JOURNAL OF AIRCRAFT
Vol. 47, No. 5, September–October 2010

Using Multiobjective Evolutionary Algorithms and Data-Mining Methods to Optimize Ornithopters' Kinematics

M. Hamdaoui,* J. Chaskalovic,† S. Doncieux,‡ and P. Sagaut§
Université Pierre et Marie Curie, 75252 Paris Cedex 5, France

DOI: 10.2514/1.45046

The aim of this work is to present a method to find and analyze maximum propulsive efficiency kinematics for a birdlike flapping-wing unmanned aerial vehicle using multiobjective evolutionary optimization and data-mining tools. For the sake of clarity and simplicity, simple geometry (rectangular wings with the same profile along the span) and simple kinematics (symmetrical harmonic dihedral motion) are used. In addition, it is assumed that the birdlike aerial vehicle (for which the span and surface area are, respectively, 1 m and 0.15 m^2) is in horizontal motion at low cruise speed (6 m/s). The aerodynamic performances of the flapping-wing vehicle are evaluated with a semi-empirical flight physics model and the problem is solved using an efficient multiobjective evolutionary algorithm called ϵ-MOEA. Groups of attractive solutions are defined on the Pareto surface, and the most efficient solutions within these groups are characterized. Given the high dimensionality of the Pareto surface in the kinematic parameters space, data-mining techniques are used to conduct the study. First, it is shown that these groups can be qualified versus the whole Pareto surface by accurate mathematical relations on the kinematic parameters. Second, the inner structure of each group is studied and highly accurate mathematical relations are found on the optimized parameters describing the most efficient solutions.

Nomenclature

A_n	=	amplitude associated with the nth harmonic, °
AR	=	wing aspect ratio
B_n	=	even Fourier coefficient of ψ for the nth harmonic, °
$b/2$	=	semispan length, m
C_{d_f}	=	section friction-drag coefficient
C_m	=	aerodynamic moment coefficient
$C_{m_{ac}}$	=	section pitching moment coefficient about its aerodynamical center
C_n	=	odd Fourier coefficient of ψ for the nth harmonic, °
C_z	=	lift coefficient
C_{zc}	=	target lift coefficient
c	=	airfoil chord, m
DC_z	=	lift lag, defined in Eq. (10)
e	=	section distance between the leading edge and the elastic axis, m
f_ψ	=	frequency of the dihedral motion, Hz
g	=	acceleration of gravity, m/s^2
L	=	lift force, N
M	=	mass, kg
M_{aero}	=	norm of the aerodynamic moment
M_p	=	pitching moment
N	=	normal force to wing's chord, N
P	=	power, W
S	=	wing area, m^2
T	=	thrust, N
t	=	time, s
V_c	=	cruise velocity, m/s
α_0	=	angle of section's zero-lift line, °
η	=	propulsive efficiency
η_s	=	leading-edge suction efficiency
$\bar{\theta}$	=	pitch angle of chord with respect to U
$\bar{\theta}_a$	=	pitch angle of flapping axis with respect to U
$\bar{\theta}_w$	=	mean pitch angle of chord with respect to flapping axis
ρ	=	atmospheric density, kg/m^3
ϕ_n	=	phase associated with the nth harmonic, rad
ψ	=	dihedral angle, °
$\bar{}$	=	time average
$\dot{}$	=	time derivative

Subscripts

ac	=	aerodynamic center
c	=	circulation
in	=	input
max	=	maximum
min	=	minimum
sep	=	separated flow

I. Introduction

THROUGHOUT time, the flight of birds has stimulated the scientific community to conceive vehicles that fly like birds [1], which led to the apparition of flapping-wing unmanned aerial vehicles (UAVs). These kinds of UAVs have many advantages over existing ones. First, they can fly with suppleness at low speeds without the loss in performance experienced by fixed-wing UAVs. Second, their acoustic signature is much more diffuse than rotorcraft UAVs, which makes them difficult to spot. Finally, they can hover and take off vertically. Many projects focused on building birdlike flapping-wing UAVs: the ornithopter project of DeLaurier [2–4], the PRF REMANTA (Research Programme on Microvehicle and New Technologies Application) [5], or the ROBUR project (the name comes from a novel by Jules Verne) [6]. Within each of these projects, the scientists tried to optimize the design of flapping-wing aerial vehicles to increase their efficiency (lift, lift to drag ratio, energy consumption, etc.). Currently, an important part of the research and development community works on conceiving the most energy-efficient airfoil adaptation and wing motion technologies able to provide the required aerodynamic loads for UAV flights. For instance, Khan and Agrawal [7] designed and optimized a flapping-wing mechanism for micro air vehicles. Hall et al. [8,9] presented a

Received 20 April 2009; revision received 2 June 2010; accepted for publication 8 June 2010.

*Ph.D. Student, Institut Jean Le Rond d'Alembert, 4 Place Jussieu; hamdaoui@lmm.jussieu.fr.

†Associate Invited Professor, Institut Jean Le Rond d'Alembert, Paris, France; Ariel University Center, 40700 Ariel, Israel; joel.chaskalovic@upmc.fr.

‡Lecturer, Institut des Systèmes Intelligents et de Robotique; Stéphane.Doncieux@isir.fr.

§Professor, Institut Jean Le Rond d'Alembert, 4 Place Jussieu; sagaut@lmm.jussieu.fr.

Betz-criterion method to compute the minimal power in forward flight and gave some useful trends of variation for the thrust, the lift, and the propulsive efficiency as functions of the flapping amplitude and the frequency. Similarly, Willis et al. [10] presented a quasi-Newton method to minimize the flight power of ornithopters in cruise flight. Berman and Wang [11] showed that insects hovering kinematics minimize energy consumption. De Margerie et al. [12] applied artificial evolution to optimize the morphology and the kinematics of a flapping-wing birdlike UAV. All of these works except de Margerie et al.'s used a single-objective approach to solve the problem in spite of its multiobjective nature. In this work, a multiobjective method to find and analyze optimal kinematics for an ornithopter in cruise flight is presented. For the sake of clarity, a simplified framework have intentionally been chosen. The geometry of the wings (rectangular rigid wings) is simplified, the flapping motion (harmonic dihedral motion) is symmetrical and the aerodynamic performances are evaluated by a semi-empirical flight physics model [13]. The simplicity of such framework enables a clear and understandable illustration of the efficiency of the method. The problem at hand involves multiple objectives and several optimization parameters. Unlike most of the approaches listed above (except [12]), all of the pertinent objectives of the problem have been simultaneously taken into account in this work. The commonly used single-objective method outputs only one solution, while the multiobjective procedure produces several solutions, as long as the objectives are contradictory (which is the case in this work). Thus, the single-objective approach does not provide the possibility to characterize solutions optimality by design rules (because it is impossible to infer rules from one solution), whereas the multiobjective strategy allows us to infer optimality rules from the cloud of Pareto-optimal solutions using statistical methods [14]. Thus, the multiobjective point of view is in two ways advantageous: it allows simultaneously taking into account all of the relevant objectives of the problem and it provides the possibility to find optimality rules for the Pareto solutions increasing by this fact the comprehension of the problem. The group of optimal solutions is found using a multiobjective evolutionary algorithm as this class of methods is, to the authors' knowledge, the best tool to solve multiobjective optimization problems [15] when the objectives can be evaluated in a reasonably short time, which is the case in our study, due to the use of a fast and simplified flight physics model. As stated previously, the multiobjective optimization procedure outputs a global Pareto set with multiple solutions. Groups of attractive solutions are defined on the Pareto surface and the most efficient solutions among them are characterized. Given the high dimensionality of the global Pareto set in the kinematic parameters space, it is not easy to explore its structure (see Bellman's "curse of dimensionality" [16]). Different techniques of high-dimensional Pareto surface analysis are commonly used (principal components analysis [17], multidimensional scaling [17], response surfaces [18], etc.), but to the authors' knowledge, none of them can simultaneously represent high-dimensional data sets accurately and allow us to visualize them comfortably. It is the goal of this work to propose a method that handles easily high dimensionality and that can characterize with accuracy selected subsets of the global Pareto set with mathematical relations or design rules. Indeed, mathematical relationships on the optimized parameters (actually low-risk inequalities also called rules) are extracted to describe the most efficient solutions. Giving rules is more profitable than just handing the value of the kinematic parameters and the objectives for the most efficient solutions. Indeed, these rules provide some useful information to the designers to understand the optimality of the most efficient solutions. In Sec. II the propulsive efficiency maximization problem is presented, then in Sec. III models and methods used to solve the problem are exposed and in Sec. IV the multiobjective optimization results are presented along with the rules qualifying the most efficient solutions.

II. Problem Statement

A 1-m-span birdlike model with rectangular rigid wings equipped with a Liebeck LPT 110 A airfoil [19] is in horizontal cruise flight at a velocity $V_c = 6$ m/s. A simple geometry has intentionally been chosen to present a method to find and analyze optimal solutions within a simplified framework. The symmetrical flapping motion can be decomposed into three basic motions: dihedral motion, sweep motion, and twist motion, also called pitch (see Fig. 1). In our case, only the dihedral motion has been considered. It is, of course, a simplification of the real motion of birds' wings, but it would have complicated the analysis to add the twist motion and/or the sweep motion, which is why, for the sake of clarity, the kinematic has been restrained to a simple dihedral motion. The dihedral angle ψ is given by a Fourier series [20] as follows:

$$\psi = C_0 + \sum_{n=1}^{2} [C_n \sin(2\pi n f_\psi t) + B_n \cos(2\pi n f_\psi t)] \qquad (1)$$

An amplitude and phase can be defined for each harmonic n as follows:

$$A_n = \sqrt{C_n^2 + B_n^2}, \qquad \phi_n = \arctan\left(\frac{C_n}{B_n}\right) \qquad (2)$$

In the literature [2,8–10,12,13,20], most of the studies on flapping wings use one harmonic or two to describe the kinematics of the flapping. In our study, one harmonic for the dihedral motion ($C_2 = B_2 = 0$) has been used, as in [2,12]. The reason for this choice is that the model of [13], which is based on DeLaurier's model [2], requires sinusoidal functions with one harmonic as the input kinematics. The flapping-wing kinematics are generated by an engine, which generally works on electricity or fuel and transforms this energy into mechanical energy. The onboard energy storage (batteries, fuel, etc.) generates an additional mass to carry and consequently reduces the payload, as the UAVs' total mass is limited for performance reasons. Souied [21] reported that half of an UAV's mass is dedicated to the carriage of the necessary energy resources for flight, which significantly reduces the payload. Thus, the optimization of the use of onboard energy appears to be an important factor to improve UAVs performances in terms of payload, energy consumption, flight time, etc. There are different ways to boost the energetic efficiency of an UAV: improvement of the efficiency of the propulsion units, reduction of the drag, energetic efficiency optimization of the wing kinematics, etc. The energetic efficiency optimization of the wing kinematics retained our attention in this work. Recently, many studies [22–24] have investigated the high propulsive efficiencies encountered in nature (insects, birds, and fish) in order to discover the principles governing these high performances. Others tried to optimize existing designs to maximize the propulsive efficiency of UAVs in flight [8–10,12,24,25]. But none of these studies has sought to characterize the optimal solutions with design rules. It is the goal of this work to present a method to find and characterize with design rules the most efficient kinematics that maximize the propulsive efficiency in horizontal cruise flight for a given velocity.

III. Models and Methods

A. Flight Physics Model

The model of [13], presented in the Appendix, is an adaptation for large incidence angles of DeLaurier's model [2]. It has been successfully used to predict the lift and thrust of a rectangular wing (1 m span) in pitch and plunge at a Reynolds number of 7×10^4 and a

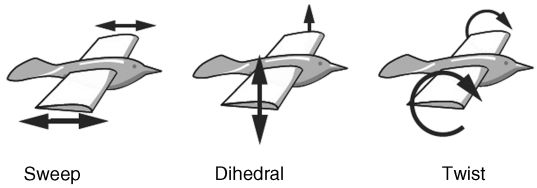

Fig. 1 Flapping motion can be decomposed into three basic motions: sweep, dihedral, and twist.

reduced frequency of 0.4 [13]. Figure 4 of [13] compares the model prediction and the experimental values of lift and thrust. One can easily see on this figure that the model of [13] is in excellent agreement with the experience. Furthermore, DeLaurier's model [2] has also been successfully compared to experimental wind-tunnel measurements at a Reynolds number of 2×10^5, as it is shown in Fig. 11 on page 160 of [26]. This figure shows that the model predictions are close to the loads measurements (lift and thrust coefficients) made on a real ornithopter wing. Given the geometry and forward velocity of our birdlike vehicle, it evolves at a Reynolds number of about 6×10^4, and according to maximum and minimum frequencies on the Pareto surface (see Table 1), it is characterized by a reduced frequency that varies between 0.34 and 0.78. It appears then that the model of [13] is suitable for our application, as it has been validated against experimental data for a Reynolds number and a reduced frequency close to those of our study. Nevertheless, as we stressed out in Sec. II, the model requires the use of sinusoidal functions with one harmonic as input kinematics. This limitation remains acceptable for our purpose, as we are not in need of more than one harmonic to expose our method, the goal of which is to find and analyze cruise optimal kinematics for a birdlike vehicle.

B. Optimization: Concepts and Tools

1. Terminology

When dealing with an optimization problem, specific vocabulary is used. On its most general form a multiobjective optimization problem reads as follows:

$$\begin{aligned} &\text{maximize } f_m(x), \qquad m = 1, \ldots, p, \quad \text{with } g_j \geq 0 \\ &j = 1, \ldots, K, \qquad h_j = 0, \qquad j = 1, \ldots, L \\ &x_i^l \leq x_i \leq x_i^u, \qquad i = 1, \ldots, q \end{aligned} \tag{3}$$

where p is the number of objectives, K the number of inequality constraints (represented by the functions g_j) and L the number of equality constraints (represented by the functions h_j). The vector $x = (x_1, \ldots, x_q)$ is the q-dimensional decision vector. The q numbers x_i^l and x_i^u are, respectively, the lower and upper bounds for the variable x_i. These bounds define the *search space* or *decision space*, **D**. The image of the decision space by the objectives is called *objectives space*. A solution is an element of the decision space and its image lies in the objectives space. A solution that satisfies the constraints is feasible and the group of feasible solutions is called *feasible space* **S**.

2. Dominance and Pareto-Optimality

A solution x^i of the problem (3) is said to dominate another solution x^j, if the following conditions are satisfied:

1) The solution x^i is not worse than x^j with respect to all objectives, which means that $f_m(x^i) \geq f_m(x^j)\ \forall\ m \in \{1, \ldots, p\}$.

2) The solution x^i is strictly better than x^j with respect to at least one objective, which means that $\exists\ m \in \{1, \ldots, p\}$ such that $f_m(x^i) > f_m(x^j)$.

The *global Pareto set* of the multiobjective optimization problem (3) is composed of the feasible solutions that are not dominated by any other solution of the feasible space. The image of the Pareto set in the objectives space is called *Pareto surface* or *Pareto front* for a biobjective problem (Fig. 2).

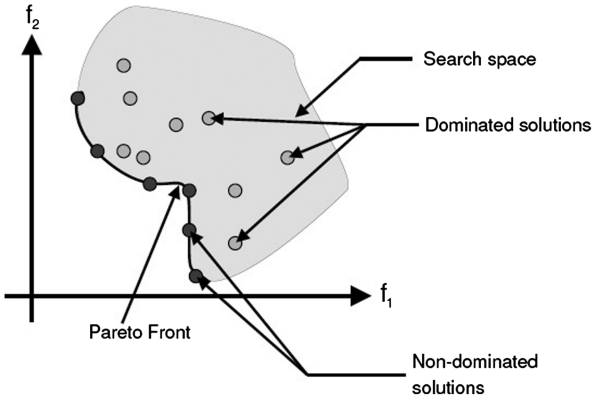

Fig. 2 Pareto front: dominated and nondominated solutions.

3. Search Procedures of the Nondominated Set

Most optimization problems involve numerous objectives in practice. A simple approach consists of transforming them into a single objective by using a weighted sum of the relevant objectives as follows:

$$F_w(x) = \sum_{m=1}^{p} w_m f_m(x) \tag{4}$$

where the weights $w_m \geq 0$ verify

$$\sum_{m=1}^{p} w_m = 1$$

Then to solve problem (3), the following optimization program is considered:

$$\max_{x \in \mathbf{S}} F_w(x) \tag{5}$$

The solution of problem (5) is included in the Pareto surface. It is given by the tangency points between the Pareto surface and the hyperplane for which the normal is the vector $w = (w_1, \ldots, w_p)$. Then by changing the weights we can get the whole front if it is convex [27]. If the Pareto surface is a hyperplane in the objectives space, then the two approaches are equivalent and one will prefer the single-criterion approach for its simplicity. But if the Pareto surface is concave or changes curvature, then this single-criterion method will be unable to get all the points on the Pareto surface by just changing the weights [27]. On the contrary, multiobjective optimization procedures produce a set of tradeoffs, among which a higher-level algorithm, or the user, may select the favorite one without the need to assign a priori weights to these alternatives. One of the most interesting properties of evolutionary optimization methods is their ability to deal with multiple objectives at once. Numerous algorithms have been proposed [15] to generate the set of such compromises. Most of them rely on the concept of Pareto dominance and generate the so-called Pareto surface. In this work, the ϵ-MOEA algorithm [28] has been used for its efficiency and robustness to generate globally Pareto-optimal sets. It is based on the ϵ-dominance concept, where ϵ is the discretization resolution of the Pareto surface.

Table 1 Statistical description of the Pareto surface and global Pareto set

	$\bar{\eta}$	DC_z	C_m	A_1	ϕ_1	f_ψ	C_0
Min	0.018	−0.38	0.02	24.75	−1.1361	4.43	−38.28
Max	0.17	3.12	0.49	56.57	1.4083	10	18.5
Mean	0.11	0.67	0.16	49.51	−0.18	7.54	−13.57
Std dev	0.04	0.72	0.11	4.99	0.8	1.36	14.41
First quartile	0.096	0.12	0.07	46.865	−0.8	6.55	−24.81
Second quartile	0.12	0.49	0.14	50.72	−0.7	7.66	−14.65
Third quartile	0.14	1.06	0.24	52.73	0.8	8.59	−3.77

4. Optimization Program

Some relevant variables for our problem need to be stated. In horizontal symmetrical flapping flight with prescribed forward velocity, the following equalities hold:

$$DC_z = 0 \tag{6}$$

$$C_T = C_d \tag{7}$$

$$C_m = 0 \tag{8}$$

where DC_z is defined in Eq. (10) and C_m is the moment coefficient. The optimal solutions should satisfy these equations. Following empirical relations established for birds [29], an average total mass M (0.69 kg) and an average wing surface area S (0.15 m^2) are computed. Then the necessary lift coefficient C_{zc} to support the weight is computed as follows:

$$C_{zc} = \frac{2Mg}{\rho S V_c^2} \tag{9}$$

where V_c is the cruise velocity, $g = 9.81\ \mathrm{m \cdot s^{-2}}$ and $\rho = 1.295\ \mathrm{kg \cdot m^{-3}}$. For $V_c = 6$ m/s, we have $C_{zc} = 1.97$. Using the empirical relations provided by [29], maximum and minimum values for the mass ($M_{\max} = 5.0$ kg and $M_{\min} = 0.1$ kg) and the cruise velocity ($V_{c_{\max}} = 30$ m/s and $V_{c_{\min}} = 6$ m/s) for a birdlike vehicle of 1 m span are estimated. Then using Eq. (9), maximum ($C_{z_{\max}} = 14.3$) and minimum ($C_{z_{\min}} = 0.01$) values for the mean lift coefficient are computed. The value of $C_{z_{\max}}$ is unexpectedly high in comparison with the commonly encountered values of aerodynamic lift coefficient. But we will show that the lift coefficient obtained in our simulations takes realistic values and that the upper and lower bounds ($C_{z_{\min}}$ and $C_{z_{\max}}$) we used for the lift coefficient are never reached on the obtained Pareto surface. With an estimation for birds of the mean body drag C_{d_b} [29] (here $C_{d_b} = 0.37$) maximum ($C_{d_{\max}} = 0.06$) and minimum ($C_{d_{\min}} = 0.0043$) values for the drag coefficient are computed via the relation $C_d = 0.00813 M^{\frac{2}{3}} (C_{d_b}/S)$. Equation (7), allows to derive upper and lower bounds for the mean thrust coefficient C_T as well [see Eq. (11)]. A reference thrust coefficient could have been computed rather than deriving upper and lower bounds for the thrust coefficient, but as the thrust coefficient is related to the drag coefficient of the body and the appendices through Eq. (7), this would have made our results dependent on the drag model and the shape of the body and the appendices. That is why we chose to impose upper and lower bounds on the thrust coefficient to avoid the use of a drag model to predict the drag on the body and to guarantee that our results are independent of the body shape. The use of Eq. (7) to derive these upper and lower bounds for the thrust coefficient means that Eq. (7) is supposed to hold and that at each predicted thrust value corresponds a drag value related to a family of body shapes. Furthermore, DC_z and DC_z^* are defined by

$$DC_z = C_z - C_{zc}, \qquad DC_z^* = \frac{DC_z}{C_{zc}} \tag{10}$$

The constraints can be summarized as follows:

$$C_{d_{\min}} \le C_T \le C_{d_{\max}} \tag{11}$$

$$C_{z_{\min}} \le C_z \le C_{z_{\max}} \tag{12}$$

$$\frac{M_{\min} - M}{M} \le DC_z^* \le \frac{M_{\max} - M}{M} \tag{13}$$

DC_z^* is the difference between the load that can be taken by the aerial vehicle and M (0.69 kg) divided by M. As the maximal mass is of 5 kg and the minimal mass is of 0.1 kg, only the solutions with a DC_z^* inferior to $DC_{z\max}^* = (M_{\max} - M)/M$ (6.5) and superior to $DC_{z\min}^* = (M_{\min} - M)/M$ (−0.85) are feasible. As we seek to maximize the propulsive efficiency $\bar{\eta}$ while satisfying Eqs. (6–8), it is clear that we have to simultaneously maximize the propulsive efficiency $\bar{\eta}$, minimize DC_z, and minimize C_m, whereas Eq. (7) is supposed to be automatically satisfied. The two constraints (11) and (12) are enforced during the optimization loop, and Eq. (13) is enforced at the end of the optimization process. The optimization variables are the kinematic parameters of the dihedral motion given in Eq. (1). These quantities vary in the following ranges:

$$C_0 \in [-40°, 40°], \qquad C_1 \in [-40°, 40°], \qquad B_1 \in [-40°, 40°]$$
$$f_\psi \in [0\ \mathrm{Hz}, 10\ \mathrm{Hz}]$$

The objectives are scaled with an asymmetrical function, defined as follows:

$$\Gamma = \begin{cases} \frac{1}{x+1} & \text{if } x > 0, \\ \frac{-1}{x-1} - 1 & \text{otherwise} \end{cases}$$

This function is used to favor finding positive values of x, which means in our case finding propulsive efficiencies $\bar{\eta}$ as close to 1.0 as possible, lift coefficients (C_z) slightly superior to the targeted lift coefficient C_{zc} and moment coefficients C_m slightly superior to zero. A penalty Π has been added to each criterion to discard uninteresting solutions (negative propulsive efficiency or negative power) and reject those that do not satisfy Eqs. (11) and (12). The penalty function is based on the Heaviside function H defined by

$$H(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{otherwise} \end{cases}$$

The penalty function is defined as follows:

$$\Pi = -H(\bar{\eta} - 1) - H(-\bar{\eta}) - H(C_T - C_{d_{\max}}) - H(C_{d_{\min}} - C_T) - H(C_z - C_{z_{\max}}) - H(C_{z_{\min}} - C_z) - H(-\bar{P}_{\text{in}})$$

The constraint (13) on DC_z^* has not been included in the penalty function but it is taken into account at the end of the optimization process by deleting the solutions that do not respect it to obtain the final global Pareto set. This constraint on the mass has not been enforced during the optimization process to obtain a greater diversity of solutions. The scaled optimization program reads as follows:

$$\text{Maximize} \begin{cases} F_1 = \Gamma(\bar{\eta} - 1) + 2\Pi - 2, \\ F_2 = \Gamma(C_z - C_{zc}) + 2\Pi - 2, \\ F_3 = \Gamma(C_m) + 2\Pi - 2 \end{cases}$$

The same penalty is applied to all the objectives to guarantee that solutions with no penalty (called nonpenalized solutions) dominate the solutions with a nonzero penalty (called penalized solutions).

5. Optimization Code

An open-source software framework, SFERES[¶] (Software Framework Enabling Research on Evolution and Simulation) written in C++ and developed by Landau et al. [30] is used for the optimization step. A module providing the value of (F_1, F_2, F_3) using the [13] model is coupled with this software.

C. Data-Mining Methods: Principles and Terminology

The postprocessing of the obtained global Pareto set is performed with data-mining techniques using Clementine software.[**] Data mining is an activity for which the goal is to discover hidden or a priori unknown facts contained in databases. Using a combination of machine learning, statistical analysis, modeling techniques and database technology, data mining finds patterns and subtle

[¶]Data available online at http://sferes.lip6.fr/ [retrieved 18 June 2010].

[**]Data available online at http://www.spss.com/software/modeling/modeler-pro [retrieved 18 June 2010]. Clementine SPSS has changed to IBM SPSS Modeler Professional since IBM bought SPSS.

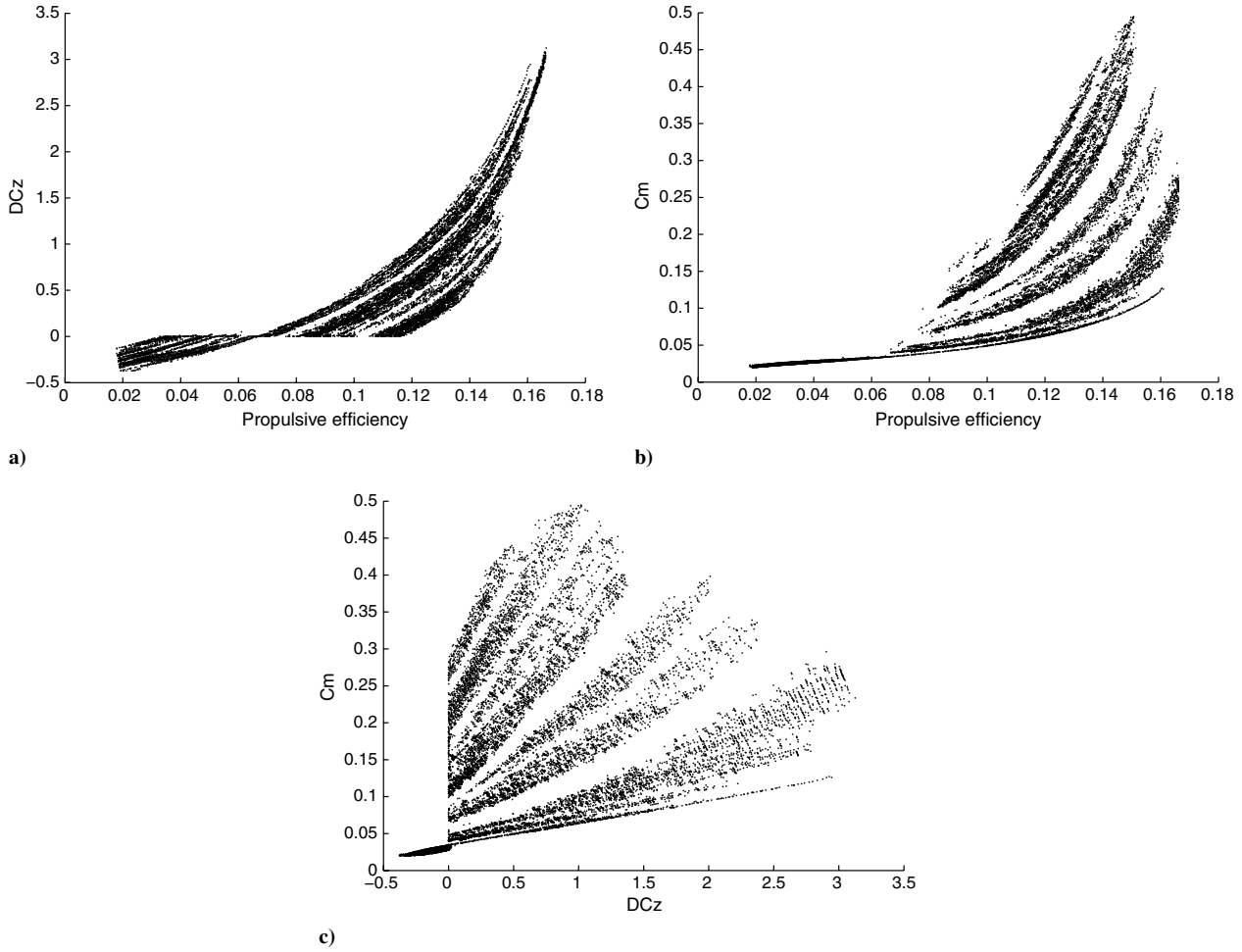

Fig. 3 Scatter-plot matrix representation of the objectives for the obtained Pareto surface: a) DC_z versus $\bar{\eta}$, b) C_m versus $\bar{\eta}$, and c) C_m versus DC_z.

relationships in data and infers rules that allow the prediction of future results.

1. Segmentation with Decision Trees

Decision trees [31] belong to the supervised data-mining tools that process segmentation (i.e., division into segments). The purpose of segmentation is to identify homogeneous subgroups inside a given population regarding a target variable that is to be explained versus predictor variables defined by the user. In the case of segmentation (that we have considered in our study), the target variable identifies a category; it describes belonging to a given group for which the characteristics are specified by the user: for example, the class of solutions for which the propulsive efficiency is superior (or inferior) to a given threshold. In this work, the segmentation have been performed by decision trees for which the operating principles are detailed throughout the presentation of binary decision trees (with only two classes) in the next paragraph (the extension to multiple classes is straightforward). A binary decision tree is composed of various subgroups (called *nodes*) from the initial population (called *root node*). These nodes are generated by identifying among the predictor variables the most discriminating one regarding the homogeneity of the resulting nodes. Each node is associated with a predictive value of the target variable; for segmentation it is usually the class. This predictive value is computed with a certain risk estimate that corresponds to the percentage of misclassified solutions in each node: i.e., the percentage of solutions that do not belong to the predicted class of the node. A tree can then be built by splitting the root node into nodes in a recursive manner. Each step splits a node into two nodes (binary decision trees), based on the most discriminating predictor variable var such that the left resulting node is characterized by the inequality var $\leq \tau$ and the right one by var $> \tau$ (τ being a threshold). This process stops when the splitting is either unfeasible or a singular segmentation can be applied to each element of the derived node. The terminal nodes are called *leaves*. The path from the root node to each leaf defines a succession of inequalities on the predictor variables that characterize the solutions belonging to the leaf with a certain risk depending on the percentage of misclassified solutions in the leaf. One single path is associated with a succession of inequalities called a *rule*. By choosing the leaves that predict the membership to the class of interest, one is able to characterize this class with a set of rules at a certain risk.

2. Kohonen Networks

Kohonen networks belong to the unsupervised data-mining tools and are based on neural networks [32,33]. No target variable is defined, but depending on the given set of predictor variables, the objective of the method is to build clusters inside a given population such that each group has to be homogeneous regarding unknown relations that have to be discovered. The objective of a Kohonen network is to form a discrete, topological mapping of an input space of arbitrary dimension Q, $\mathbf{x}_i$ being the ith input vector. Close input vectors should be close in the map (topology preservation). A Kohonen map is composed of output units often arranged in a 2-D rectangular or hexagonal grid, $\mathbf{z}_i$ being the 2-D coordinate vector of the ith output unit on the map. These units are also defined by Q-dimensional weight vectors, also called code books, with $\boldsymbol{\lambda}_i$ being the Q-dimensional weight vector of unit i. The output units are connected to each other by a neighborhood relation based on a neighborhood function $\upsilon(z_i, z_j)$ that takes as arguments the coordinates of the output units on the map. The learning process is as follows:

1) Initialize the weights for each output unit to small random numbers.

2) Loop until weights changes are negligible.

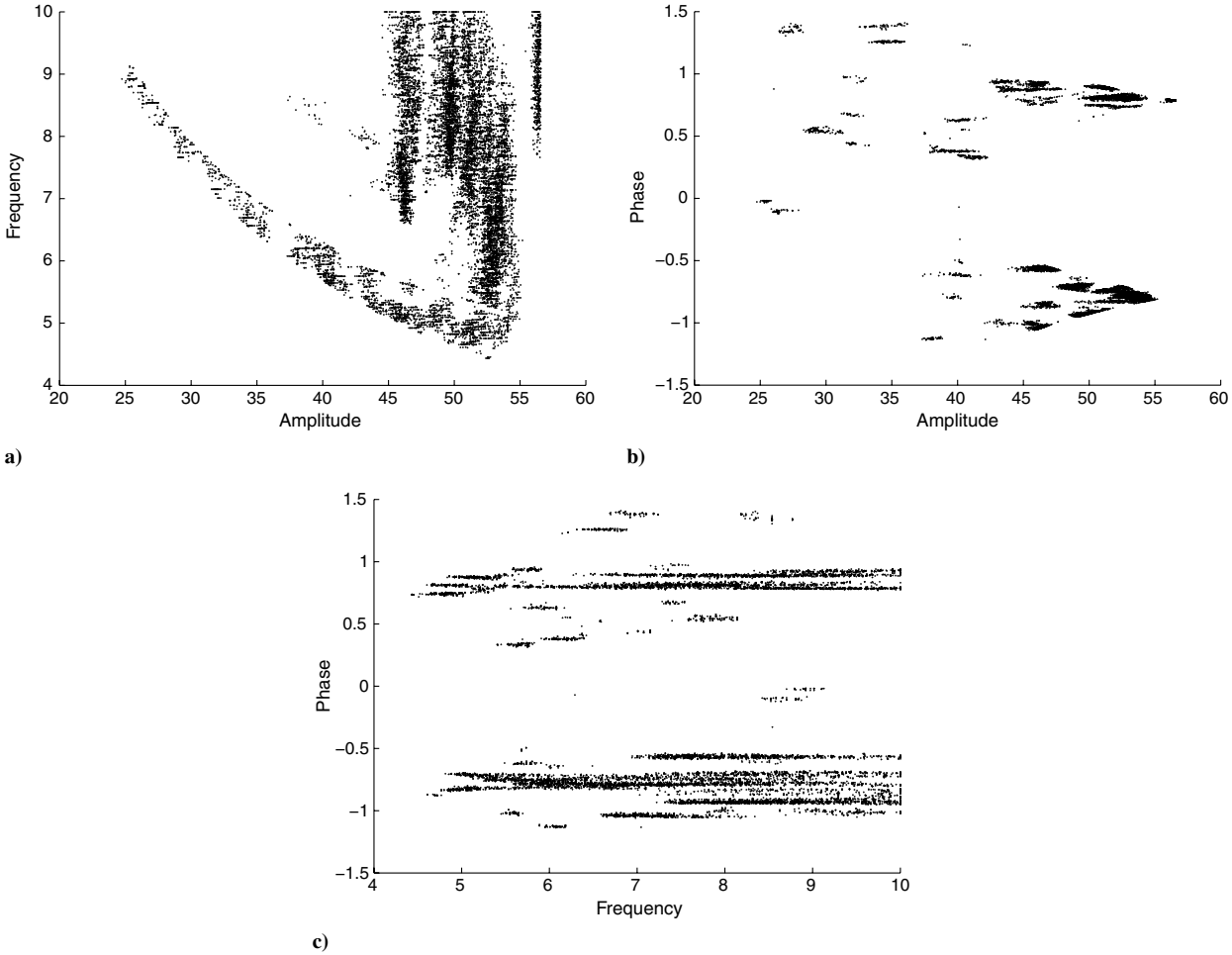

Fig. 4 Scatter-plot matrix representation of the parameters for the obtained global Pareto set: a) frequency versus amplitude, b) phase versus amplitude, and c) phase versus frequency.

3) For each input vector $\mathbf{x}_i$,
   a) Feed the input vector to the output units.
   b) Find the winning output unit i^*.
   c) Find all units in the neighborhood of the winner.
   d) Update the weight vectors for these units.
4) Reduce the size of the neighborhoods if required.

The winning output unit is simply the unit with the weight vector that has the smallest distance (usually Euclidian) to the input vector. The neighborhood of a unit is composed of the units within some distance to that unit on the map (the distance is computed with the neighborhood function, given the coordinates of the units on the map [32,33]). For example, if the size of the neighborhood is one, then all units no more than one unit away (either horizontally or vertically) fall within its neighborhood. The weights of every unit in the neighborhood of the winning unit (including the winning unit itself) are updated using

$$\boldsymbol{\lambda}_i = \boldsymbol{\lambda}_i + \mu\upsilon(z_i, z_{i^*})(\mathbf{x}_i - \boldsymbol{\lambda}_i)$$

where μ is a control parameter of the method, called the learning rate. This will move each unit in the neighborhood closer to the input vector. As time progresses the learning rate and the neighborhood size are reduced. If the parameters are well chosen the final network should capture the intrinsic characteristics of the input data. The built map is a low-dimensional representation of the Q-dimensional input space. It can be used for visualization or to extract some interesting properties of the input space. One can group the output units that are similar to distinguish groups or patterns in the data. This grouping procedure is called *clustering*.

3. Postprocessing Steps

Decision trees and Kohonen maps have been combined to analyze the global Pareto set obtained at the end of the optimization loop. Groups of attractive solutions are defined by mathematical relationships in order to select subsets from the Pareto surface. The goal is to discover the mathematical relationships on kinematic parameters (global Pareto set) to qualify the most efficient solutions of these subsets. Therefore, the following strategy is used:

1) Define groups of attractive solutions.
2) Characterize the subset of interest with respect to the entire global Pareto set using a binary decision tree: The target variable is the membership to the subset and the output is the decision-tree rules, called rule_t.
3) Study the inner structure of each subset with Kohonen maps: the output is a certain number of clusters.
4) Select the cluster with the highest proportion of efficient solutions.
5) Run the decision tree on the subset: the target variable is the membership to the selected cluster and the output is the decision-tree rules, called rule_k.
6) Describe the most efficient solutions of the subset by taking the intersection of rule_t and rule_k.

IV. Results and Discussion

A. General Considerations

Evolutionary algorithms are stochastic optimization algorithms. Their output may then vary, even when restarted with the same parameters. To tackle this problem, eight identical runs are launched and the eight Pareto surfaces are aggregated. Then the final Pareto surface is obtained after a Pareto nondomination sort on the

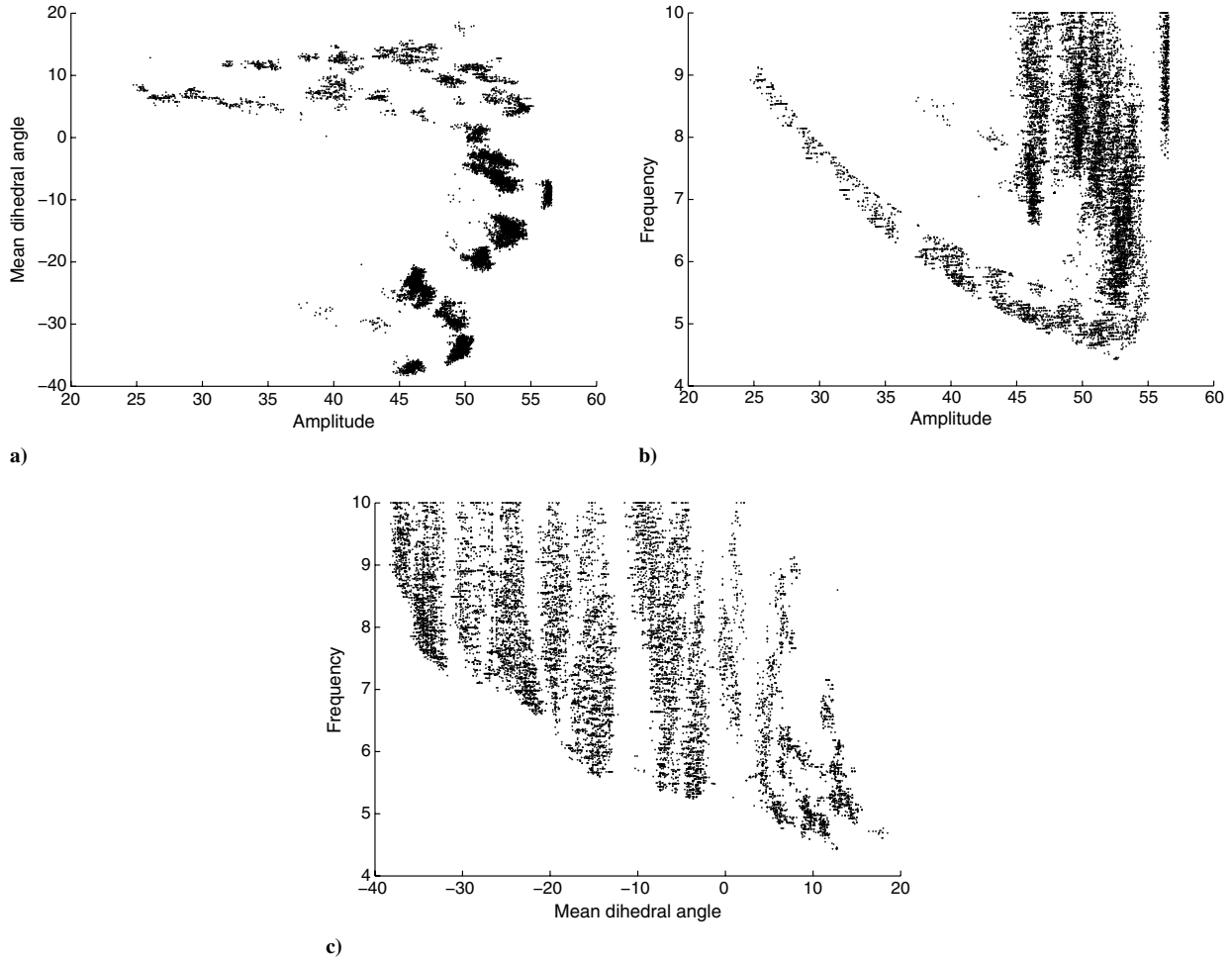

Fig. 5 Scatter-plot matrix representation of the parameters for the obtained global Pareto set: a) C_0 versus amplitude, b) frequency versus amplitude, and c) frequency versus C_0.

aggregated set (Sec. III.B.2). For each run, the optimization loop is stopped when no visual change on the output is detected within a period of at least 10 generations or when the maximum number of generations is reached (5000).

Space integrals have been computed with the Simpson formula [34] using 1024 points and classical averaging have been employed to compute time averages with 100 points (the number of points in space have been determined so that the change in value of the flight physics model's output was less than 0.0001 by doubling the number of points). A value of 0.025 is used for $C_{m_{ac}}$ and a common forward speed ($V_c = 6$ m/s) for birds in slow flight with a 1 m wing span [12,29] is considered.

B. Global Pareto Set and Pareto Surface

First of all, it appears that $DC_z^* = DC_z/C_{zc}$ ($C_{zc} = 1.97$) takes values that are between $DC^*_{z_{\max}} \sim 6.5$ and $DC^*_{z\min} \sim -0.85$. The maximal value of DC_z^* (1.63) is inferior to $DC^*_{z_{\max}}$ and the minimal value of DC_z^* (−0.2) is superior to $DC^*_{z_{\min}}$, thus all the Pareto-optimal solutions obey to Eq. (13). By rewriting Eq. (10) as follows,

$$C_z = C_{zc}(1 + DC_z^*)$$

upper (5.2) and lower (1.5) bounds for the lift coefficient on the Pareto surface are computed. It appears that these bounds are in the interval $[C_{z_{\min}}, C_{z_{\max}}]$ and are distinct from $C_{z_{\min}}$ and $C_{z_{\max}}$: thus, all Pareto-optimal solutions are characterized by realistic lift coefficients. Furthermore, some Pareto-optimal solutions have very low propulsive efficiencies, as can be seen in Table 1. There are two main reasons to explain that: the tendency of propulsive efficiencies to be low for slow cruise velocities and the underestimation of propulsive efficiencies for slow cruise velocities by the [13] simplified flight physics model. Indeed, the variation of the flight power versus the velocity is U-shaped [29], which means that the flight power takes high values for low as well as high velocities. As the propulsive power increases with the velocity, the propulsive efficiency is low for low and high speeds and peaks for an intermediate speed. In our case, the cruise velocity is low, so it is normal to get low propulsive efficiencies. Nevertheless, as the flight physics model of [13] is simplified, it does not take into account low Reynolds number phenomena that affect the performances at slow cruise speed such as laminar-turbulent transition, laminar separation bubbles, and wake capture [35–37]. This may explain why the predicted propulsive efficiencies are very low in our case, but as there are no reliable measurements of the propulsive efficiency for real birds [38], it is difficult to compare the propulsive efficiencies we obtained to real ones in the case of slow cruise flight. As our objective is to present a method to find and analyze maximum cruise-flight propulsive efficiency kinematics for a flapping-wing vehicle using a multiobjective approach and data-mining tools, the low values of the propulsive efficiency are not a problem, since they do not affect the methodology. In the following, the optimization parameters (C_0, C_1, B_1, f_ψ) will be replaced by the equivalent set of parameters $(C_0, A_1, \phi_1, f_\psi)$. Actually, the amplitude of the dihedral motion A_1, computed in degrees, and the phase ϕ_1, computed in radians [see Eq. (2)], are more meaningful than C_1 or B_1. For the sake of clarity and simplicity, the expressions *optimization*, *optimized*, and *optimal parameters* will refer to $(C_0, A_1, \phi_1, f_\psi)$, given that the two sets of parameters are strictly equivalent from the optimization point of view because of Eq. (2).

Some statistical indicators are computed on the Pareto surface and the global Pareto set to get an idea of the distribution of the values of objectives and optimal parameters. These values can be found in

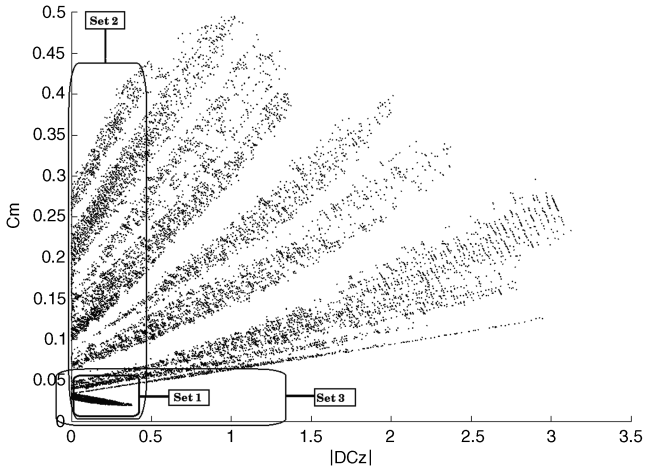

Fig. 6 $|DC_z|$ versus C_m for the set obtained at $V_c = 6$ m/s.

Table 1. For each objective $(\bar{\eta}, DC_z, C_m)$ and kinematic parameter $(C_0, A_1, \phi_1, f_\psi)$, four different groups of values have been defined. The values inferior to the first quartile are qualified by very low values, those between the first and the second quartiles are designated by low values, those between the second and the third quartiles are called high values and those superior to the third quartile are identified by very high values. Using the scatter-plot matrix method [39] (see Fig. 3) on the Pareto surface, it comes that high propulsive efficiency values are associated with high C_m values (see Fig. 3b) and high DC_z values (see Fig. 3a).

One can also find the same representation for the corresponding optimal parameters (see Figs. 4 and 5). Nevertheless, as these high propulsive efficiency solutions are associated with high C_m values and high DC_z values they do not satisfy the cruise-flight Eqs. (6–8). If the high DC_z values are acceptable, as they mean that the predicted lift is greater than the needed lift, the high C_m values are unacceptable, because they do not guarantee flight stability. This poses the question of the choice of the selection criteria to adopt for assessing Pareto-optimal solutions attractiveness. Within this work, we define groups of solutions according to the values taken by the objectives and study their specific characteristics in terms of rules on the kinematic parameters (global Pareto set) using data-mining methods such as decision trees and Kohonen maps.

C. Data-Mining Approach

Our goal is to find the Pareto-optimal solutions that maximize $\bar{\eta}$ and satisfy the constraints (6–8). In Fig. 6, the projection of the Pareto surface on the plane $(|DC_z|, C_m)$ allows us to distinguish four different sets as follows:

1) Set 1 is the set of solutions that simultaneously minimize $|DC_z|$ and C_m.

2) Set 2 is the set of solutions that minimize only C_m.

3) Set 3 is the set of solutions that minimize only $|DC_z|$.

4) Set 4 is the remaining solutions.

Set 1 corresponds to the solutions that fulfill the three constraints (6–8), set 2 refers to the solutions that meet only a couple of constraints (8) and (7), and set 3 consists of the solutions that satisfy only the couple of constraints (6) and (7).

Therefore, the sets 2 and 3 partially fulfill the constraints (6–8). Set 3 includes solutions that can have high C_m values, as it does not satisfy the constraint on the moment coefficient, which is unacceptable from the flight stability point of view. Set 2 includes low C_m solutions (ensures flight stability) that have positive DC_z (possibility to take additional payload), which makes it interesting to consider. Now the three sets can be ordered in order of importance: set 1 is the most interesting because it fulfills Eqs. (6–8); then comes set 2, because it contains stable solutions that can take additional payload; finally, set 3, which does not contain potentially interesting solutions. In the following, the characteristics of the sets 1 and 2 are studied. Each set can be defined more precisely with mathematical inequalities as follows:

$$(|DC_z| - \min(|DC_z|)) \leq C_{zc}/10 \tag{14}$$

$$(C_m - \min(C_m)) \leq C_{m_{ac}} \tag{15}$$

Set 1 is defined by the two inequalities (14) and (15), and set 2 is defined by the inequality (15). The idea behind the inequalities (14) and (15) is to select a neighborhood, if it does exist, of the point that conjointly minimizes $|DC_z|$ and C_m. One can easily understand that for set 2, the idea is to select a neighborhood of the point that minimizes only C_m. $C_{zc}/10$ has been chosen as amplitude variation in the $|DC_z|$ direction, because C_{zc} is the reference lift coefficient of our problem and $C_{zc}/10$ is a small lift coefficient variation compared to C_{zc}. In the C_m direction, $C_{m_{ac}}$ has been chosen as amplitude variation, because $C_{m_{ac}}$ is the reference moment coefficient of our problem and is small enough compared to the mean C_m value (see Table 1) on the Pareto surface. To get an idea of the distribution of the values of the objectives and the kinematic parameters for the sets 1 and 2, we computed some statistical indicators (see, respectively, Tables 2 and 3). It is worth noting that all the data-mining techniques used in this paper produce stable and reliable results in the sense that one is able to get reproducible results with Clementine software.

Table 2 Statistical description of set 1 in the objectives and parameters space

	$\bar{\eta}$	DC_z	C_m	A_1	ϕ_1	f_ψ	C_0
Min	0.018	−0.19	0.021	24.75	−1.14	4.7	−5.93
Max	0.08	0.19	0.045	54.97	1.41	9.13	13.17
Mean	0.04	−0.07	0.0295	42.53	0.16	6.08	6.75
Std dev	0.015	0.08	0.0052	8.89	0.81	1.116	3.96
First quartile	0.03	−0.14	0.026	34.83	−0.74	5.21	5.74
Second quartile	0.04	−0.07	0.028	43.156	0.38	5.76	6.83
Third quartile	0.05	−0.01	0.031	51.41	0.8	6.69	9.16

Table 3 Statistical description of set 2 in the objectives and parameters space

	$\bar{\eta}$	DC_z	C_m	A_1	ϕ_1	f_ψ	C_0
Min	0.018	−0.38	0.02	24.75	−1.14	4.43	−5.93
Max	0.09	0.34	0.045	55.05	1.41	9.13	18.5
Mean	0.04	−0.12	0.028	43.73	0.25	5.79	8.57
Std dev	0.016	0.12	0.0055	7.90	0.78	1.03	4.32
First quartile	0.028	−0.21	0.024	39.32	−0.71	5.07	6.31
Second quartile	0.036	−0.135	0.0266	45.285	0.55	5.44	9.02
Third quartile	0.048	−0.04	0.0295	50.92	0.87	6.10	11.96

1. Analysis of Set 1

The solutions that belongs to set 1 have very low C_m values and low to very low DC_z values. Rules on the kinematic parameters are found to characterize the solutions of set 1, called the *gold* group, versus the other solutions of Pareto surface, called the *else* group.

a. Characterization of Set 1 Versus the Global Pareto Set by Decision Trees. A decision-tree algorithm is ran to find rules on the optimization parameters $(C_0, A_1, \phi_1, f_\psi)$ (called the predictor variables), to explain the belonging to set 1 or gold group. Indeed, the purpose of the decision tree is to divide into segments the whole population of solutions into homogeneous subgroups where the standard deviation of the target variable is minimum. Here, the target variable is categorical. It describes the membership to set 1: solutions that simultaneously minimize $|DC_z|$ and C_m or not (for more detail on the decision tree, see, for example, [31]). Because the target variable is binomial, the minimization of the standard deviation consists of maximizing the percent of solutions that belong to one of the two classes (gold or else) in a given node. The processed tree is shown in detail in (Fig. 7). It is worth noting that given that set 1 contains only one-tenth of the whole Pareto set, the modeling techniques based on decision trees (as almost all of the data modeling techniques), would have trouble with such biased data, because it will tend to learn only the group composed by 90% of the solutions that do not belong to set 1, since it is more rare. If the data are well-balanced with approximately equal numbers of solutions in each group, the decision tree will have a better chance to find patterns that distinguish the two groups. It is the reason why we processed a rebalancing operation on the data of set 1 by overweighting the corresponding solutions for more accurate and significant results. The precision of the tree is assessed by a contingency table that gives the risk estimate of the tree (3.56%). Risks tell the chances of misclassification at any level in a decision tree. For categorical predictions as in our study, the risk is the proportion of cases incorrectly classified. So if the risk estimate is equal to 3.56, it means that 96.44% of the data are correctly classified by the segmentation model computed by the decision tree. Here, for a binomial modeling prediction, the risk describes the proportion of cases incorrectly classified by the tree. Therefore, the quality level of the decision tree is very high and its reading is as follows. The highest group in the decision tree (see Fig. 7), called the root node, corresponds to the whole population that is to be segmented. One can note that the two classes of solutions each represent 50% of the whole database, given that we proceeded to a rebalancing operation, as we explained before. Then a first split appears due to the identification by the decision-tree algorithm that the variable C_0, with a threshold of -4.11 is the most discriminating one. That is to say that any another predictor variable will produce a worse segmentation in terms of homogeneity. The result due to this first level of segmentation is that the two corresponding nodes that appear at the second level of the tree (see Fig. 7) are more homogeneous: 98.8% of the solutions of the node 1 are in the else group and 85.7% of the solutions of the node 2 are in the gold group. This result corresponds to a significant improvement of the degree of homogeneity regarding the root node constituted by 50% of gold solutions against 50% of else solutions. Moreover, if we explore the tree we find two very homogeneous nodes (nodes 4 and 7, see Fig. 7) regarding the gold group (around 95% of the solutions in each node correspond to set 1), for which the two exclusive rules are given as follows:

1) Rule 1:

$$-4.11 \leq C_0 \leq 2.32 \quad \text{and} \quad f_\psi \leq 5.67 \tag{16}$$

2) Rule 2:

$$2.32 \leq C_0 \leq 13.05 \tag{17}$$

One can note that these rules are given in terms of C_0 and f_ψ, with C_0 being the most discriminating variable.

b. Characterization of the Inner Structure of Set 1 by Kohonen Maps. Once set 1 has been characterized by comparison to the rest of the global Pareto set, we analyze its inner structure by using a typology method based on self-organizing Kohonen networks [32,33] to discover the role played, if any, by the predictor variables (kinematic parameters) in the composition of the set. The predictor variables that have been taken into account to build theses clusters are $(C_0, A_1, \phi_1, f_\psi)$. The use of this data-mining technology available in Clementine software yielded eight clusters, within which each member is distinguished by its belonging to one of the four propulsive efficiency's quartiles formed on set 1. This mode of qualification allows us to identify a posteriori homogeneous groups characterizing the highest level of the propulsive efficiency $\bar{\eta}$. We select the cluster, called cluster 1, that contains a high proportion of solutions for which the values of $\bar{\eta}$ are superior to the second and the

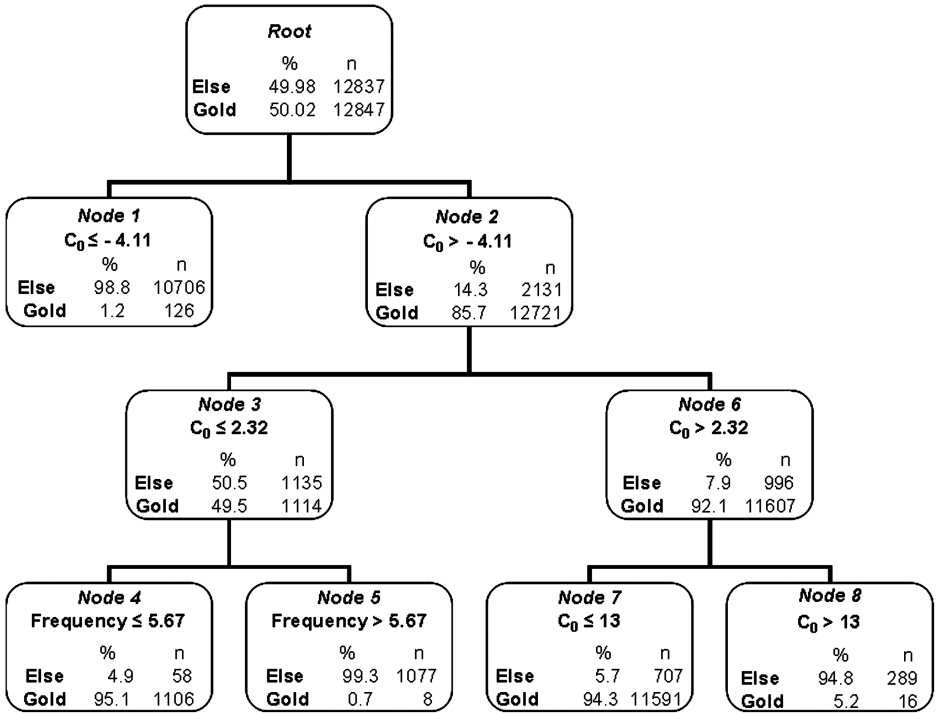

Fig. 7 Decision tree and the rules characterizing set 1 by comparison to the whole global Pareto set.

third quartile computed on set 1; this corresponds to the most efficient solutions of set 1 and is of particular interest to our study. Furthermore, the identification of the discriminant rules that characterize each one of the eight clusters is obtained by processing a decision tree for which the target variable is the cluster's belonging. The resulting decision tree processed under Clementine reveals with high accuracy the belonging of each solution to a given cluster by finding rules on the kinematic parameters. It appears that cluster 1 is qualified by two inequalities produced by the decision-tree algorithm as follows:

3) Rule 3:

$$A_1 \geq 43.40 \quad \text{and} \quad \phi_1 \leq -0.61 \tag{18}$$

2. Interpretation

The comparison of the thresholds on f_ψ found in rule (16) and the statistical indicators computed for the whole global Pareto set (see Table 1), shows that the frequency is inferior to the first quartile of the frequency values for the global Pareto set, which means that f_ψ takes very low values. On the other hand, one can see that in Eq. (16) C_0 takes high values and in Eq. (17) it takes very high values. Thus, we can say that to belong to set 1, which is equivalent to fulfilling Eqs. (6–8), either Eq. (16) or Eq. (17) has to be satisfied: either the upper-bounded mean dihedral angle takes high values and the frequency takes very low values or the mean dihedral angle takes very high values. One could think that rule (16) characterizes completely the highly efficient solutions of set 1. Actually, this rule tells us why this group of solutions is different from the other solutions in the global Pareto set in terms of predictor variables $(C_0, A_1, \phi_1, f_\psi)$, but it does not characterize the most efficient solutions of set 1. In this study, as we are interested in discovering the rules that qualify the most efficient solutions that fulfill Eqs. (6–8), we have performed a second step to characterize the most efficient solutions of set 1 by Kohonen maps and decision trees, which led to the characterization of cluster 1 by rule (18). If we compare the thresholds obtained in rule (18), which characterizes the most efficient solutions of set 1 (solutions of cluster 1), to the statistical indicators computed on the global Pareto set (see Table 1), we can see that the solutions of cluster 1 are characterized by negative low values to very low values of phase ϕ_1 and high values to very high values of amplitude A_1. Therefore, the most efficient solutions of set 1 (solutions of cluster 1) are characterized by two sets of rules. As they belong to set 1, they are qualified by rules (16) or (17) obtained by the decision tree, and as they are the most efficient solutions of set 1, they are also described by rule (18). We can then say that the most efficient solutions that fulfill Eqs. (6–8) are associated with high values to very high values of amplitude A_1 and to negative low values to very low values of phase ϕ_1 [see rule (18)] and either positive very high values of mean dihedral angle C_0 [Eq. (17)] or high values of mean dihedral angle C_0 and very low values of frequency f_ψ [Eq. (16)]. These rules are summarized in Table 4 in the column corresponding to set 1.

3. Analysis of Set 2

We recall here that the solutions that belong to set 2 have very low C_m values. As noted previously (see Sec. IV.C.1), we found rules to characterize set 2 by comparison to the rest of the solutions in the global Pareto set. The solutions of set 2 are called the gold group, and the other solutions in the database constitute the else group.

a. Characterization of Set 2 Versus the Global Pareto Set by Decision Trees. One more time, a rebalanced decision tree is processed and the precision of the tree is assessed by a contingency table that gives the risk estimate, which is equal to 1.33%. Set 2 (see Fig. 8) is characterized by two rules:

1) Rule 1:

$$C_0 \leq 2.32 \quad \text{and} \quad f_\psi \leq 5.58 \tag{19}$$

2) Rule 2:

$$C_0 \geq 2.32 \tag{20}$$

Table 4 Summary of the rules for set 1 and set 2

	Set 1	Set 2
Qualification		
Rule 1	$-4.11 \leq C_0 \leq 2.32$, $f_\psi \leq 5.67$	$C_0 \leq 2.32$, $f_\psi \leq 5.58$
Rule 2	$2.32 \leq C_0 \leq 13.05$	$C_0 \geq 2.32$
Most efficient solutions		
Rule 3	$A_1 \geq 43.40$, $\phi_1 \leq -0.61$	$A_1 \geq 44.07$, $\phi_1 \leq 0.45$

One can note that these two rules are quite similar to those rules (16) and (17) found for set 1. Rule (16) is almost identical to rule (19), except for the frequency threshold, and rule (17) is quite similar to rule (20), except that there is no upper limit for rule (20), which shows that the inequality on $|DC_z|$ that characterizes set 1 introduces an upper limit on the mean dihedral angle C_0. Indeed, we note that the dihedral angles between 13.05 and 18.5° have not been selected in set 1 (see Table 2), which shows that there is an influence of the mean dihedral angle on $|DC_z|$.

b. Characterization of the Inner Structure of Set 2 by Kohonen Maps. As previously noted (see Sec. IV.C.1), the inner structure of set 2 is analyzed with a typology method based on self-organizing Kohonen maps to qualify the most efficient solutions of set 2. Among the eight clusters produced by the algorithm, cluster 2 has been selected as it contains a high proportion of solutions for which the propulsive efficiencies are superior to the second and the third quartile; cluster 2 corresponds to the most efficient solutions of set 2 and is of particular interest to our study. By using a decision tree, the membership of each solution to a given cluster is explained by finding the rules on the predictor variables with a risk estimate of 0.45%. It appears that cluster 2 is characterized by the two following inequalities:

3) Rule 3:

$$A_1 \geq 44.07 \quad \text{and} \quad \phi_1 \leq 0.45 \tag{21}$$

If Eqs. (21) and (18) are compared, it appears that the two rules are quite similar. Indeed, at each time, the amplitude A_1 and the phase ϕ_1 are involved and the threshold obtained on the amplitude A_1 is quite the same in the two cases, but for the phase, even if the value of the threshold is different, the nature of the inequality remains the same (it is a smaller or equal inequality).

4. Interpretation

The comparison of the thresholds on f_ψ found in rule (19) to the statistical indicators computed for the whole global Pareto set (see

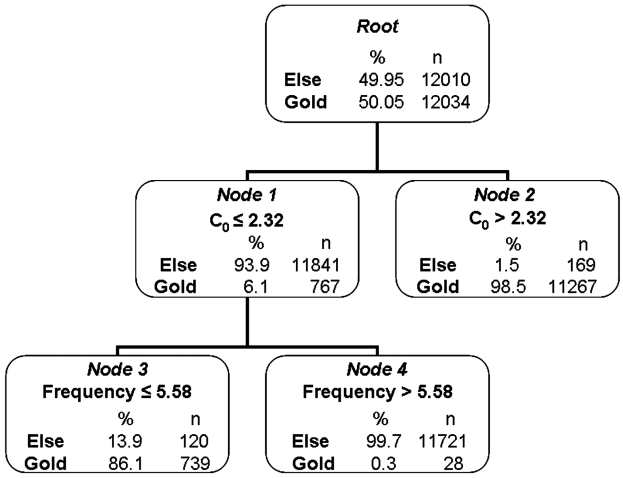

Fig. 8 Decision tree and the rules characterizing set 2 by comparison to the whole global Pareto set.

Table 1), shows that the frequency values of these solutions are inferior to the first quartile of the frequency values for the global Pareto set, which means that the solutions satisfying rule (19) are characterized by very low frequency values. C_0 takes high values in the case of rule (19) and very high values in the case of rule (20). Rules (19) and (20) show that to belong to set 2, either the frequency f_ψ has to be very low and the mean dihedral angle C_0 takes high values, or the mean dihedral angle just takes very high values. If we confront the thresholds obtained in rule (21) to the statistical indicators computed on the global Pareto set (see Table 1), we can see that the solutions of cluster 2 are characterized by low values to very low values of phase ϕ_1 and high values to very high values of amplitude A_1. The most efficient solutions of set 2 are characterized by two sets of rules. As they belong to set 2 they are qualified by the rules (19) or (20) obtained by the decision tree, and as they are the most efficient solutions of set 2 they are also described by rule (21). We can then say that the most efficient solutions that fulfill Eqs. (7) and (8) are associated with high values to very high values of amplitude A_1 and low values to very low values of phase ϕ_1 [see rule (21)] and either positive very high values of mean dihedral angle C_0 [Eq. (20)] or high values of mean dihedral angle C_0 and very low values of frequency f_ψ [Eq. (19)]. These rules are summarized in Table 4 in the column corresponding to set 2.

V. Conclusions

In conclusion, we have performed a constrained multiobjective optimization to find kinematics maximizing the propulsive efficiency for a simplified birdlike aerial vehicle in horizontal motion at slow cruise speed (6 m/s). The model in [13] has been used to describe the physics of the flapping-wing flight and evolutionary algorithms have been employed to perform the multiobjective constrained optimization. The study of the obtained global Pareto set showed that we can distinguish two groups of interest: the group of solutions that obey cruise-flight constraints (set 1) and the group of solutions that are stabilized in the moment (set 2). Using data-mining techniques embedded in the commercial software Clementine of SPSS, Inc. (decision trees and Kohonen maps), we identified the rules that characterize the most efficient solutions from the two sets. On one hand, it appeared that the most efficient solutions of set 1 are characterized by high values to very high values of amplitude A_1 and negative low values to very low values of phase ϕ_1 [see rule] and either positive very high values of mean dihedral angle C_0 [Eq. (17)] or high values of mean dihedral angle C_0 and very low values of frequency f_ψ [Eq. (16)]. On the other hand, it appeared that the most efficient solutions of set 2 are characterized by high values to very high values of amplitude A_1 and low values to very low values of phase ϕ_1 [see rule (21)] and either positive very high values of mean dihedral angle C_0 [Eq. (20)] or high values of mean dihedral angle C_0 and very low values of frequency f_ψ [Eq. (19)]. These rules are summarized in Table 4. This work presents a method to find and characterize by design rules the most efficient kinematics for birdlike flapping-wing UAVs: it has been demonstrated that highly accurate design rules for the most efficient Pareto-optimal solutions can be obtained by combining data-mining techniques such as Kohonen maps and decision trees. Simple kinematics and simple geometry have been considered in this study. One can think of complexifying the geometry or the kinematics to get a more realistic representation. Indeed, a pitching motion can be added to the dihedral motion and/or take a more realistic geometry, as described in [20], and/or use an articulated wing or take into consideration more harmonics in the description of the dihedral motion.

Appendix: Flight Physics Model

The flow's relative velocity V at $\frac{1}{4}$-chord location for each section is given by

$$V_x^2 = [V_c\cos(\theta) - \dot{h}\sin(\theta - \bar{\theta}_a)]^2$$

$$V_n^2 = \left[\dot{h}\cos(\theta - \bar{\theta}_a) - w_0 + \frac{1}{4}c\dot{\theta} + V_c\sin(\theta)\right]^2$$

$$V = \sqrt{V_x^2 + V_n^2}, \qquad \frac{w_0}{V_c} = \frac{2(\alpha_0 + \bar{\theta})}{2 + \mathrm{AR}} \tag{A1}$$

where h is the plunging displacement of leading edge in the flapping direction, $\bar{\theta} = \bar{\theta}_w + \bar{\theta}_a$ is the mean pitch angle, and w_0 is the downwash velocity at $\frac{3}{4}$-chord location. For pure dihedral motion we have

$$\dot{h} = r\dot{\psi}\cos(\psi)$$

$$\dot{\psi} = \sum_{n=1}^{2} 2\pi n f_\psi [C_n\cos(2\pi n f_\psi t) - B_n\sin(2\pi n f_\psi t)]$$

where r is the distance taken from the root flapping axis to the considered wing section. The angle of attack γ at the $\frac{3}{4}$-chord location are given by

$$\gamma = \tan^{-1}\left[\frac{\dot{h}\cos(\theta - \bar{\theta}_a) + \frac{1}{4}c\dot{\theta} + V_c\sin(\theta) - w_0}{V_c\cos(\theta) - \dot{h}\sin(\theta - \bar{\theta}_a)}\right]$$

The section's representative vortex $\mathrm{d}\Gamma$ (see Fig. A1) and the circulatory aerodynamic force $\mathrm{d}F_c$ are given by

$$\mathrm{d}\,\Gamma = V_c c\pi(\alpha' + \alpha_0 + \bar{\theta}), \qquad \mathrm{d}F_c = \rho V\,\mathrm{d}\Gamma\,\mathrm{d}r$$

with

$$\alpha' = \left[\frac{\dot{h}\cos(\theta - \bar{\theta}_a) + \frac{3}{4}c\dot{\theta} + V_c(\theta - \bar{\theta}) - w_0}{V_c}\right]$$

Then the circulatory normal force $\mathrm{d}N_c$ is

$$\mathrm{d}\,N_c = \mathrm{d}F_c\cos(\gamma)$$

An additional normal midchord force $\mathrm{d}N_a$, due to apparent mass effect, is given by

$$\mathrm{d}N_a = \frac{\rho\pi c^2}{4}\dot{v}_2\,\mathrm{d}r$$

$$\dot{v}_2 = \ddot{h}\cos(\theta - \bar{\theta}_a) - \dot{h}\,\dot{\theta}\sin(\theta - \bar{\theta}_a) + \frac{1}{2}c\ddot{\theta} + V_c\dot{\theta}\cos(\theta)$$

Therefore the section's total attached-flow normal force $\mathrm{d}N$ is

$$\mathrm{d}\,N = \mathrm{d}N_a + \mathrm{d}N_c$$

The chordwise forces due to camber $\mathrm{d}D_c$, leading-edge suction $\mathrm{d}T_s$, and friction drag $\mathrm{d}D_f$ are given by

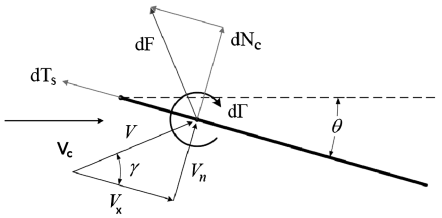

Fig. A1 Incidence correction adopted within the [13] model.

$$\mathrm{d}D_c = -2\pi\alpha_0(\alpha' + \bar{\theta})\cos(\gamma)\frac{\rho V_c V}{2}c\,\mathrm{d}r$$

$$\mathrm{d}D_f = C_{d_f}\frac{\rho V_x^2}{2}c\,\mathrm{d}r, \qquad \mathrm{d}T_s^1 = \left(\alpha' + \bar{\theta} - \frac{1}{4}\frac{c\dot{\theta}}{V_c}\right)\cos(\gamma)$$

$$\mathrm{d}T_s^2 = (\alpha' + \alpha_0 + \bar{\theta})\sin(\gamma)$$

$$\mathrm{d}T_s = \eta_s 2\pi\left[(\mathrm{d}T_s^1 + \mathrm{d}T_s^2)\frac{\rho V_c V}{2}c\,\mathrm{d}r\right]$$

The section friction-drag coefficient C_{d_f} is computed by [2]

$$C_{d_f} = \frac{0.89}{[\log(Re)^{2.58}]}$$

with the Reynolds number Re based on the chord and V. The leading-edge suction efficiency η_s is fixed to 0.98. Thus, the total chordwise force $\mathrm{d}F_x$ is

$$\mathrm{d}F_x = \mathrm{d}T_s - \mathrm{d}D_c - \mathrm{d}D_f$$

The poststall behavior is locally modeled (for each section) by using a stall criterion. The stall occurs when

$$\gamma - \frac{3}{4}\left[\frac{c\dot{\theta}}{V_c}\right] \geq \alpha_{\mathrm{sep}_{\max}}$$

where $\alpha_{\mathrm{sep}_{\max}} = 13^\circ$ or 0.23 rad. Then it is assumed that $\mathrm{d}D_c = 0$ and the other forces are given by

$$\mathrm{d}N = \mathrm{d}N_{c_{\mathrm{sep}}} + \mathrm{d}N_{a_{\mathrm{sep}}}, \qquad V_n = \dot{h}\cos(\theta - \bar{\theta}_a) + \frac{1}{2}c\dot{\theta} + V_c\sin(\theta)$$

$$\hat{V} = (V_x^2 + V_n^2)^{\frac{1}{2}}, \qquad \mathrm{d}N_{c_{\mathrm{sep}}} = C_{d_{cf}}\frac{\rho\hat{V}V_n}{2}c\,\mathrm{d}r, \qquad \mathrm{d}N_{a_{\mathrm{sep}}} = \frac{1}{2}\mathrm{d}N_a$$

$$\mathrm{d}D_{f_{\mathrm{sep}}} = C_{d_{f_{\mathrm{sep}}}}\frac{\rho V_x^2}{2}c\,\mathrm{d}r$$

$$\mathrm{d}T_{s_{\mathrm{sep}}} = \eta_{s_{\mathrm{sep}}}2\pi\left(\alpha' + \bar{\theta} - \frac{1}{4}\frac{c\dot{\theta}}{V_c}\right)\cos(\gamma)\frac{\rho V_c V}{2}c\,\mathrm{d}r$$

The dynamic stall phenomena is accounted for through $\mathrm{d}T_{s_{\mathrm{sep}}}$ and $\mathrm{d}D_{f_{\mathrm{sep}}}$, which contribute to the normal force (see Fig. A2). The constants $\eta_{s_{\mathrm{sep}}}$ and $C_{d_{f_{\mathrm{sep}}}}$ were taken equal to, respectively, 1.491 and 0.065 [13]. Now we can derive the equations for section lift $\mathrm{d}L$ and thrust $\mathrm{d}T$ forces as follows:

$$\mathrm{d}L = \mathrm{d}N\cos(\theta) + \mathrm{d}F_x\sin(\theta), \qquad \mathrm{d}T = \mathrm{d}F_x\cos(\theta) - \mathrm{d}N\sin(\theta)$$

On integration along the span and in time, we get the time-averaged lift $\bar{L}$ and thrust $\bar{T}$ forces as follows:

$$\bar{L} = 2f_\psi\int_0^{f_\psi}\int_0^{\frac{b}{2}}\cos(\psi)\,\mathrm{d}L\,\mathrm{d}t, \qquad \bar{T} = 2f_\psi\int_0^{f_\psi}\int_0^{\frac{b}{2}}\mathrm{d}D\,\mathrm{d}t$$

One may also compute the instantaneous power required to move the section against its aerodynamic loads and the norm of the aerodynamic moment $\mathrm{d}M_{\mathrm{aero}}$ about the center of gravity of the vehicle, assuming permanent symmetrical dihedral motion. For attached flow it is given by

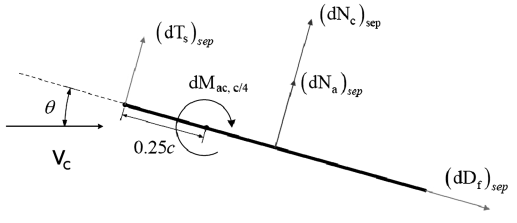

Fig. A2 Modelization of dynamic stall within the [13] model.

$$\mathrm{d}P_{\mathrm{in}} = \mathrm{d}F_x\dot{h}\sin(\theta - \bar{\theta}_a) + \mathrm{d}N\left[\dot{h}\cos(\theta - \bar{\theta}_a) + \frac{1}{4}c\dot{\theta}\right]$$
$$+ \mathrm{d}N_a\left[\frac{1}{4}c\dot{\theta}\right] - \mathrm{d}M_{\mathrm{ac}}\dot{\theta} - \mathrm{d}M_a\dot{\theta}$$

$$\mathrm{d}M_p = \mathrm{d}M_{\mathrm{ac}} + \mathrm{d}M_a - \mathrm{d}N_a\left[\frac{1}{4}c - e\right] - \mathrm{d}N_c\left[\frac{1}{2}c - e\right]$$

$$\mathrm{d}M_{\mathrm{ac}} = \frac{1}{2}\rho U^2 C_{m_{\mathrm{ac}}}c^2\,\mathrm{d}r$$

$$\mathrm{d}M_a = -\left[\frac{1}{16}\rho\pi c^3\dot{\theta}V_c + \frac{1}{128}\rho\pi c^4\ddot{\theta}\right]\mathrm{d}r$$

$$\mathrm{d}M_{\mathrm{aero}}^2 = [\sin(\psi)r\mathrm{d}T]^2 + 4[\mathrm{d}M_p\sin(\psi)]^2$$

For stalled flow we have

$$\mathrm{d}P_{\mathrm{in}_{\mathrm{sep}}} = \mathrm{d}N_{\mathrm{sep}}\left[\dot{h}\cos(\theta - \bar{\theta}_a) + \frac{1}{2}c\dot{\theta}\right]$$

$$\mathrm{d}M_{p_{\mathrm{sep}}} = -[\mathrm{d}N_{c_{\mathrm{sep}}} + \mathrm{d}N_{a_{\mathrm{sep}}}]\left[\frac{1}{2}c - e\right]$$

$$\mathrm{d}M_{\mathrm{aero}}^2 = [\sin(\psi)r\mathrm{d}T]^2 + 4[\mathrm{d}M_{p_{\mathrm{sep}}}\sin(\psi)]^2$$

with $e = 0.25$. We can also derive expressions for the span-integrated and time-averaged power and norm of the aerodynamic moment as follows:

$$\bar{P}_{\mathrm{in}} = 2f_\psi\int_0^{f_\psi}\int_0^{\frac{b}{2}}\mathrm{d}P_{\mathrm{in}}\,\mathrm{d}t, \qquad \bar{M}_{\mathrm{aero}} = f_\psi\int_0^{f_\psi}\left\{\int_0^{\frac{b}{2}}\mathrm{d}M_{\mathrm{aero}}^2\right\}^{\frac{1}{2}}\mathrm{d}t$$

Finally, the average propulsive efficiency is defined as follows:

$$\bar{\eta} = \frac{\bar{T}V_c}{\bar{P}_{\mathrm{in}}}$$

References

[1] Herzog, K., "Flapping Wing Flight in Nature and Science," *Aeromodeller Annual 1964-65*, Model Aeronautical Press, Watford, England, U.K., 1965, pp. 44–57.

[2] DeLaurier, J., "An Aerodynamic Model for Flapping-Wing Flight," *The Aeronautical Journal*, Vol. 97, No. 964, 1993, pp. 125–130.

[3] DeLaurier, J., "The Development and Testing of a Full-Scale Piloted Ornithopter," *Canadian Aeronautics and Space Journal*, Vol. 45, No. 2, 1999, pp. 72–82.

[4] Larijani, R., and DeLaurier, J., "A Nonlinear Aeroelastic Model for the Study of Flapping Wing Flight," *Fixed and Flapping Wing Aerodynamics for Micro Air Vehicle Applications*, Progress in Astronautics and Aeronautics, Vol. 195, AIAA, Reston, VA, 2001, pp. 399–428.

[5] Rakotomamonjy, T., Ouladsine, M., and Le Moing, T., "Modelization and Kinematics Optimization for a Flapping-Wing Microair Vehicle," *Journal of Aircraft*, Vol. 44, No. 1, 2007, pp. 217–231. doi:10.2514/1.22960

[6] Doncieux, S., Mouret, J., Angeli, A., Barate, R., Meyer, J., and de Margerie, E., "Building an Artificial Bird: Goals and Accomplishments of the ROBUR Project," *Proceedings of the European Micro Aerial Vehicles Conference (EMAV 2006)*, 2006.

[7] Khan, Z., and Agrawal, S., "Force and Moment Characterization of Flapping Wings for Micro Air Vehicle Application," *Proceedings of the 2005 American Control Conference*, American Control Council, Evanston, IL, 2005, pp. 1515–1520.

[8] Hall, K., Pigott, S., and Hall, S., "Power Requirements for Large-Amplitude Flapping Flight," *Journal of Aircraft*, Vol. 35, No. 3, 1998, pp. 352–361. doi:10.2514/2.2324

[9] Hall, K. C., and Hall, S. R., "Minimum Induced Power Requirements for Flapping Flight," *Journal of Fluid Mechanics*, Vol. 323, No. -1, 1996, pp. 285–315. doi:10.1017/S0022112096000924

[10] Willis, D., Peraire, J., Drela, M., and White, J., "A Numerical Exploration of Parameter Dependence in Power Optimal Flapping Flight," 24th AIAA Applied Aerodynamics Conference, San Francisco,

AIAA Paper 2006-2994, 2006.
[11] Berman, G., and Wang, Z., "Energy-Minimizing Kinematics in Hovering Insect Flight," *Journal of Fluid Mechanics*, Vol. 582, 2007, pp. 153–168. doi:10.1017/S0022112007006209
[12] de Margerie, E., Mouret, J., Doncieux, S., and Meyer, J., "Artificial Evolution of the Morphology and Kinematics in a Flapping-Wing Mini-UAV," *Bioinspiration & Biomimetics*, Vol. 2, 2007, p. 65. doi:10.1088/1748-3182/2/4/002
[13] Kim, D., Lee, J., Lee, J., and Han, J., "An Aeroelastic Analysis of a Flexible Flapping Wing Using Modified Strip Theory," *Active and Passive Smart Structures and Integrated Systems*, Proceedings of SPIE, Vol. 6928, SPIE, Bellingham, WA, 2008, Paper 69281O. doi:10.1117/12.776137
[14] Deb, K., and Srinivasan, A., "Innovization: Discovery of Innovative Design Principles Through Multiobjective Evolutionary Optimization," *Multiobjective Problem Solving from Nature: From Concepts to Applications*, Nature Computing Series, Springer, New York, 2007, p. 243.
[15] Deb, K., *Multi-Objective Optimization using Evolutionary Algorithms*, Wiley-Interscience Series in Systems and Optimization, Wiley, Chichester, England, U.K., 2001.
[16] Bellman, R., *Adaptive Control Processes: A Guided Tour*, Vol. 1, Princeton University Press, Princeton, NJ, 1961.
[17] Zilinskas, A., Fraga, E., and Mackute, A., "Data Analysis and Visualisation for Robust Multi-Criteria Process Optimisation," *Computers and Chemical Engineering*, Vol. 30, Nos. 6–7, 2006, pp. 1061–1071. doi:10.1016/j.compchemeng.2006.02.003
[18] Goel, T., Vaidyanathan, R., Haftka, R., Shyy, W., Queipo, N., and Tucker, K., "Response Surface Approximation of Pareto-Optimal Front in Multi-Objective Optimization," *Computer Methods in Applied Mechanics and Engineering*, Vol. 196, Nos. 4–6, 2007, pp. 879–893. doi:10.1016/j.cma.2006.07.010
[19] Liebeck, R., "Design of the Blended Wing Body Subsonic Transport," *Journal of Aircraft*, Vol. 41, No. 1, 2004, pp. 10–25. doi:10.2514/1.9084
[20] Liu, T., Kuykendoll, K., Rhew, R., and Jones, S., "Avian Wing Geometry and Kinematics," *AIAA Journal*, Vol. 44, No. 5, 2006, pp. 954–963. doi:10.2514/1.16224
[21] Souied, H., "Optimization of the Kinematics of Flapping Airfoils for Micro Air Vehicle Applications," Ph.D. Thesis, Università di Genova, Genova, Italy, and Université Paul Sabatier, Toulouse, France, 2008.
[22] Anderson, J., Streitlien, K., Barrett, D., and Triantafyllou, M., "Oscillating Foils of High Propulsive Efficiency," *Journal of Fluid Mechanics*, Vol. 360, 1998, pp. 41–72. doi:10.1017/S0022112097008392
[23] Pedro, G., Suleman, A., and Djilali, N., "A Numerical Study of the Propulsive Efficiency of a Flapping Hydrofoil," *International Journal for Numerical Methods in Fluids*, Vol. 42, No. 5, 2003, pp. 493–526. doi:10.1002/fld.525
[24] Isogai, K., and Harino, Y., "Optimum Aeorelastic Design of a Flapping Wing," *Journal of Aircraft*, Vol. 44, No. 6, 2007, pp. 2040–2048. doi:10.2514/1.27142
[25] Harada, M., "Calculation Method for Optimal Circulation Distribution on a Finite Span Flapping Wing," 1st Unmanned Aerospace Vehicles, Systems, Technologies, and Operations Conference and Workshop, AIAA, Paper 2002-3414, 2002.
[26] DeLaurier, J., "The Development of an Efficient Ornithopter Wing," *The Aeronautical Journal*, Vol. 97, 1993, pp. 153–153.
[27] Andersson, J., "A Survey of Multiobjective Optimization in Engineering Design," Dept. of Mechanical Engineering, Linköping University, Rept. LiTH-IKP-R-1097, Linköping, Sweden, 2000.
[28] Deb, K., Mohan, M., and Mishra, S., "Towards a Quick Computation of Well-Spread Pareto-Optimal Solutions," *Evolutionary Multi-Criterion Optimization*, Springer, New York, 2003, pp. 68–68.
[29] Shyy, W., Berg, M., and Ljungqvist, D., "Flapping and Flexible Wings for Biological and Micro Air Vehicles," *Progress in Aerospace Sciences*, Vol. 35, No. 5, 1999, pp. 455–505. doi:10.1016/S0376-0421(98)00016-5
[30] Landau, S., Doncieux, S., Drogoul, A., and Meyer, J.-A., "SFERES: Un Framework Pour la Conception de Systmes Multi-Agents Adaptatifs," *Technique et Science Informatiques*, Vol. 21, No. 4, 2002, pp. 427–446.
[31] Rokach, L., and Maimon, O., *Data Mining with Decision Trees: Theory and Applications*, World Scientific, Singapore, ROK, 2008.
[32] Kohonen, T., *Self-Organizing Maps*, 3rd ed., Vol. 1, Springer Series in Information Sciences, Vol. 30, Springer, Berlin, 2001.
[33] Seiffert, U., and Jain, L., *Self-Organizing Neural Networks: Recent Advances and Applications*, Physica Verlag, Heidelberg, Germany, 2002.
[34] Clenshaw, C., and Curtis, A., "A Method for Numerical Integration on an Automatic Computer," *Numerische Mathematik*, Vol. 2, No. 1, 1960, pp. 197–205. doi:10.1007/BF01386223
[35] Radespiel, R., Windte, J., and Scholz, U., "Numerical and Experimental Flow Analysis of Moving Airfoils with Laminar Separation Bubbles," *AIAA Journal*, Vol. 45, No. 6, 2007, pp. 1346–1356. doi:10.2514/1.25913
[36] Shyy, W., Aono, H., Chimakurthi, S., Trizila, P., Kang, C.-K., Cesnik, C., and Liu, H., "Recent Progress in Flapping Wing Aerodynamics and Aeroelasticity," *Progress in Aerospace Sciences* (to be published); available online Feb. 2010. doi:10.1016/j.paerosci.2010.01.001
[37] Shyy, W., Lian, Y., Tang, J., Viieru, D., and Liu, H., *Aerodynamics of Low Reynolds Number Flyers*, Cambridge Aerospace Series, Vol. 22, Cambridge University Press, New York, Oct. 2007.
[38] Tobalske, B., and Dial, K., "Flight Kinematics of Black-Billed Magpies and Pigeons over a Wide Range of Speeds," *Journal of Experimental Biology*, Vol. 199, No. 2, 1996, p. 263; available online at http://jeb.biologists.org/cgi/reprint/199/2/263.pdf.
[39] Roudenko, O., "Application Des Algorithmes Évolutionnaires aux Problèmes d'Optimisation Multi-Objectif avec Contraintes," Ph.D. Thesis, Ecole Polytechnique, Palaiseau, France, 2004.