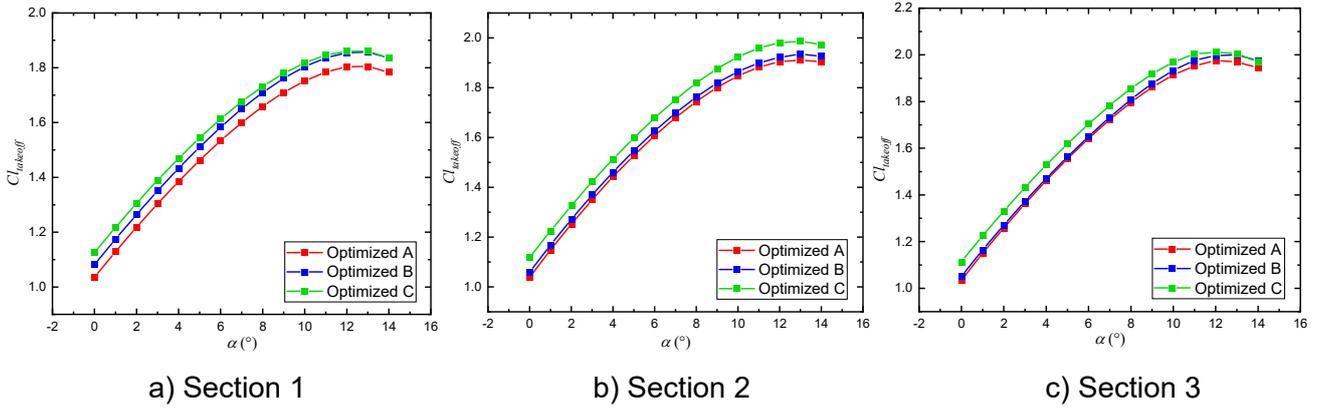
Figure 14 – $Cl_{land} - Cl_{takeoff}$ distribution of the samples.

a) Section 1 b) Section 2 c) Section 3
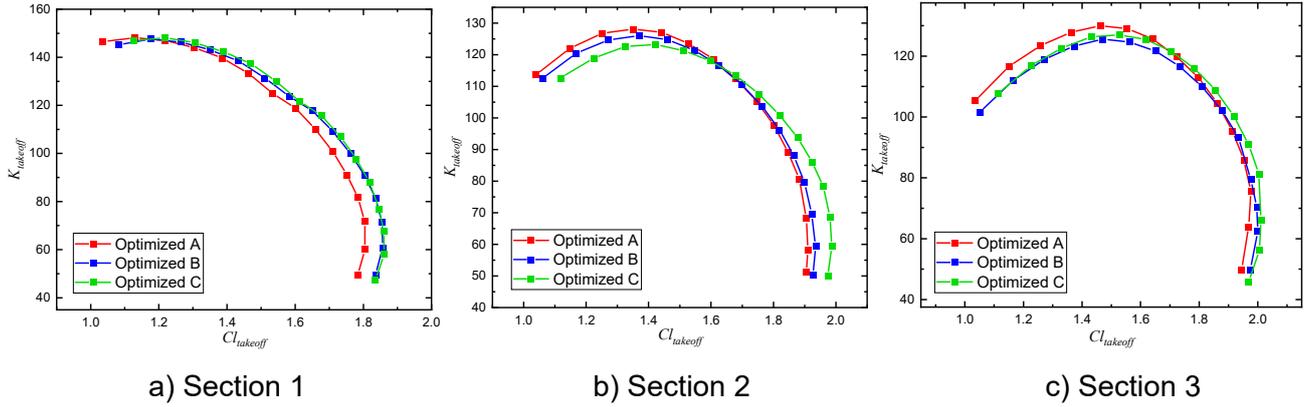
Figure 15 – $Cl_{land} - Cl_{takeoff}$ distribution of the samples.

a) Section 1 b) Section 2 c) Section 3
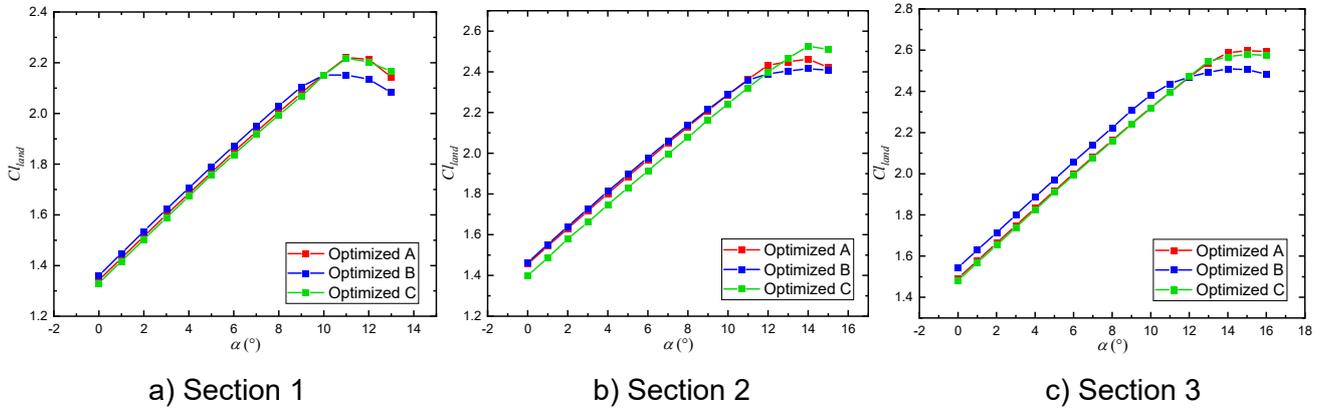
Figure 16 – $Cl_{land} - Cl_{takeoff}$ distribution of the samples.

6. Conclusion and Prospect

A multi-objective optimization algorithm based on DRL is proposed. The algorithm adopts multi-actor parallel training to search the complete Pareto solution in the variable space. Three reward functions, ($MS-reward$, $ED-reward$, $HV-reward$), are tested by Fonseca function. The results show that more complete and uniform Pareto solutions can be found by using $HV-reward$ function. The multi-objective algorithm is applied to the 2D optimization of the high-lift device of a commercial aircraft.

The optimization targets are to maximize the lift coefficient and lift-to-drag ratio of the takeoff configuration and the lift coefficient of the landing configuration at 8° angle of attack. The results show that the negative correlation between the lift coefficient and the lift-to-drag ratio of the take-off configuration is obvious. It is difficult to find the non-dominated solutions with better performance considering both the performances of take-off and landing. In the following work, the 3D optimization of the aircraft model will be carried out based on the same optimization algorithm, and the

11

experience of 2D optimization will be used for the pre-training of the agents and compared with the 3D optimization results.

7. Contact Author Email Address

Peiqing Liu: lpq@buaa.edu.cn

8. Copyright Statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS proceedings or as individual off-prints from the proceedings.

References

[1] Reckzeh D. Aerodynamic design of the high-lift-wing for a megaliner aircraft. *Aerospace Science and Technology*, Vol. 7, No. 2, pp 107-119, 2003.

[2] Part 25: Airworthiness Standard: Transport Category Airplanes.

[3] Smith A M O. High-lift aerodynamics. *Journal of Aircraft*, Vol. 12, No. 6, pp 501-530, 1975.

[4] Tian Y, Quan J C, et al. Mechanism/structure/aerodynamic multidisciplinary optimization of flexible high-lift devices for transport aircraft. *Aerospace Science and Technology*, Vol. 93, 2019.

[5] Trapani G, Kipouros T, Savill M. Computational aerodynamic design for 2D high-lift airfoil configurations. *V Pegasus-AIAA Student Conference*, 2010.

[6] Kanazaki M, Tanaka K, et al. Multi-objective aerodynamic optimization of elements' setting for high-lift airfoil using kriging model. *44th AIAA Aerospace Science Meeting and Exhibit*, Reno, Nevada, 2006.

[7] Weissinger J. The lift distribution of swept-back wings. *Technical Report Archive & Image Library*, 1947.

[8] Minervino M, Iannelli P, Quagliarella D. 3d flap design using navier-stokes equations and evolutionary optimization techniques on an industrial platform. *Evolutionary and Deterministic Methods for Design, Optimization and Control*, Capua, Italy, 2011.

[9] Nakayama H, Kim H, et al. Aerodynamic Optimization of Multi-Element Airfoil. *44th AIAA Aerospace Science Meeting and Exhibit*, Reno, Nevada, 2006.

[10] Yan X H, Zhu J H, et al. Aerodynamic shape optimization using a novel optimizer based on machine learning techniques. *Aerospace Science and Technology*, Vol. 86, pp 826-835, 2019.

[11] Dai J H, Liu P Q, et al. Aerodynamic optimization of high-lift devices using a 2D-to-3D optimization method based on deep reinforcement learning and transfer learning. *Aerospace Science and Technology*, Vol. 121, 2022.

[12] Wang W H, Liu P Q, et al. Numerical study of the aerodynamic characteristics of high-lift droop nose with the deflection of Fowler flap and spoiler. *Aerospace Science and Technology*, Vol. 48, pp 75-85, 2016.

[13] White D J. Multi-objective infinite-horizon discounted Markov decision processes. *Journal of Mathematical Analysis and Applications*, Vol. 89, No. 2, 1982.

[14] Gabor Z, Kalmar Z, Szepesvari C. Multi-criteria reinforcement learning. *Proceedings of the 15th International Conference on Machine Learning*, Madison, 1998.

[15] Mannor S, Shimkin N. A geometric approach to multi-criterion reinforcement learning. *Journal of Machine Learning Research*, Vol. 5, pp 325-360, 2004.

[16] Vamplew P, Yearwood J, et al. On the limitations of scalarisation for multi-objective reinforcement learning of Pareto fronts. *Proceedings of the 21st Australasian Joint Conference on Artificial Intelligence: Advances in Artificial Intelligence*, Berlin, pp 372-378, 2008.

[17] Holland J, Adaptation in natural and artificial systems: an introductory analysis with application to biology. *Control & Artificial Intelligence*, 1975.

[18] Eberhart R, Kennedy J, A new optimizer using particle swarm theory. *MHS95 Sixth International Symposium on Micro Machine & Human Science, IEEE*, 2002.

[19] Moffaert K V, Nowe A, Multi-Objective reinforcement learning using sets of pareto dominating policies. *Journal of Machine Learning Research*, Vol. 15, 2014.

[20] Sutton R S, Barto A G. *Reinforcement learning: an introduction*. MIT Press, 1998.

[21] Lillicrap T P, Hunt J J, et al. Continuous control with deep reinforcement learning. *Computer Science*, 2015.

[22] Fonseca C M, Paquete L, Lopez-Ibanez M. An improved dimension-sweep algorithm for the hypervolume indicator. *2006 IEEE Congress on Evolutionary Computation*, Vancouver, Canada, pp 1157-1163, 2006.

[23] Fonseca C M, Fleming P J. An overview of evolutionary algorithms in multi-objective optimization. *Evolutionary Computation*, Vol. 3, No. 1, pp 1-16, 1995.

[24] Wild J, Brezillon J, et al. Realistic high-lift design of transport aircraft by applying numerical optimization. *European Conference on Computational Fluid Dynamics*, 2006.

[25] Chin V D, Peters D W, et al. Flowfield measurements about a multi-element airfoil at high Reynolds numbers, 1993.

[26] ANSYS Inc. ANSYS ICEM CFD programmer's guide release 18.2, 2017.