AIAA JOURNAL
Vol. 49, No. 2, February 2011

Airfoil Data Fitting Using Multivariate Smoothing Thin Plate Splines

Rita Ponza and Ernesto Benini*
University of Padova, 35131 Padova, Italy

DOI: 10.2514/1.J050585

Smoothing thin plate splines, a fitting technique based on a rigorous roughness penalty approach, have been recently investigated as a promising tool for bivariate interpolation of aerodynamic data. In this paper, this technique is implemented and extended to multivariate fitting. In particular, the method is applied for estimating the aerodynamic polars of well-known two-dimensional symmetrical and nonsymmetrical airfoils as functions of some geometric parameters describing the airfoil shape and a further variable defining the flow regime (either the Mach or the Reynolds number). Therefore, the simultaneous influence of five independent variables on three responses (lift, drag, and pitching moment coefficients) is investigated. To this purpose, a large database is generated via numerical simulations (using a validated flow solver) containing all information required to build a reliable response surface. Then, the model is built and its performance validated by performing queries on complete aerodynamic polars at various flow regime conditions of a series of airfoils not included into the database. Results show a very good matching between predicted and calculated curves, thus demonstrating the remarkable predictive capability of the implemented tool.

Nomenclature

$\mathbf{a}$	=	vector of coefficients in thin plate splines
CD	=	drag coefficient
CL	=	lift coefficient
CM	=	pitching moment coefficient
d	=	dimension of the independent variables vector
$f(\mathbf{x})$	=	systematic component of the functional relationship between $\mathbf{y}$ and $\mathbf{x}$
$J_m(f)$	=	functional in function f
M	=	Mach number
m	=	order of derivative used to measure roughness
n	=	number of observations
R	=	determination coefficient
Re	=	Reynolds number
$S_{md}(f)$	=	functional in function f
$\mathbf{x}$	=	vector of independent variables, independent parameters, predictors
$\mathbf{y}$	=	vector of observed values, or output variables
$\hat{\mathbf{y}}$	=	estimated response vector
$y1_v, \ldots y6_v$	=	geometric independent variables for the volume distribution curve
$x_{\max}$camb	=	chordwise location of maximum camber
$y_{\max}$camb	=	maximum camber value
α	=	smoothing parameter in thin plate splines, angle of attack
η_{md}	=	function of the Euclidean distance in the independent variable space
v	=	integer number
ϕ	=	polynomial in $\mathbb{R}^d$ with degree less than m
$\boldsymbol{\delta}$	=	vector of weights in thin plate splines

I. Introduction and Background

RELIABLE fitting of aerodynamic data, such as airfoil coefficients derived from either experiments or simulations, is still one of the biggest issues in aeronautics. In fact, the outcome of test campaigns often consists in large databases including complex dependencies between responses and multiple independent variables, thus claiming for robust multivariate fitting techniques in order for the dataset to be profitably used for aerodynamic data modeling. Mathematically speaking, the problem above can be viewed as finding the "best" approximation of an input–output function which maps from the available sample data into the response space.

There have been many ways in which researchers have dealt with such a problem.

Parametric fitting, for example, involves determination of a linear or nonlinear response model (often referred to as a response surface) using standard regression techniques [1–3] and have been successfully applied, for instance in [4], to multivariate fitting of aerodynamic characteristics of a complex aircraft configuration. Furthermore, in [5,6] a linear parametric model was built with the aim of estimating the drag polars at different attitudes of a joined wing concept aircraft, where the parametric model was eventually used to provide reliable data to the flight mechanics tools. The common basis for constructing a parametric model is that some a priori knowledge of the phenomenon being investigated is mandatory to get reliable predictions. Therefore, in most cases the experience of the aerodynamicist plays a dominant role in model building: while this is fundamental to preserve adherence of the model to its physical background, it could provide a somewhat forced and undesirable bias in the model itself. Moreover, when highly multivariate problems are tackled, i.e., featuring a large number of independent variables and multiple response functions, an effective method to bring an a priori knowledge into a viable model becomes very difficult to be identified.

On the other hand, conventional nonparametric fitting deals with the problem of finding a response surface without making general assumptions with respect to its analytical form [7–9]. The basic hypothesis relies in the fact that the response surface is assumed to obey to a few and rather general smoothness conditions. The very attractive feature of this approach is that data to be fitted is not forced into a prescribed mathematical structure in order for the unknown model parameters to be determined, but they are left free to build the statistical model on their own without being trapped into a predefined, constrained formulation. Some relevant examples of nonparametric methods for data fitting or metamodel construction in aeronautics have already been documented, which include multivariate adaptive regression splines [10,11], neural networks [12–14], and radial basis functions [15–17], each of which has some

Received 31 March 2010; revision received 9 September 2010; accepted for publication 24 October 2010.

*Research Professor, Department of Mechanical Engineering, Via Venezia; ernesto.benini@unipd.it (Corresponding Author). Member AIAA.

advantages and drawbacks [18]. Other more recently developed nonparametric techniques, whose application to aeronautics is far less diffused, encompass support vector regression [19–21], regression Kriging [22–25], and moving least-squares or local polynomial regression [26–30].

In a previous paper [31], the authors proposed an application to aerodynamic data fitting based on a "smoothing thin plate spline" technique, a nonparametric methodology which proved to be robust and reliable for a bivariate problem, i.e., the fitting of aerodynamic data of a sample three-dimensional-shaped body over a wide range of two attitude angles (pitch and yaw). In that work, the method demonstrated its superiority with respect to all the cited concurrent strategies, especially when available data was sparse, as resulted from a cross-validation analysis. The reader is referred to [31] for a complete treatment of the smoothing thin plate splines technique, including comparison with concurrent parametric and nonparametric fitting methods.

In the present paper, a true multivariate problem is investigated using the smoothing thin plate splines approach, where the influence of more than two independent variables is considered at the same time, with the aim of showing how well the methods work when more complex predictive models are to be constructed. Specifically, an application of the smoothing thin plate spline technique for determining the aerodynamic polars of 2-D airfoils as functions of some geometric parameters describing the shape of the airfoils and a further variable defining the flow regime will be illustrated. This example has been selected because of its general significance in aeronautics, as well as because a large amount of airfoil data can be resumed from the open literature or can be generated throughout numerical analyses by almost any researcher, therefore encompassing the so-called *curse of dimensionality* problem [3]. In fact, the amount of data needed to build a multivariate model is an exponential function of the number of independent dimensions to be accounted for, i.e., the number of independent variables. Since almost an uncountable number of airfoil shapes and related aerodynamic coefficients are easily producible, this seems the best trial to verify the goodness of the proposed approach in a multidimensional domain.

It is worth underlying that the purpose here is not to compare the results of the various fitting approaches (to this respect, the reader is referred to [31] for a detailed discussion), rather to highlight the capabilities of the method to deal with a complex interpolating problem without producing under or overfitting, and to demonstrate its smoothing and adaptive features which allow to produce very reliable responses.

II. Brief Review of Smoothing Thin Plate Splines

Thin plate splines (TPS) have been known for quite a long time in the field of applied mathematics, where they were originally introduced for geometric design: specifically, the name *thin plate spline* refers to a physical analogy involving the bending of a thin sheet of metal. What is still unexplored is their application to the fitting of multivariate aerodynamic data, which is the subject of the present work. In particular, we will refer to what is often called a smoothing thin plate spline (STPS) in the works by Wahba [32] and Green and Silverman [9].

The STPS approach for multivariate regression is based on the idea of penalizing a response surface based on its roughness: basically, this method originates from the relaxation of some of the assumptions needed when building a model with the classical polynomial regression techniques. Specifically, when searching for an approximation function $\hat{\mathbf{y}} = f(\mathbf{x})$ with a least-squares approach, only the surface fitting to the observed data is to be maximized, and usually no additional constraint is imposed to f. Obviously, the least-squares sum goes down to zero if f exactly interpolates the data; however, many different surfaces exist passing through the observed values and, even if some generic smoothness conditions are imposed to f to make it continuous up to a specified derivative order, the resulting surface may not be satisfactory, being potentially subject to excessive fluctuations. Actually, a good data fitting should not be the unique goal in the identification of a response model: a further and often conflicting objective should be to obtain a surface free from undesired fluctuating behavior (absence of distortions).

This rather qualitative notion is actually quantified through the "roughness penalty approach." Hence, the problem of identifying an adequate model of the observed data may be stated so as to clearly highlight the need for a tradeoff between the above-mentioned objectives of minimization of residual errors and abatement of the local fluctuations. Specifically, what acts like a balancing factor between the two objectives in the TPS formulation is a smoothing parameter α.

The main result of the roughness penalty approach is that natural thin plate splines can be demonstrated to be the unique solutions to the two-objective data fitting problem described above, and this makes the identification of the response surface a fully deterministic problem, where nothing needs to be imposed by the analyst except for the smoothing parameter, which on the other hand may be automatically determined.

For mathematical details on the univariate STPS formulation the reader is referred to [31].

III. Multivariate STPS

The STPS approach can be extended quite in straightforward way to functions of more than one or two variables [9]. Consider for instance the problem of finding out a proper, sufficiently smooth estimator f to build up a model of the form $y_i = f(\mathbf{x_i}) + \text{error}$ $i = 1, \ldots, n$ for n observations y_i, where $\mathbf{x_i}$ is a d-dimensional vector.

As for the two-dimensional case, f turns out to be the solution of a minimization problem. Specifically, the function to be minimized is a penalized sum of squares:

$$S_{md}(f) = \sum_{i=1}^{n} \{y_i - f(\mathbf{x_i})\}^2 + \alpha J_m(f) \tag{1}$$

where $\alpha > 0$ is a smoothing parameter and $J_m(f)$ is a functional quantifying the smoothness of f, being m the order of the derivative (higher than the second) used to measure roughness.

The expression for $J_m(f)$ in Eq. (1), that indicates a penalty in d dimensions based on the mth derivative, is:

$$J_m(f) = \int \ldots \int_{\mathbb{R}^d} \sum \frac{m!}{\nu_1! \ldots \nu_d!} \left(\frac{\partial^m f}{\partial x_1^{\nu_1} \ldots \partial x_d^{\nu_d}} \right)^2 \mathrm{d}x_1 \ldots \mathrm{d}x_d \tag{2}$$

where the sum within the integral is extended over all the nonnegative integers $\nu_1, \nu_2, \ldots, \nu_d$ such that $\nu_1 + \nu_2 + \ldots + \nu_d = m$. According to this definition, the only surfaces for which $J_m(f) = 0$ are polynomials of degree less than m. Moreover, it is necessary to impose the condition $2m > d$, so that roughness functional J_m based on integrated first derivatives can be used only for one-dimensional problems, those based on integrated second derivatives only for three or less dimensions, and so on. The reason for that is expressed mathematically in terms of Beppo Levi and Sobolev spaces [33].

The optimization problem formulated in Eq. (1) can be faced by considering a peculiar, finite dimensional class of functions f: let us define a function η_{md}

$$\eta_{md}(r) = \begin{cases} \theta r^{2m-d} \log r & \text{if } d \text{ is even} \\ \theta r^{2m-d} & \text{if } d \text{ is odd} \end{cases} \tag{3}$$

where the constant of proportionality θ has the following expression:

$$\theta = \begin{cases} (-1)^{m+1+\frac{d}{2}} 2^{1-2m} \pi^{-\frac{d}{2}} (m-1)!^{-1} \left(m - \frac{d}{2} \right)!^{-1} & \text{if } d \text{ is even} \\ \Gamma\left(\frac{d}{2} - m \right) 2^{-2m} \pi^{-\frac{d}{2}} (m-1)!^{-1} & \text{if } d \text{ is odd} \end{cases} \tag{4}$$

Finally, we define

$$M = \binom{m+d-1}{d}$$

and we focus our attention on a class $\{\phi_j, j = 1, 2, \ldots, M\}$ of linearly independent polynomials spanning the M-dimensional space of polynomials in $\mathbb{R}^d$ with degree less than m. A function f on $\mathbb{R}^d$ is called a natural thin plate spline of order m if it has the form:

$$f(\mathbf{x}) = \sum_{i=1}^{n} \delta_i \eta_{md}(\|\mathbf{x} - \mathbf{x_i}\|) + \sum_{j=1}^{M} a_j \phi_j(\mathbf{x}) \tag{5}$$

with the coefficient vector δ satisfying the condition $\mathbf{T}\delta = 0$, being $T_{ij} = \phi_i(\mathbf{x_j})$.

Analogously to the two-dimensional case, it can be demonstrated that, provided that the points $\mathbf{x_i}$ are distinct and sufficiently dispersed to determine a unique least-squares polynomial surface of degree $m-1$, and under the condition $2m > d$, the function f that minimizes $J_m(f)$ under the constraint $f(\mathbf{x_i}) = y_i$ is a natural thin plate spline of order m.

It follows that the function that minimizes S_{md} is a natural thin plate spline of order m, with coefficient vectors $\mathbf{a}$ and δ uniquely identified by:

$$\begin{bmatrix} \mathbf{E} + \alpha\mathbf{I} & \mathbf{T}^T \\ \mathbf{T} & 0 \end{bmatrix} \begin{pmatrix} \delta \\ a \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix} \tag{6}$$

being the generic $\mathbf{E}$ matrix element $E_{ij} = \eta_{md}(\|\mathbf{x_i} - \mathbf{x_j}\|)$.

The identification of the optimal smoothing parameter value α for an assigned observation set can be devolved to automatic procedures, the most common of them being based on a generalized cross-validation (GCV) approach, which privileges the predictive capabilities of the model over unsampled input vectors rather than data fitting [31]. In fact, from the minimization of the GCV score:

$$\text{GCV}(\alpha) = n^{-1} \sum_{i=1}^{n} \left\{ \left(\frac{1 - A_{ii}(\alpha)}{1 - n^{-1}\text{tr}\, A(\alpha)} \right)^2 \{y_i - \hat{f}^{(-i)}(\mathbf{x}_i)\}^2 \right\} \tag{7}$$

the optimal value of the smoothing parameter can be obtained.

When applying STPS to a single data set, usually no additional observations are available over which prediction capabilities of the model may be assessed: this obstacle is moved around within the GCV approach, by considering each of the assigned data as a potential "new observation"; specifically, a response surface may be derived from the whole data set except one observation pair, and then the surface's predictive capability may be judged over the input vector-point that has been omitted from the data when building the model. Extending this procedure to all the input point-vectors, while keeping fixed the value of the smoothing parameter, gives a global measure of the model predictive efficiency. In particular, a cross-validation score function may be defined which represents the target function to be minimized to find the optimal value for α.

An example of pseudocode for STPS costruction is the following:

1) Acquire and store n-observed data $\mathbf{x_i}$ (d-dimensional) and y_i $\forall\ i = 1, \ldots, n$.

2) Calculate $\eta_{md}(r)$ based on the m-th derivative order and space dimensionality d.

3) Calculate elements of matrix $\mathbf{T}$.

4) Calculate the smoothing parameter α based on the GCV criterion.

a) Calculate GCV using Eq. (7).

b) Minimize GCV iteratively using tentative pseudodata $\hat{f}$ or use Gu and Wahba [34] and Wood's [35] decomposition methods.

5) Solve system of Eqs. (6) using Turk and O'Brien's lower and upper (LU) triangular matrices factorization [36].

6) Evaluate interpolating function.

A first significant portion of the computational effort involved in calculating STPS is the cost required to build the matrix $\mathbf{E}$. Such matrix is entirely nonzero except along the diagonal (where $i = j$), requiring the calculation of all interpoint distances within the set including all x_i-s. Such complexity can be evaluated in the order of $\mathcal{O}(n^2 d)$.

Another troublesome term is $\text{tr}(\mathbf{A})$, which would take $\mathcal{O}(nd^2)$ operations to evaluate directly, for each trial set of smoothing parameters. A simple grid search for iterative m smoothing parameters using k grid points per parameter would require $\mathcal{O}(nd^2k^m)$ floating point operations. To allow practical use of models with multiple tentative penalties, an efficient method proposed by Wood [35] can be used for ($\mathcal{O}(d^3)$ complexity). Turk and O'Brien's LU factorization to solve Eq. (6) requires $\mathcal{O}(n^3)$ computations by iterative means [36]. Thus, while solution of the system may appear to be the limiting step, it needs only be as computationally expensive as constructing the system. Finally evaluating resulting function at potentially many points to extract the isosurface, calculate normals, or other derivative quantities, etc., requires $\mathcal{O}(nd)$ evaluations.

IV. Curse of Dimensionality

When trying to map a high-dimensional input variables' space into an output space, a series of difficulties usually arise: specifically, multivariate data is difficult to work with because of the relevant number of observations that are necessary to get good estimates. Furthermore, adding more features to the explanatory variables' space increases their interdependency relationships and can also cause an augmented noise, which may adversely affect prediction reliability. This is usually referred to as the "curse of dimensionality" [37].

Actually, statistical data fitting tools infer knowledge, or information, from available samples. Obviously, the models built through data fitting are only valid in the space region where sampling data is available. Whatever the model or class of models, generalization on data too much different from all available observations is impossible.

Hence, one of the key issues in a successful development of a model building technique is to have enough available data to properly fill the domain or part of the domain where the model is required to be valid. It is apparent that, every other constraint being kept unchanged, the number of observations should grow exponentially with the input space dimensions (e.g., if 10 data points seems reasonable for a smooth one-dimensional model, 100 will be necessary for a two-dimensional model with the same smoothness, 1000 for a three-dimensional model, and so on). This exponential increase is the first consequence of the curse of dimensionality. More generally, the curse of dimensionality is the expression of all phenomena that arise with high-dimensional data, and that have most often undesirable consequences on the behavior and performance of data fitting algorithms. Specifically, a nonparametric model building approach may show dramatically deteriorated prediction performance, unless it is fit with a proper number of independent observations.

In this work, the effectiveness of STPS as a model building technique in a multivariate input data scenario will be assessed. Specifically, an application of the method for determining the aerodynamic polars of 2-D airfoils as functions of some geometric parameters describing the shape of the airfoils and a further variable defining the flow regime will be illustrated. It is now apparent that, even keeping to a minimum the amount of geometric parameters to be introduced into the model, a relevant number of data (either experimental or numerical) on different airfoil shapes is needed to properly fit the model and get a reliable analysis of the geometric configuration's effects on the aerodynamic performance. For this reason, 2-D airfoils are suitable for such an application, since large databases are available either in the open literature or by simulating them by means of 2-D codes.

V. Airfoil Data Interpolation

A. Parameterization

To identify effective independent variables for geometric parameterization, the following arguments have been introduced. The basic idea is that lift, drag, and moment coefficients of a generic

airfoil should be governed by some variables describing its area (or volume, if one thinks of area per unit length) distribution in the chord direction, besides its angle of attack and the flow regime at which it is flown. Moreover, for nonsymmetrical airfoils, also the camber law is an important factor to be taken into account.

Concerning the angle of attack, in this work it was regarded as a parameter rather than an independent variable of the statistical model: in other words, a series of different STPS models were built, each of which refers to a peculiar airfoil attitude. In fact, in a typical design problem, an estimation of aerodynamic coefficients at a fixed, specific attitude is needed. Moreover, the resulting decrease of explanatory variables of the statistical model allows reducing the number of independent observations required to fit the model.

Two distinct databases, one including symmetrical and the other nonsymmetrical airfoils, were built.

Starting from the nondimensional airfoil coordinates, each profile was characterized with the corresponding nondimensional volume (area per unit length) distribution, using a trapezoidal approximation, as illustrated in Fig. 1. Given a series of points in the chordwise direction, the area of the preceding trapezium was attributed to each point: the whole airfoil was covered in the same way and finally the nondimensional area distribution as a function of the dimensionless chordwise coordinate was found. For nonsymmetrical airfoils, the camber was also determined and plotted as a function of the dimensionless chordwise coordinate. Being both the above-mentioned distributions highly dependent on the curve resolution, i.e., on the number and position of points chosen for their representation in the chordwise direction, a fixed distribution of those points was selected and kept constant for all the airfoils, both symmetric and nonsymmetric, in order for the obtained curves to be independent from the selected discretization criterion. Specifically, the chordwise discretization for the volume and camber curves was made up of 41 points with a cosine-type distribution.

Once all the airfoils were characterized with their own volume and camber distributions, it was necessary to select a parameterization approach in order for those distributions to be described with the minimum number of effective parameters.

Regarding the volume curve, the selected approach was to approximate it through a cubic spline, so a proper number and distribution of the spline control points was searched for. It was found that 8 knots in the chordwise direction, including the lower and upper limits, were sufficient to reconstruct the typical volume curve shape with the required accuracy: actually, the parameterization chosen was capable of accurately covering all the analyzed airfoil shapes, both symmetric and nonsymmetric. The selected control points' distribution is reported in Table 1 and illustrated graphically in Fig. 2. Since both the lower and upper bound coordinates in the y direction are always constant and equal to 0, the number of effective volume parameters is reduced to 6: they are the y coordinates of the central

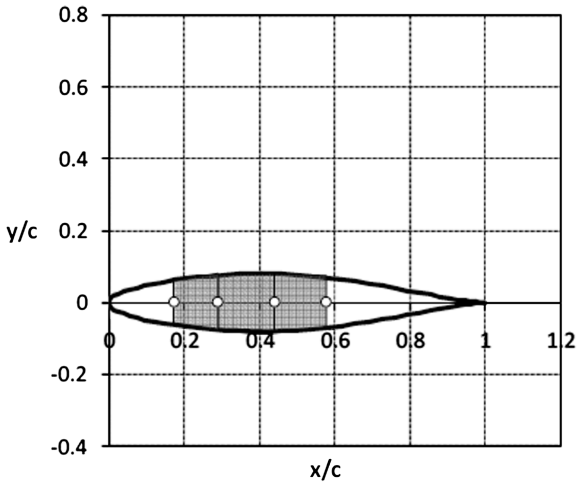

Fig. 1 Determination of a generic airfoil area distribution.

Table 1 2-D airfoils volume spline control points' distribution

Control point number	1	2	3	4
Chordwise coordinate	0.0	0.04	0.18	0.4
Control point number	5	6	7	8
Chordwise coordinate	0.57	0.8	0.95	1.0

spline knots, hereafter referred to as $y1_v$, $y2_v$, $y3_v$, $y4_v$, $y5_v$, and $y6_v$, respectively, as illustrated in Fig. 2.

On the other hand, as far as the camber law is concerned, it was decided not to describe the overall curve, with the aim of reducing the total number of independent variables in the STPS model (and hence the amount of observations required). Therefore, only the chordwise location of maximum camber and the corresponding maximum camber value were taken as the mean line parameters: they will be indicated in the following with $x_{\max}$camb and $y_{\max}$camb, respectively.

B. Calculation of Aerodynamic Performance

All the airfoils were simulated using the 2-D open source code XFOIL,[†] an interactive program for the design and analysis of subsonic isolated airfoils. The formulation of XFOIL combines the speed and accuracy typical of high-order panel methods with a fully coupled viscous/inviscid interaction. The inviscid formulation of XFOIL is a simple linear-vorticity stream function panel method where the equations are closed with an explicit Kutta condition. Moreover, a Karman–Tsien compressibility correction is incorporated. As it is well known, the theoretical foundation of the Karman–Tsien correction breaks down in supersonic flow [38]. As a result, the code accuracy rapidly degrades as the transonic regime is entered. As far as the viscous formulation is concerned, the boundary layers and wake are described with a two-equation lagged dissipation integral boundary layer formulation and an envelope e^n transition criterion. The entire viscous solution (boundary layers and wake) is strongly interacted with the incompressible potential flow via the surface transpiration model, thus allowing proper calculation of limited separation regions. The total velocity at each point on the airfoil surface and wake, with contributions from the freestream, the airfoil surface vorticity, and the equivalent viscous source distribution, is obtained from the panel solution with the Karman–Tsien correction added. This is incorporated into the viscous equations, yielding a nonlinear elliptic system which is readily solved by a full-Newton method. XFOIL was chosen for this application for its substantial easiness of use and because it offers a good compromise between execution rapidity and accuracy of results.

A total of 60 panels for the upper side and 60 for the lower side, both with a cosine-type distribution, were used for the discretization of each airfoil, as illustrated in Fig. 3. All the airfoils were simulated at six different flow regimes, with simultaneously varying Reynolds and Mach numbers, in order for both viscosity and compressibility effects to be captured and included into the statistical model. Variations of Reynolds and Mach numbers were achieved by acting on the undisturbed flow speed only: therefore, increasing Reynolds numbers corresponded to growing Mach numbers as well, according to a linear relationship between the two. It follows that the ratio between the Reynolds and Mach number was kept fixed throughout all the simulated flow regimes, so in the STPS model the independent variable accounting for the flow regime could be either the Reynolds or the Mach number indifferently, being the one left unspecified immediately derivable from the ratio of the two. The complete list of independent variables used to build up the statistical model is reported in Table 2. However, not all of the nine independent variables listed in Table 2 were included simultaneously in the STPS model, since this would have meant an unaffordable number of observations (i.e., of different airfoil geometries) to be introduced into the database. For both symmetric and nonsymmetric profiles, a

[†]Data available online at http://web.mit.edu/drela/Public/web/xfoil/ [accessed Nov. 2010].

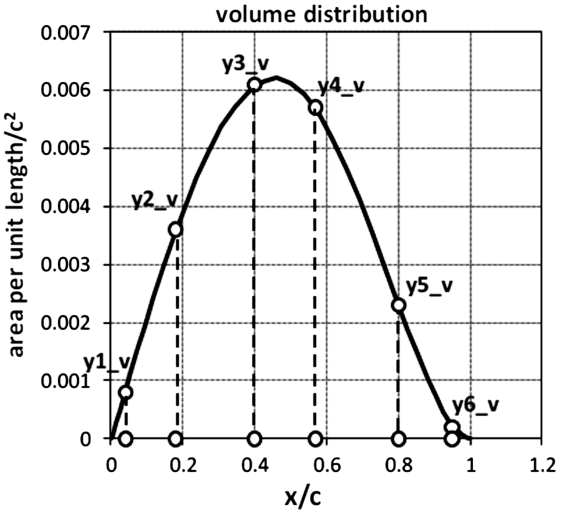

Fig. 2 Distribution of the control points over a typical volume curve.

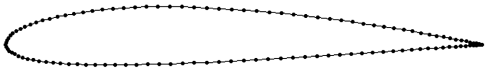

Fig. 3 Distribution of panels used for the XFOIL simulation on a generic airfoil.

reduction in the independent variables was carried out to bring back the necessary number of variables to five, which was shown with various in-depth tests to be sufficient for the polars of a generic airfoil to be predicted by the statistical model with satisfactory accuracy.

As a consequence, a total of 65 different symmetric airfoils and 65 nonsymmetric airfoils were simulated. In fact, various tests demonstrated that this was an adequate number of observations for a STPS statistical model with at least five independent variables to be reliable from the predictive point of view. Because of the six different flow regimes at which each airfoil was tested, a total of $6 * (65 + 65) = 780$ XFOIL runs were needed to fill in the symmetrical and nonsymmetrical airfoil database. Because of the Karman–Tsien compressibility correction applicability range, only Mach numbers up to 0.6 were simulated to keep the calculations far enough from the transonic regime. The Reynolds per unit length (based on the airfoil chord length) and Mach number conditions at which the airfoils were tested are summarized in Table 3.

Table 2 Independent variables for classification of airfoils

Independent variable id	Variable description
1	M (or Re)
2	$y1_v$
3	$y2_v$
4	$y3_v$
5	$y4_v$
6	$y5_v$
7	$y6_v$
8 (for nonsymmetric airfoils)	$x_{\max}$camb
9 (for nonsymmetric airfoils)	$y_{\max}$camb

Table 3 Flow regime conditions used in the XFOIL simulations

Flow regime condition number	Mach number	Reynolds number per unit length
1	0.1	1,380,000
2	0.2	2,760,000
3	0.3	4,140,000
4	0.4	5,520,000
5	0.5	6,900,000
6	0.6	8,280,000

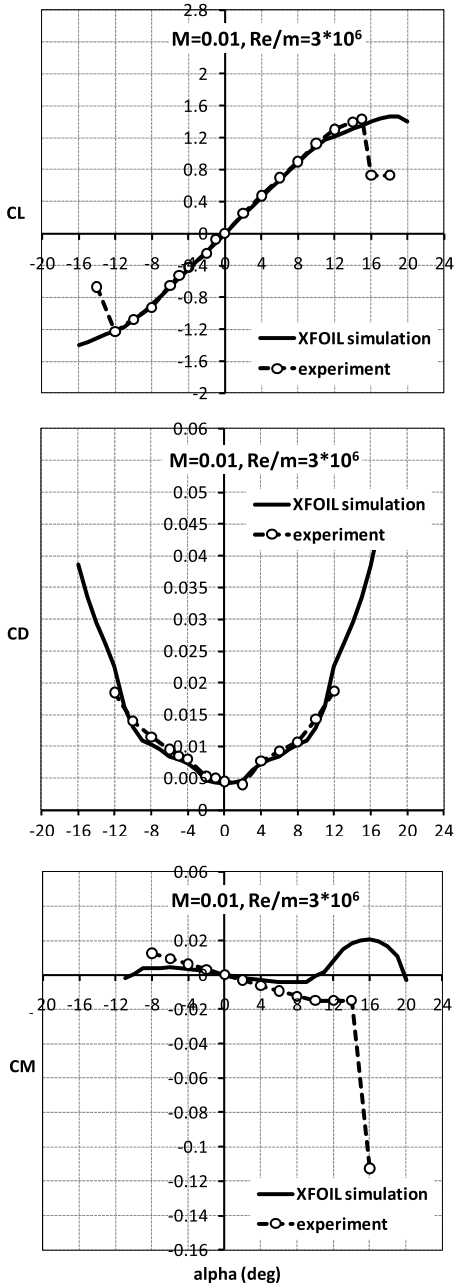

Fig. 4 Validation results of XFOIL on the symmetrical airfoil NACA64-012: comparison with experimental data at incompressible flow regime and at Reynolds number per unit length $3 * 10^6$.

Airfoil polars were calculated for a maximum range of angles of attack between $-16°$ and $+16°$: in fact, reliable results can be obtained from XFOIL simulations only up to the stall onset. Obviously, for airfoils stalling at a smaller angle of attack, simulations did not reach convergence and therefore the polars' extension was reduced. A whole of 500 iterations were prescribed for each point of the polar to reach solution convergence and transition location was left free over both the pressure and suction side of the profile.

With the simulations on all the airfoils completed, a number of different look-up tables (distinct for symmetric and nonsymmetric profiles), one for each value of the angle of attack, were filled in, reporting simultaneously the values of the independent variables (input variables of the statistical model) and the lift, drag, and moment coefficients values at that attitude (responses of the statistical model) for all the analyzed airfoils.

C. Validation of the Flow Solver

A validation analysis of XFOIL was carried out: specifically, some comparisons were performed of simulation results using XFOIL and experimental data published in the literature [39] over a couple of airfoils, one symmetrical and one nonsymmetrical, each of which was tested at an incompressible flow regime and at two different Reynolds number values.

The two profiles, namely the NACA64-012 and the NACA2412, were discretized adopting the same criterion used for the databases' construction, with a whole of 60 panels for the upper side and 60 for the lower side, both with a cosine-type distribution. Moreover, a whole of 500 iterations were prescribed for each point of the polar to reach solution convergence and transition location was left free over both the pressure and suction side of the airfoils.

The obtained validation results are illustrated in Figs. 4–7. Regarding the lift curves, the agreement between XFOIL and the wind tunnel data at both the considered flow conditions is very satisfactory before stall onset for both symmetrical and nonsymmetrical airfoils, with the slope of the linear portion of the curves and the angle of zero lift being excellently reproduced. On the other hand, the stall angle predicted by XFOIL is higher than the experimental one. Actually, XFOIL stall model is quite simplified and fails in case of massive flow separations: therefore, some mismatch may arise with experimentally acquired values on the simulated airfoils. A very good match is also obtained when comparing the simulated and experimental drag coefficients, with XFOIL predictions being slightly underestimated only for the nonsymmetrical airfoil all along the described polar and for both the analyzed flow regimes. Finally, the pitching moment polars are satisfactory captured as a trend, however, featuring some zones where discrepancy becomes significant for both airfoils at each of the considered Reynolds number values. In particular, some major disagreements with experiments are evidenced in the poststall region.

On the whole, it may be concluded that, despite some intrinsic limitations, XFOIL is a sufficiently reliable tool for construction of databases to be used for the statistical model determination.

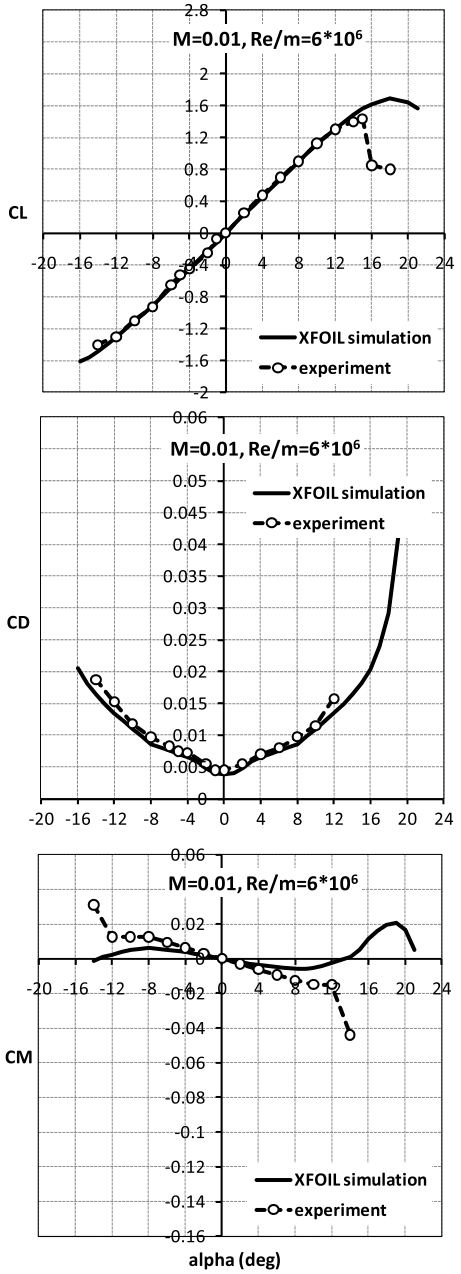

Fig. 5 Validation results of XFOIL on the symmetrical airfoil NACA64-012: comparison with experimental data at incompressible flow regime and at Reynolds number per unit length $6 * 10^6$.

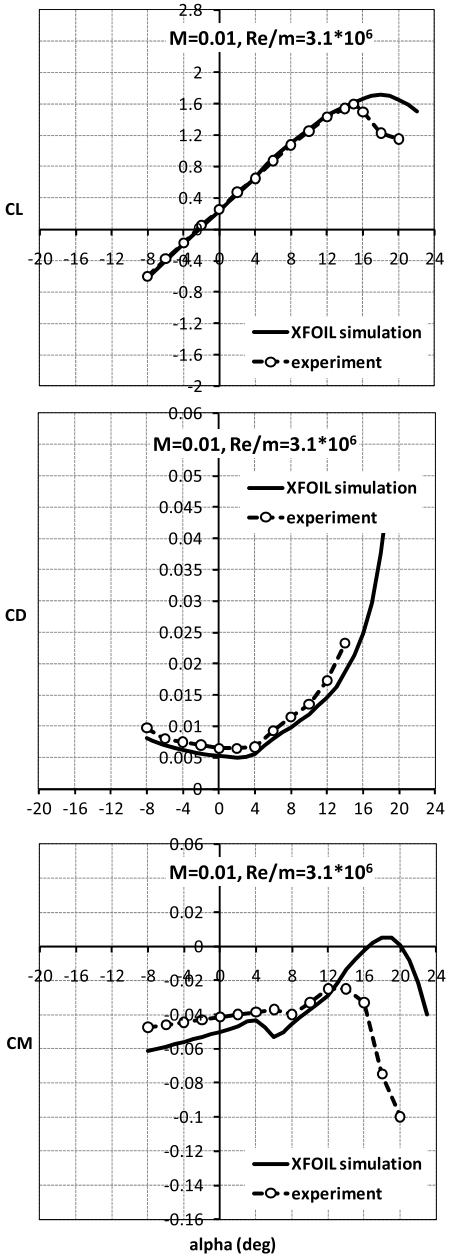

Fig. 6 Validation results of XFOIL on the nonsymmetrical airfoil NACA2412: comparison with experimental data at incompressible flow regime and at Reynolds number per unit length $3.1 * 10^6$.

VI. Discussion of Results

The STPS models built for each value of the angle of attack were used to predict the lift, drag, and moment polars of airfoils not included into the database. The airfoils selected for validation were characterized with their independent variables values, which were then used, at prescribed flow regime and attitude, as new input points where the response of the statistical model was queried. Finally, the obtained polars were compared with those coming from XFOIL simulations over the same profile.

In the following, obtained results are presented and discussed, for both symmetrical and nonsymmetrical airfoils.

A. Symmetrical Airfoils

Validation of STPS model was carried out on three different symmetrical airfoils with low, medium and moderately high percentage thickness, respectively, namely the NACA64-009, the NACA63A-015 and the NACA64-020.

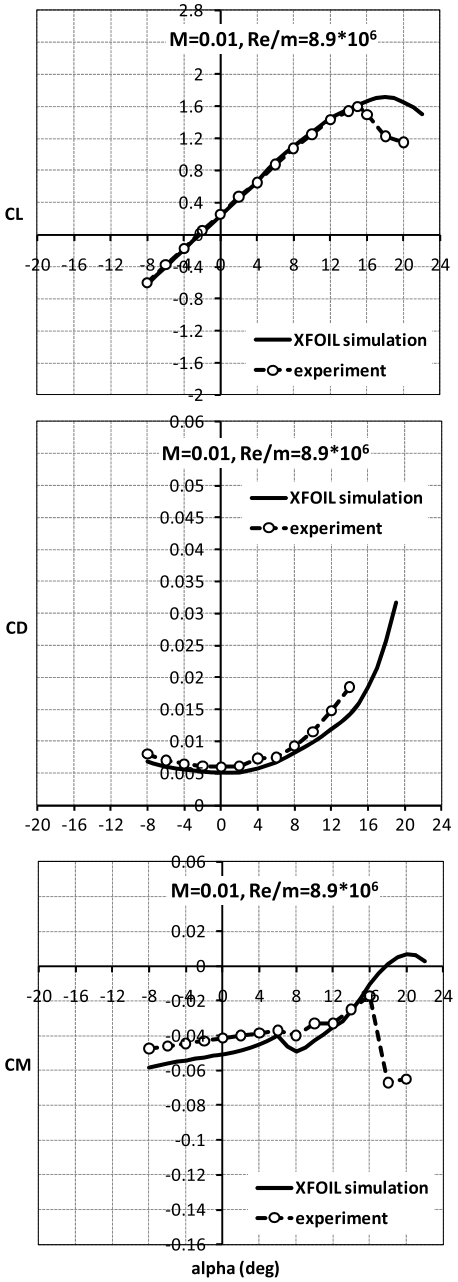

Fig. 7 Validation results of XFOIL on the nonsymmetrical airfoil NACA2412: comparison with experimental data at incompressible flow regime and at Reynolds number per unit length $8.9 * 10^6$.

Because of the symmetric aerodynamic behavior of symmetrical airfoils, the validation work was limited to positive angles of attack only, ranging from 0 to 16°. However, as already mentioned, for airfoils stalling at a smaller angle of attack, simulations did not reach convergence and therefore the polars' extension was reduced.

On the whole, $17 * 3$ different STPS models were built, one for each of the aerodynamic coefficients and at each value of the angle of attack; then, each of the models was fed with the independent variables of the three selected airfoils and the corresponding response was extracted.

A statistical significance study was carried out to identify the most suitable combination of input variables to be introduced into the statistical model to get a high correlation with the simulated polars. Specifically, the Student parameter for each geometric variable was calculated for lift, drag and moment coefficient distributions over the whole range of examined angles of attack, and a choice was made according to its values. The results of the Student tests, that are reported in Fig. 8, highlighted that the geometric parameters $y2_v$, $y3_v$, $y4_v$, and $y5_v$ have the highest statistical significance and therefore they are likely to allow the best correlation on aerodynamic coefficients over the analyzed airfoils. On the other hand, the Mach (or Reynolds) number was not tested statistically, but it was rather included directly into the independent variable set, because it defines the flow regime condition and as such it has no correlation with the geometrical variables.

As an example, some diagnostic plots of the statistical models for CL, CD, and CM obtained for symmetrical airfoils at an angle of attack equal to 3° as a function of the above-mentioned input

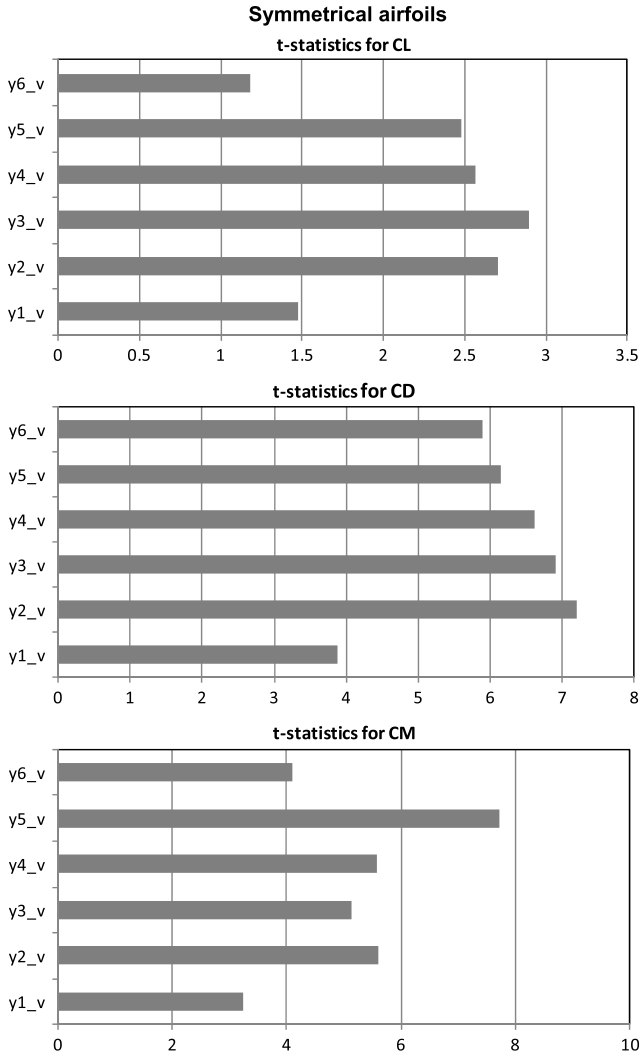

Fig. 8 *t*-statistics of the geometrical parameters for symmetrical airfoils.

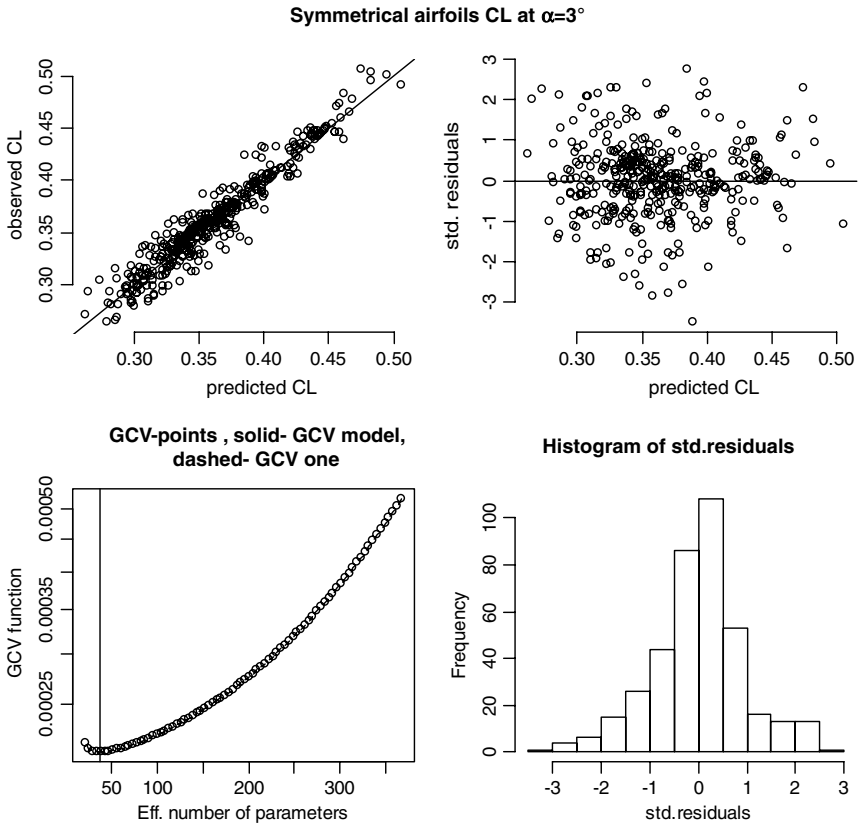

Fig. 9 Diagnostic plots of the statistical model for CL of symmetrical airfoils at $\alpha = 3°$.

variables are reported in Figs. 9–11, respectively. The correlation coefficients for lift, drag and moment coefficients are

$$R^2_{\mathrm{CL}} = 0.923, \qquad R^2_{\mathrm{CD}} = 0.973, \qquad R^2_{\mathrm{CM}} = 0.837$$

and they represent a measure of how close the point cloud is to the first quadrant bisector in the plots at the top left of Figs. 9–11, depicting the values predicted by the statistical model vs the simulated ones using XFOIL. Correlation on CM is the lowest of the

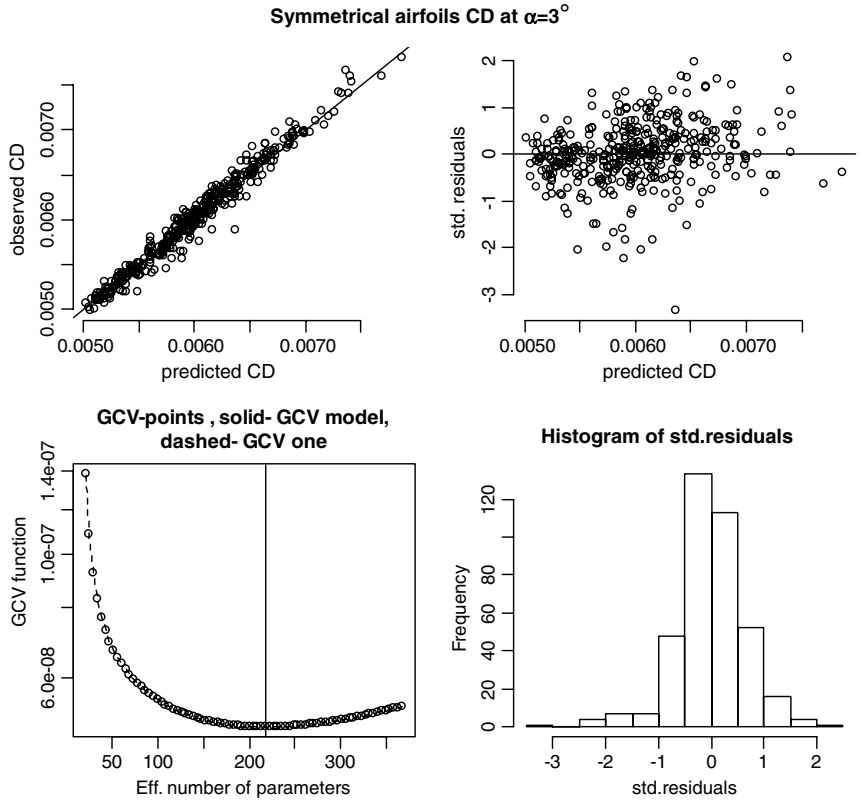

Fig. 10 Diagnostic plots of the statistical model for CD of symmetrical airfoils at $\alpha = 3°$.

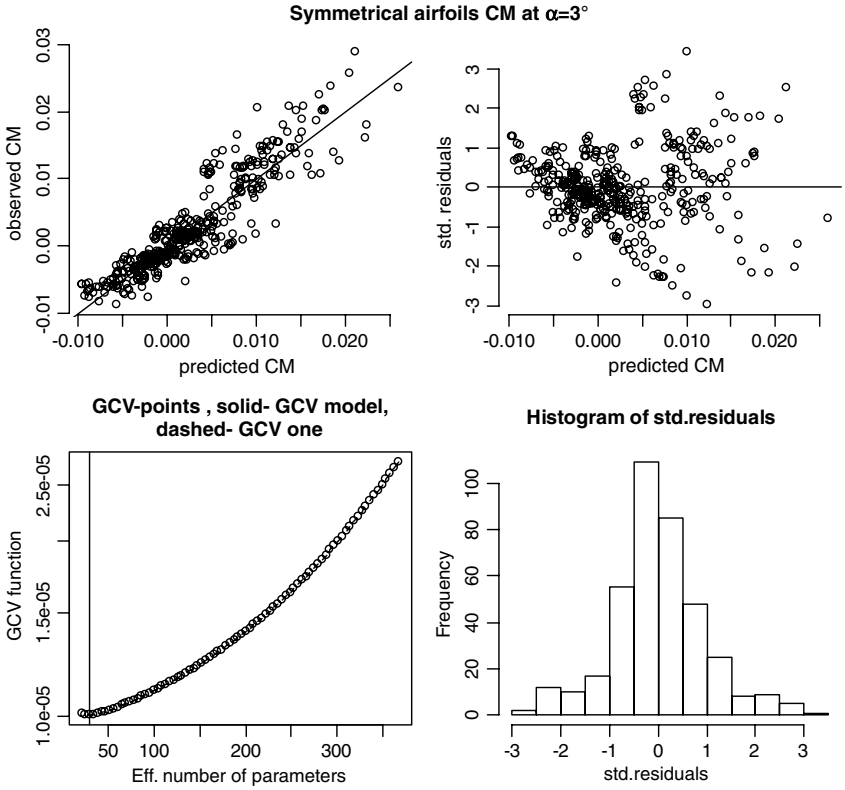

Fig. 11 Diagnostic plots of the statistical model for CM of symmetrical airfoils at $\alpha = 3°$.

three, due to fact that the moment coefficients of the database are much sparser than the lift and drag ones. A normal distribution of residuals around the null value is evidenced for each of the aerodynamic coefficients by both the histograms at the bottom right of Figs. 9–11 and the plots at the top right of the same figures, representing the residual distribution as a function of the predicted values. Finally, the plots at the bottom right of Figs. 9–11 illustrate the GCV(α) score behavior as a function of the equivalent degrees of freedom of the model, which is directly correlated to the adopted value of α ([9]). Actually, the minimum of GCV(α) identifies the effective value of the smoothing parameters used in the statistical model.

For the sake of brevity, only validation results on NACA63A-015 at two different flow regimes, one at low and the other at moderately high Mach and Reynolds numbers, are reported. The geometry and volume distribution curve of this airfoil are depicted in Fig. 12, while the values of the pertinent geometric explanatory variables are reported in Table 4. Validation results are illustrated in Figs. 13 and 14, where the CL, CD, and CM polars predicted with STPS statistical models are compared with those simulated using XFOIL, for the flow conditions referred to as numbers 1 and 5, respectively, in Table 3. The dashed lines on the same figures represent the range of variability of the aerodynamic coefficients of all the airfoils included in the database, for each analyzed Mach number and Reynolds (per unit chord length) number conditions, so they, in a sense, delimit the search space over which the statistical model operates to give its responses.

Validation results show an excellent correlation of the statistically predicted values with XFOIL simulations for all the analyzed airfoils.

Specifically, the lift coefficient curves are very well reconstructed, with the curve slope being accurately captured; even the stall onset is reproduced with a remarkable accuracy, while some major discrepancies are evidenced in the poststall portion of the curves. To this purpose, it is worth noting that the reliability of the stall prediction compared with experimental values goes beyond the scope of this work and therefore is not under discussion here; actually, we are only interested in the statistical model ability to predict stall based on the data included into the database. Those data were obtained using XFOIL, in which the stall model fails in case of massive flow

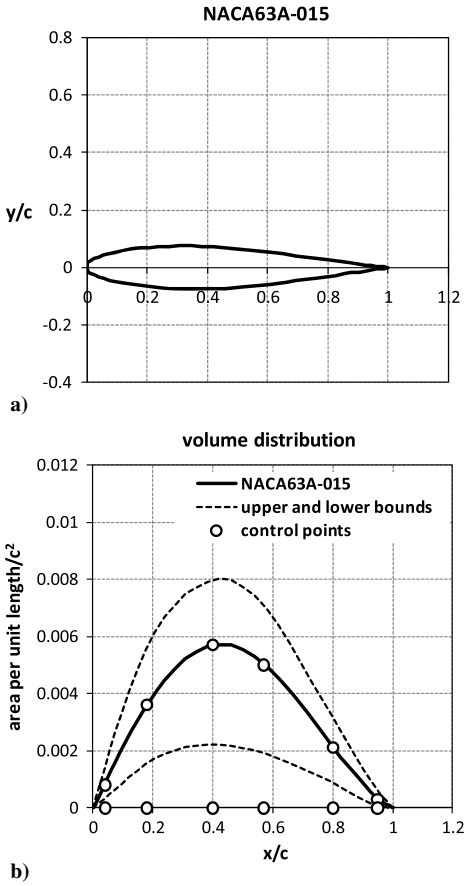

Fig. 12 NACA63A-015: a) geometry of the airfoil, and b) volume distribution curve.

Table 4 NACA63A-015: values of the explanatory geometric variables

NACA63A-015						
Variable name	$y1_v$	$y2_v$	$y3_v$	$y4_v$	$y5_v$	$y6_v$
Variable value	0.0008	0.0036	0.0057	0.005	0.0021	0.0003

separations, as already mentioned, and therefore some mismatch may arise with experimentally acquired values on the simulated airfoils. Despite this, the important thing here is that the STPS model is capable of predicting the stall onset in a coherent way with the available data.

The drag curves are very well reconstructed too, with both the minimum drag values and the curve envelope at higher angles of attack being satisfactory captured. Some discrepancy with the XFOIL values is shown only for the NACA64-009 at the highest angles of attack and for increasing Mach and Reynolds numbers.

Finally, the moment coefficient curves show an outstanding correlation with simulated values, and this is much more surprising, considering that moment curves included into the database are much sparser than lift and drag ones.

Moreover, it is important to note that variations on the curves of all coefficients with the flow regime are very well captured: specifically, the increasing lift curve slope and the growing slope of the drag curve (at the highest angles of attack) with growing Reynolds and Mach numbers are excellently reproduced. The prediction accuracy tends to slightly deteriorate at the flow regime with the highest value of Mach and Reynolds numbers, due to the fact that much less observations are available in the database at those conditions: in fact, many analyzed airfoils (particularly thin ones) failed to reach convergence with XFOIL simulations at increasing Reynolds and Mach numbers and hence they had to be removed from the database. As a consequence, at these flow regimes the set of acquired data is statistically less significant for the STPS model to be constructed with a high reliability.

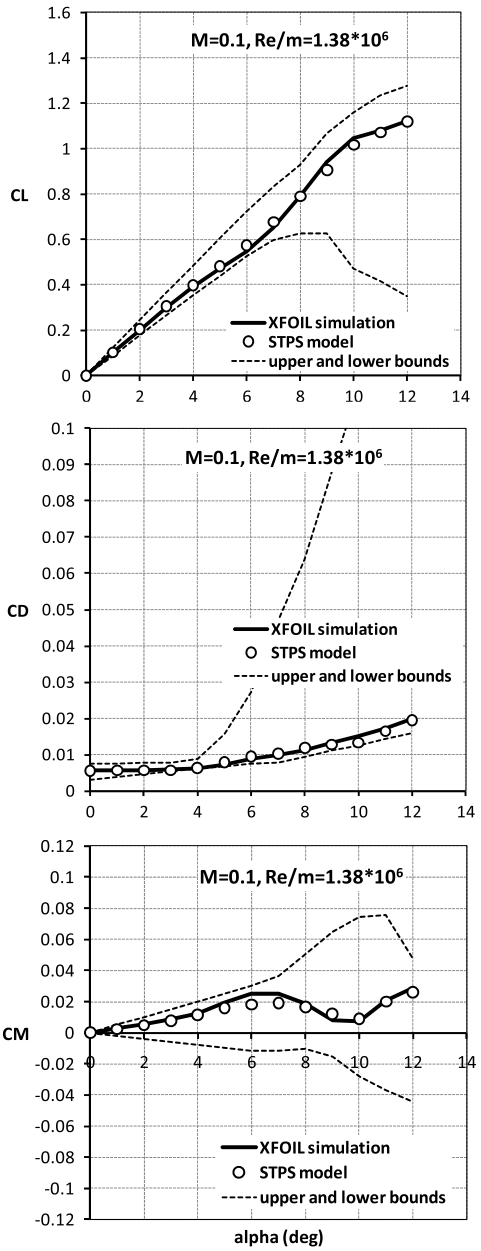

Fig. 13 Validation results on NACA63A-015 airfoil: CL, CD, and CM polars at low Mach and Reynolds (per unit chord length) number conditions; STPS model compared with XFOIL simulations.

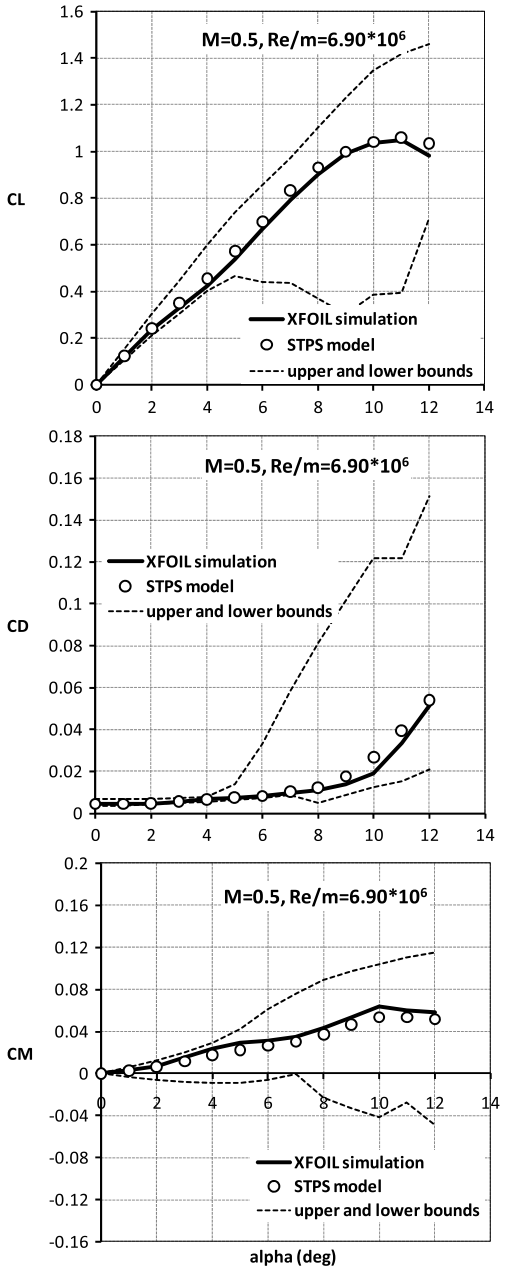

Fig. 14 Validation results on NACA63A-015 airfoil: CL, CD, and CM polars at moderately high Mach and Reynolds (per unit chord length) number conditions; STPS model compared with XFOIL simulations.

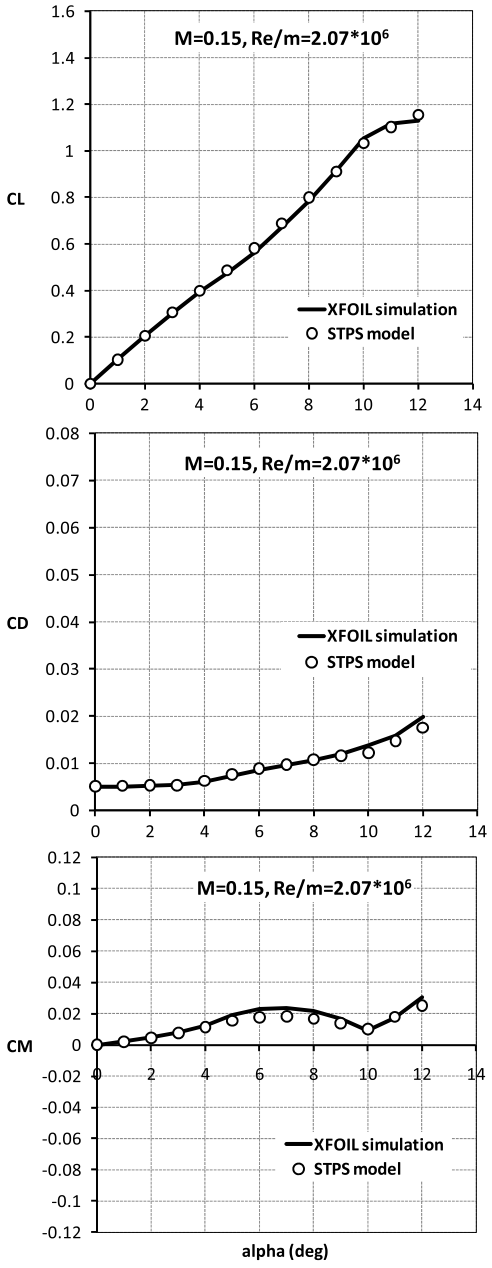

Fig. 15 Validation results on NACA63A-015 airfoil: CL, CD, and CM polars at a low Mach and Reynolds (per unit chord length) number conditions not included into the database; STPS model compared with XFOIL simulations.

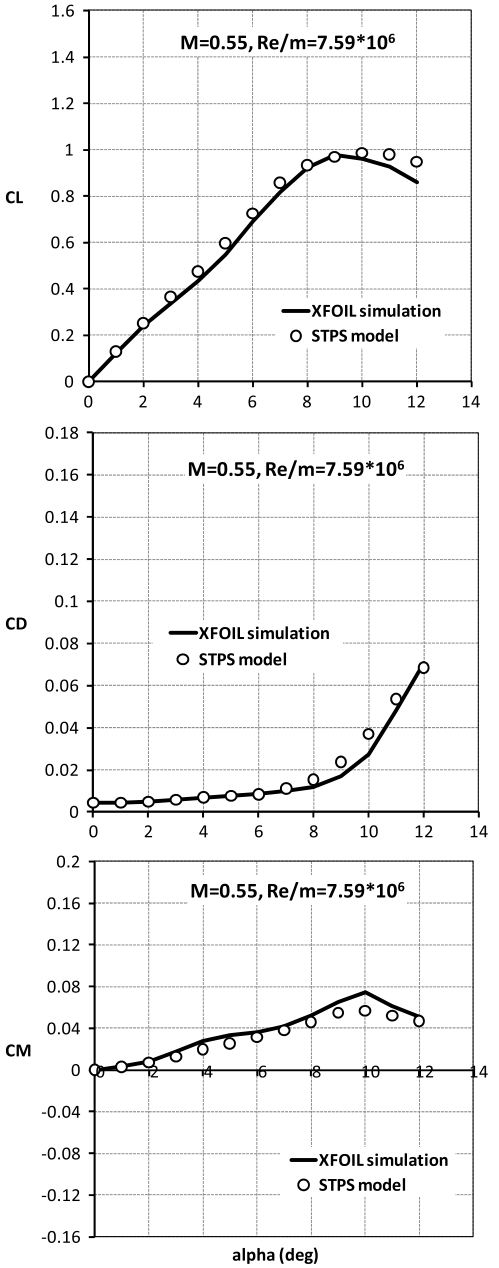

Fig. 16 Validation results on NACA63A-015 airfoil: CL, CD, and CM polars at a moderately high Mach and Reynolds (per unit chord length) number conditions not included into the database; STPS model compared with XFOIL simulations.

Furthermore, the capability of the model to predict aerodynamic performance at flow regimes not directly included into the database, although encompassed by the analyzed conditions, was verified. To this purpose, three intermediate Mach and Reynolds numbers conditions were tested over the NACA63A-015 airfoil: the obtained results are illustrated in Figs. 15 and 16, and once again the correlation with simulations is excellent for all the aerodynamic coefficients.

In Table 5 the mean square error (MSE) of the model fitting for lift, drag, and moment coefficients at each of the examined flow regimes is reported and compared with the maximum values the coefficients exhibit over the analyzed incidence range according to XFOIL

Table 5 MSE of the STPS model on aerodynamic coefficients NACA63A-015

NACA63A-015					
M		0.1	0.15	0.5	0.55
CL	MSE	0.01807	0.01714	0.03061	0.04519
	Max value	1.123	1.1312	1.0497	0.9792
CD	MSE	1.157E − 5	1.455E − 5	9.62E − 5	1.435E − 4
	Max value	0.01982	0.01977	.0513	0.07052
CM	MSE	8.554E − 5	9.542E − 5	3.936E − 4	5.585E − 4
	Max value	0.0291	0.0305	0.0639	0.0747

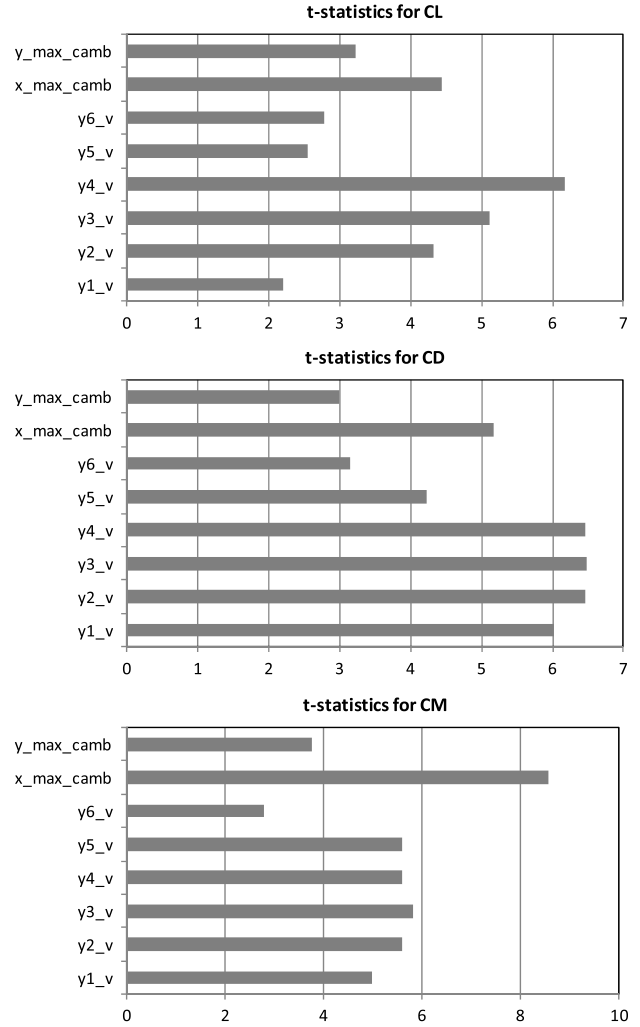

Fig. 17 ***t*-statistics of the geometrical parameters for nonsymmetrical airfoils.**

simulations: the MSE values are at least 2 orders of magnitude less than the maximum value for all the tested coefficients and flight regimes.

The validation results clearly demonstrated that the choice of describing the airfoil geometry by means of the chordwise volume (or area per unit length) distribution is suitable for the purpose of constructing an accurate nonparametric statistical model for polars' determination. Actually, the selected geometric variables for airfoil description were proven to be appropriate for a complete characterization of the airfoil behavior at different angles of attack and flow regimes. This was not so clear since the beginning of the work, due to this choice being rather unusual for aerodynamic applications.

B. Nonsymmetrical Airfoils

Analogously to symmetrical airfoils, validation of STPS model on nonsymmetrical profiles was carried out on three different airfoils with low, medium and moderately high percentage thickness, respectively, namely the NACA23009, the NACA3415, and the NACA23120.

The range of angles of attack selected for validation was wider than for symmetrical airfoils, with values of α spreading from -7 to $16°$. Therefore, $24 * 3$ different STPS models were built, one for each aerodynamic coefficient and at each value of the above-mentioned angles of attack; then, the independent variables of the three selected airfoils were used as input values for each of the statistical models, and the corresponding response was extracted.

Also in this case, a statistical significance study was carried out to identify the most suitable combination of input variables to be introduced into the statistical model to get a high correlation with the simulated polars. The Student parameter for each geometric variable was calculated for lift, drag and moment coefficient distributions over the whole range of examined angles of attack, and results are reported in Fig. 17. As can be inferred from the figure, the geometric parameters $y2_v$, $y3_v$, $y4_v$, and $x_{\max}$camb were shown to have the largest significance and hence they were selected as independent variables of the statistical model. Once again, to define the flow conditions the airfoils were flown, the Mach (or Reynolds) number was included directly into the independent variable set without testing its significance.

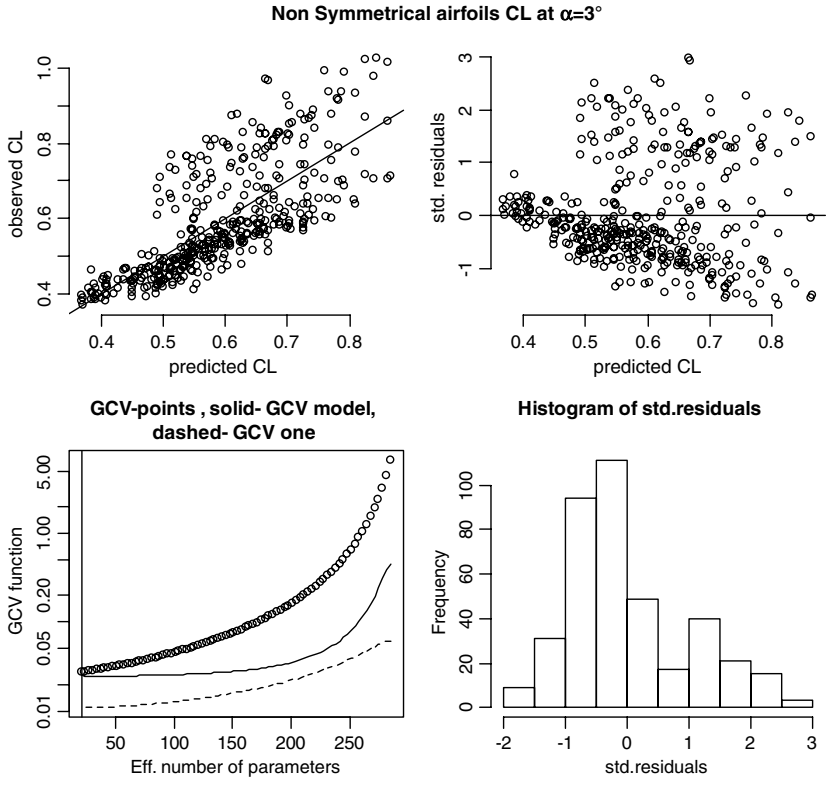

Fig. 18 Diagnostic plots of the statistical model for CL of nonsymmetrical airfoils at $\alpha = 3°$.

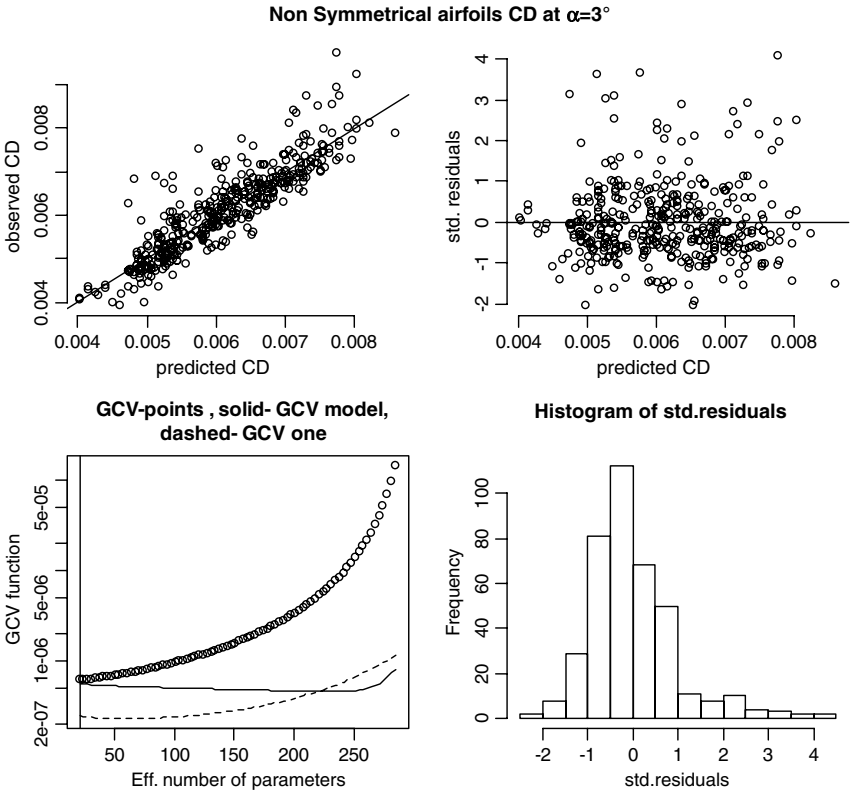

Fig. 19 Diagnostic plots of the statistical model for CD of nonsymmetrical airfoils at $\alpha = 3°$.

As for symmetrical airfoils, some diagnostic plots of the statistical models for CL, CD, and CM obtained at an angle of attack equal to 3° as a function of the above-mentioned input variables are illustrated in Figs. 18–20, respectively. The correlation coefficients are

$$R^2_{\mathrm{CL}} = 0.551, \qquad R^2_{\mathrm{CD}} = 0.789, \qquad R^2_{\mathrm{CM}} = 0.655$$

as usual, they represent a measure of the point cloud proximity to the first quadrant bisector in the plots at the top left of Figs. 18–20, where the values predicted by the statistical model are depicted vs the observed ones coming from XFOIL simulations. The correlation coefficients' values are worse than for symmetrical airfoils, especially for CL, while the CM cloud features a somewhat

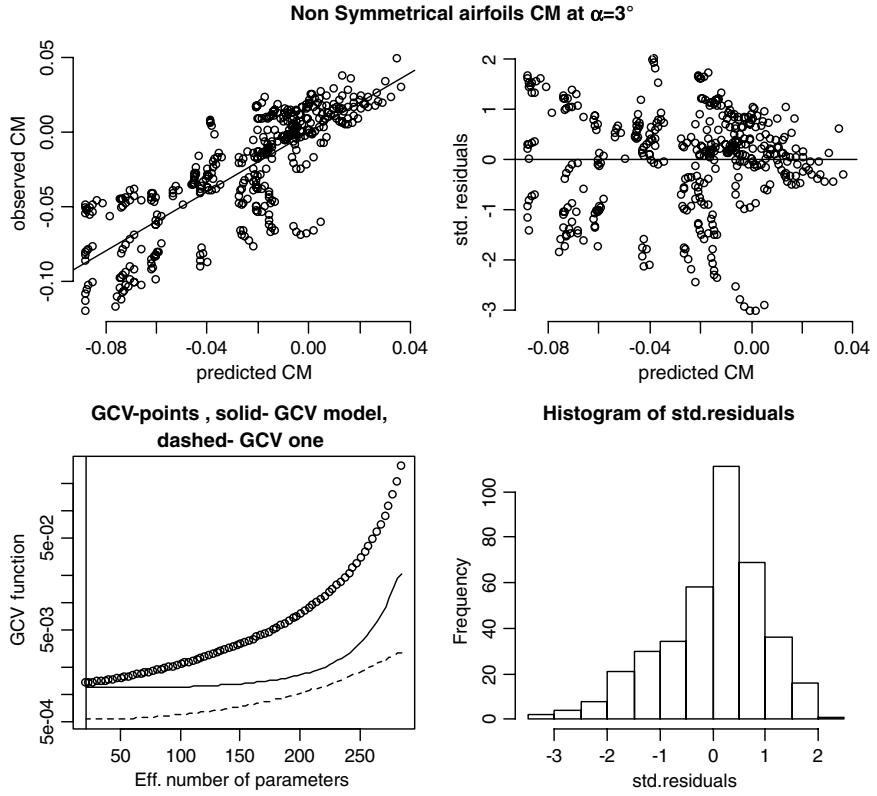

Fig. 20 Diagnostic plots of the statistical model for CM of nonsymmetrical airfoils at $\alpha = 3°$.

clusterized behavior, with the various groups of points corresponding to the different airfoil families introduced into the database. Despite this, the prediction of aerodynamic performance of all the airfoils selected for validation is very satisfactory, as will be illustrated later on. In fact, unlike parametric techniques, the smoothing thin plate splines approach does not search for the statistical model with the highest possible correlation coefficient: the reconstructed model comes rather from a compromise between correlation and smoothness of the response hypersurface, to avoid data overfitting. A normal distribution of residuals around the null value is evidenced for each of the aerodynamic coefficients by both the histograms at the bottom right of Figs. 18–20 and the plots at the top right of the same figures, representing the residual distribution as a function of the predicted values. Once again, the plots at the bottom right of Figs. 18–20 illustrate the GCV(α) score behavior as a function of the equivalent degrees of freedom of the model, thus identifying the specific smoothing parameter value to be used in the response surface.

Only results of validation on NACA23120 at two different flow regimes, one at low and the other at moderately high Mach and

Table 6 NACA23120: values of the explanatory geometric variables

NACA23120				
Variable name	$y1_v$	$y2_v$	$y3_v$	$y4_v$
Variable value	0.0014	0.0053	0.0074	0.0064
Variable name	$y5_v$	$y6_v$	x_{max}camb	y_{max}camb
Variable value	0.0029	0.0005	0.1198	0.0207

Table 7 MSE of the STPS model on aerodynamic coefficients NACA23120

NACA23120			
M		0.1	0.5
CL	MSE	0.06032	0.03917
	Max value	1.4764	1.5712
CD	MSE	5.93E − 5	8.47E − 5
	Max value	0.03073	.07614
CM	MSE	4.65E − 4	4.45E − 4
	Max value	0.0514	0.0839

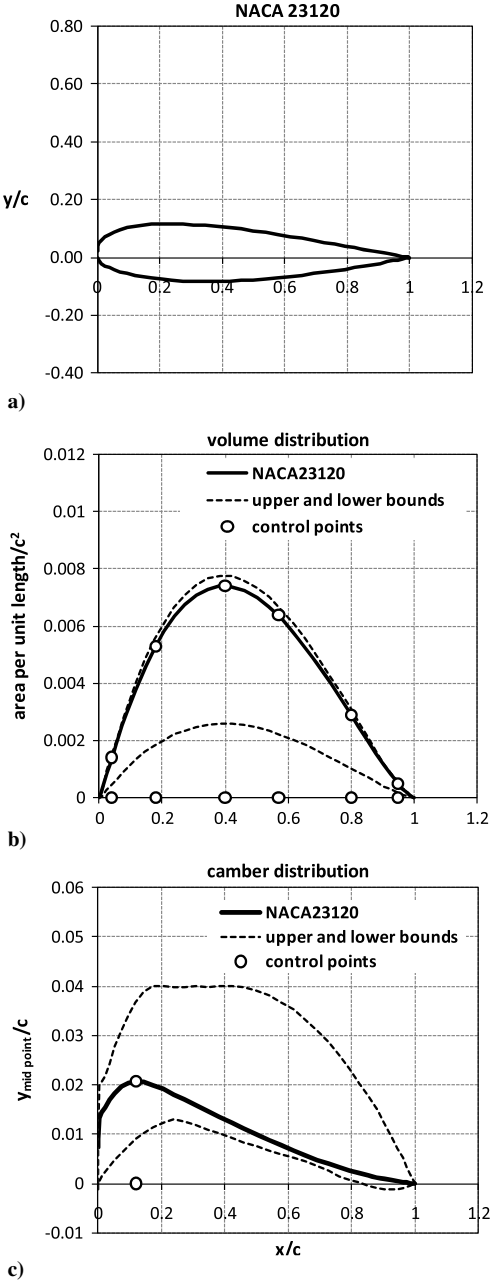

Fig. 21 NACA23120: a) geometry of the airfoil, b) volume distribution, and c) camber law.

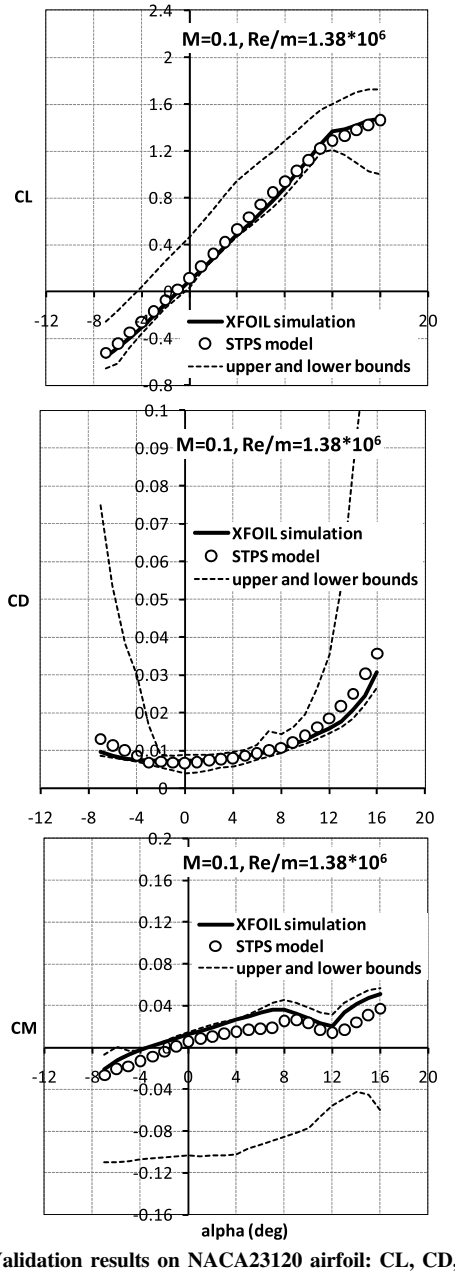

Fig. 22 Validation results on NACA23120 airfoil: CL, CD, and CM polars at low Mach and Reynolds (per unit chord length) number conditions; STPS model compared with XFOIL simulations.

Reynolds numbers, are reported. The geometry, volume distribution and camber curves of the airfoil are illustrated in Fig. 21, while the values of the pertinent geometric explanatory variables are reported in Table 6. Validation results are depicted in Figs. 22 and 23, where the response of the statistical model is compared with the lift, drag, and moment polars simulated using XFOIL, for the flow conditions referred to as numbers 1 and 5, respectively, in Table 3. Once again, the dashed lines on the same figures represent the range of variability of the aerodynamic coefficients of all the airfoils included in the database, for each analyzed Mach number and Reynolds (per unit chord length) number conditions.

Also for nonsymmetrical airfoils, an outstanding correlation of the statistically predicted values with XFOIL simulations was obtained for the selected profiles.

In particular, the lift coefficient curves are very well reproduced, with both the curve slope and the angle of zero lift being accurately captured: this means that the STPS model is capable, through the selected independent variables, to correctly identify the peculiar family to which the airfoil under analysis belongs among those included into the database. Also in this case, the stall onset is reproduced with outstanding accuracy for all the selected airfoils, and even the poststall portion of the curves is very well correlated with XFOIL simulations throughout all the tested flow regimes.

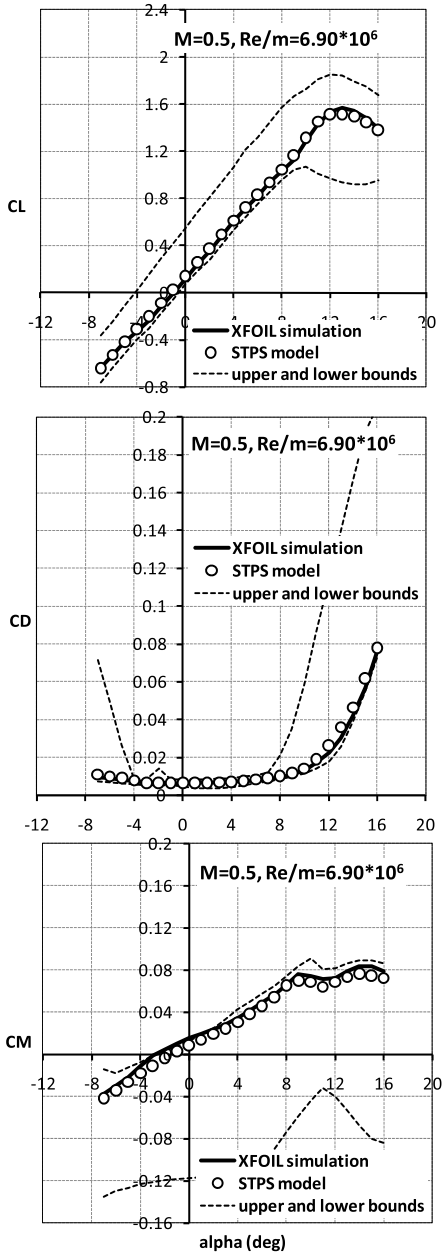

Fig. 23 Validation results on NACA23120 airfoil: CL, CD, and CM polars at moderately high Mach and Reynolds (per unit chord length) number conditions; STPS model compared with XFOIL simulations.

Also the drag curves are very well reconstructed, with both the minimum drag coefficient (and the corresponding angle of attack of minimum drag) and the curves' divergence at higher angles of attack being captured with outstanding accuracy.

Finally, the moment coefficient curves show an excellent correlation with simulated values, and once again this is much more surprising, considering that, as for symmetrical airfoils, moment curves included into the database are much sparser than lift and drag ones. Hence, this comes as a further confirmation of the validity of the statistical model.

Once again, the effects of Reynolds and Mach numbers variations on the curves of all aerodynamic coefficients are very well captured; specifically, the increasing lift curve slope, the anticipated stall onset and the growing slope of the divergent portion of the drag curves (at the highest angles of attack) with growing Reynolds and Mach numbers are excellently reproduced. Unlike the symmetrical airfoil case, the prediction accuracy does not seem to deteriorate at the highest values of Mach and Reynolds numbers, despite less observations are available in the database at those conditions, due to many of the analyzed airfoils (particularly thin ones) to fail convergence with XFOIL simulations at increasing Reynolds and Mach numbers.

Again, the MSE of the model fitting for lift, drag and moment coefficients at each of the examined flow regimes is reported in Table 7 and compared with the maximum values of the coefficients coming from XFOIL simulations over the analyzed incidence range: similarly to the symmetric airfoil case, the MSE values are at least 2 orders of magnitude less than the corresponding maximum value for all the tested coefficients and flight regimes.

Also in this case, the effectiveness of the selected explanatory variable set for the statistical model was clearly demonstrated, allowing for a complete characterization of the airfoils' behavior at different angles of attack and flow regimes.

VII. Conclusions

A method for multivariate fitting of both symmetric and nonsymmetric airfoils polars was successfully implemented, and is able to account for simultaneous influence of five independent geometric variables (chordwise distribution of area) on three responses (lift, drag, and pitching moment coefficients). The method was based on the implementation of multivariate smoothing thin plate splines, which were built on a dataset containing airfoil performance, as calculated from a validated 2-D flow solver. Results can be summarized as follows:

1) Validation clearly demonstrated that the choice of describing the airfoil geometry by means of the chordwise volume (or area per unit length) distribution is suitable for the purpose of constructing an accurate nonparametric statistical model for polars' determination.

2) Symmetric airfoils: the lift coefficient curves, their slope and stall onset angles are very well reconstructed, while some major discrepancies are evidenced in the poststall portion of the curve (it is worth recalling, however, that the scope of the work was not to estimate the poststall behavior, a fact which depends merely on accuracy of available data). The drag curves are very well reconstructed too, with both the minimum drag values and the curve envelope at higher angles of attack being satisfactorily captured. Some discrepancy with the XFOIL values is shown only for the NACA64-009 at the highest angles of attack and for increasing Mach and Reynolds numbers. Moment coefficient curves show an outstanding correlation.

3) Nonsymmetric airfoils: also in this case, prediction of aerodynamic performance of all the airfoils selected for validation is very satisfactory. Lift coefficient curves are very well reproduced, with both the curve slope and the angle of zero lift being accurately captured; stall onset is reproduced with remarkable accuracy for all the selected airfoils, and even the poststall portion of the curves is very well correlated with XFOIL simulations throughout all the

tested flow regimes. Also the drag curves are very well reconstructed, along with both the minimum drag coefficient and the curves' divergence at higher angles of attack. Finally, moment coefficient curves show an excellent correlation with simulated values. Effects of Reynolds and Mach number variations on the curves of all aerodynamic coefficients are very well captured; specifically, the increasing lift curve slope, the anticipated stall onset, and the growing slope of the divergent portion of the drag curves (at the highest angles of attack) with growing Reynolds and Mach numbers are excellently reproduced. Unlike the symmetrical airfoil case, the prediction accuracy does not seem to deteriorate at the highest values of Mach and Reynolds numbers.

Acknowledgment

The authors gratefully acknowledge Antonio Saporiti of AgustaWestland for the invaluable technical support and deep inspiration he has given to this work.

References

[1] Hill, T., and Lewicki, P., "Statistics: Methods and Applications," *StatSoft*, StatSoft, Tulsa, OK, 2007.

[2] Myers, R. H., and Montgomery, D. C., *Response Surface Methodology*, 2nd ed., Wiley-Interscience, New York, 2002.

[3] Box, G. E., and Draper, N. R., *Empirical Model-Building and Response Surfaces*, Wiley, New York, 1987.

[4] Landman, D., Simpson, J., Vicroy, D., and Parker, P., "Efficient Methods for Complex Aircraft Configuration Aerodynamic Characterization Using Response Surface Methodologies," AIAA 2006-922, Jan. 2006.

[5] Rallabhandi, S. K., Cagaty, E., and Mavris, D. N., "An Improved Procedure for Prediction of Drag Polars of a Joined Wing Concept Using Physics-Based Response Surface Methodology," SAE Paper 2001-01-3015, 2001.

[6] Mavris, D. N., and Qiu, S., "An Improved Process for the Generation of Drag Polars for use in Conceptual/Preliminary design," AIAA Paper 1999-01-5641, 1999.

[7] Eubank, R. L., *Nonparametric Regression and Spline Smoothing*, Marcel Dekker, Inc., New York, 1999.

[8] Bowman, A. W., and Azzalini, A., *Applied Smoothing Techniques for Data Analysis*, Clarendon Press, Oxford, 1997.

[9] Green, P. J., and Silverman, B. W., *Nonparametric Regression and Generalized Linear Models*, Chapmal & Hall, London, 1994.

[10] Friedman, J. H., "Multivariate Adaptive Regression Splines," *Annals of statistics*, Vol. 19, No. 1, 1991, pp. 1–67. doi:10.1214/aos/1176347963

[11] Klein, V., "Estimation of Aircraft Aerodynamic Parameters from Flight Data," *Progress in Aerospace Sciences*, Vol. 26, No. 1, 1989, pp. 1–77. doi:10.1016/0376-0421(89)90002-X

[12] Raisinghani, S. C., Ghosh, A. K., and Kalra, P. K., "Two New Techniques for Parameter Estimation Using Neural Networks," *The Aeronautical Journal*, Vol. 102, No. 1011, 1998, pp. 25–29.

[13] Raisinghani, S. C., Ghosh, A. K., and Khubchandani, S., "Estimation of Aircraft Lateral-Directional Parameters Using Neural Networks," *Journal of Aircraft*, Vol. 35, No. 6, 1998, pp. 876–881. doi:10.2514/2.2407

[14] Vijaykumar, M., Omkar, S. N., Ganguli, R., Sampath, P., and Suresh, S., "Identification of Helicopter Dynamics Using Recurrent Neural Networks and Flight Data," *Journal of the American Helicopter Society*, Vol. 51, No. 2, 2006, pp. 164–174. doi:10.4050/JAHS.51.164

[15] Mullur, A. A., and Messac, A., "Extended Radial Basis Functions: More Flexible and Effective Metamodeling," *AIAA Journal*, Vol. 43, No. 6, 2005, pp. 1306–1315. doi:10.2514/1.11292

[16] Rocha, H., Li, W., and Hahn, A., "Principal Component Regression for Fitting Wing Weight Data of Subsonic Transports," *Journal of Aircraft*, Vol. 43, No. 6, 2006, pp. 1925–1936. doi:10.2514/1.21934

[17] Praveen, C., and Duvigneau, R., "Radial Basis Functions and Kriging Metamodels for Aerodynamic Optimization," INRIA Rept. No. 6151, 2007.

[18] Forrester, A. J., and Keane, A. J., "Recent Advances in Surrogate-Based Optimization," *Progress in Aerospace Sciences*, Vol. 45, Nos. 1–3, 2009, pp. 50–79. doi:10.1016/j.paerosci.2008.11.001

[19] Cristianini, N., and Shawe–Taylor, J., *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge Univ. Press, Oxford, 2000.

[20] Smola, A. J., and Schölkopf, B., "A Tutorial on Support Vector Regression," *Statistics and Computing*, Vol. 14, No. 3, 2004, pp. 199–222. doi:10.1023/B:STCO.0000035301.49549.88

[21] Suykens, J. A. K., Horvath, G., Basu, S., Micchelli, C., and Vandewalle, J. (eds.), *Advances in Learning Theory: Methods, Models and Applications*, NATO Science Series III: Computer & Systems Sciences, Vol. 190, IOS Press, Amsterdam, 2003.

[22] Cressie, N. A. C., "The Origins of Kriging," *Mathematical geology*, Vol. 22, No. 3, 1990, pp. 239–252. doi:10.1007/BF00889887

[23] Clark, I., and Harper, W. V., *Practical Geostatistics 2000*, Ecosse North America, Columbus, OH, 2000.

[24] Forrester, A. I. J., Keane, A. J., and Bressloff, N. W., "Design and Analysis of 'Noisy' Computer Experiments," *AIAA Journal*, Vol. 44, No. 10, 2006, pp. 2331–2339. doi:10.2514/1.20068

[25] Hu, J., "Methods of Generating Surfaces in Environmental GIS Applications," *Proceedings of 1995 ESRI Conference*, Palm Springs, CA, 1995.

[26] Toropov, V. V., Schramm, U., Sahai, A., Jones, R. D., and Zeguer, T., "Design Optimization and Stochastic Analysis Based on the Moving Least Squares Method," *6th World Congress of Structural and Multidisciplinary Optimization*, Rio de Janeiro, May–June 2005.

[27] Cleveland, W. S., "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, Vol. 74, No. 368, 1979, pp. 829–836. doi:10.2307/2286407

[28] Fan, J., Gasser, T., Gijbels, I., Brockmann, M., and Engel, J., "Local Polynomial Regression: Optimal Kernels and Asymptotic Minimax Efficiency," *Annals of the Institute of Statistical Mathematics*, Vol. 49, No. 1, 1997, pp. 79–99.

[29] Fan, J., and Gijbels, I., *Local Polynomial Modelling and Its Applications*, Chapman & Hall, London, 1996.

[30] Ruppert, D., Sheather, S. J., and Wand, M. P., "An Effective Bandwidth Selector for Local Least Squares Regression," *Journal of the American Statistical Association*, Vol. 90, No. 432, 1995, pp. 1257–1270. doi:10.2307/2291516

[31] Benini, E., and Ponza, R., "Nonparametric Fitting of Aerodynamic Data Using Smoothing Thin Plate Spline," *AIAA Journal*, Vol. 48, No. 7, 2010, pp. 1403–1419. doi:10.2514/1.J050028

[32] Wahba, G., *Spline Models for Observation Data*, Society for Industrial and Applied Mathematics, Philadelphia, 1990.

[33] Meinguet, J., "Multivariate Interpolation of Arbitrary Points Made Simple," *Zeitschrift fur Angewandte Mathematik und Physik*, Vol. 30, No. 2, 1979, pp. 292–304. doi:10.1007/BF01601941

[34] Gu, C., and Wahba, G., "Minimizing GCV/GML Scores with Multiple Smoothing Parameters via the Newton Method," *SIAM Journal on Scientific and Statistical Computing*, Vol. 12, No. 2, 1991, pp. 383–398. doi:10.1137/0912021

[35] Wood, S. N., "Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties," *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 62, Part 2, 2000, pp. 413–428. doi:10.1111/1467-9868.00240

[36] Turk, G., and O'Brien, J. F., "Variational Implicit Surfaces," Georgia Inst. of Technology Tech. Rept. GIT-GVU-99-15, 1999.

[37] Bellmann, R., *Adaptive Control Processes: A Guided Tour*, Princeton Univ. Press, Princeton, NJ. 1961.

[38] Shapiro, A. H., *The Dynamics and Thermodynamics of Compressible Fluid Flow*, Vol. 1, Wiley, New York, 1977.

[39] Abbott, I. H., and Von Doenhoff, A. E., *Theory of Wing Sections*, Dover Publications, New York, 1959.

A. Messac
Associate Editor