

Surrogate modelling for aerodynamic coefficients prediction in aeronautical configurations

Carlos Paulete-Periáñez*, Esther Andrés-Pérez**, Carlos Lozano*

*Theoretical and Computational Aerodynamics Branch, National Institute for Aerospace Technology (INTA), Ctra. Ajalvir, km. 4. 28850 Torrejón de Ardoz (Spain), pauletepc@inta.es, lozanorc@inta.es, www.inta.es

** Engineering Department, ISDEFE & Technical University of Madrid UPM Ctra. Ajalvir, km. 4. 28850 Torrejón de Ardoz (Spain), eandres@isdefe.es, www.isdefe.es, www.upm.es

Abstract

European aeronautical industries are pursuing to reduce both fuel consumption and pollutant emissions in order to meet the exigent goals established by ACARE (Vision 2020 and FlightPath 2050) and, therefore, maintain the competitiveness of the European Aviation in the global world market. For that purpose, efficient and fast aerodynamic optimization methods are crucial, but the cost of CFD simulations is still a serious bottleneck. This paper investigates support vector machines as surrogate models for quick prediction of aerodynamic coefficients. Several tests are performed, focusing on the influence of the database selection and size on the precision of the surrogate model.

1. Brief introduction and literature review

Surrogate models approximate a function based on a set of training points and can then predict the function at new points. In engineering, these models are widely used because they are fast to train and are able to provide quick and accurate predictions. However, the prediction time of these techniques usually increases with the size of the dataset, and the high computational cost of CFD simulations (required to build the training database) is still a bottleneck in aerodynamic shape optimization processes. This paper analyses how surrogate modelling can be used to generate a quick prediction of the aerodynamic coefficients [1-3] and therefore be a candidate to be integrated in optimization frameworks. In particular, support vector machines (SVMs) will be used to predict the aerodynamic behaviour on 2D (NACA0012 and RAE2822 airfoils) and 3D (DPW wing) configurations.

Generally speaking, surrogate modeling [1–5] refers to a group of techniques that make use of previously obtained sampled data in order to build surrogate models, which are subsequently used to predict the value of variables at new points in the design space. This group of techniques has been developed from many different fields, such as chemical and materials engineering [6, 7], agriculture and ecology [8, 9], medicine [10–12], economy [13] and the aerospace field [14-16]. The application to the aerospace field, and particularly, to aerodynamic data prediction from CFD, wind tunnel and flight testing data, can allow a first stage exploration of new areas in the design space, without the need of expensive simulations, wind tunnel or flight testing, and in this way, reduce the number of required experiments. There are well-documented examples of the applicability of soft computing approaches (neural networks and evolutionary-based techniques) in a broad range of prediction and optimization problems including some parts of aerodynamic or multidisciplinary optimization processes. The majority of published studies use some type of evolutionary algorithms hybridized with neural networks as metamodels: in [17] a complete study of different types of neural networks working as metamodels in an aerodynamic shape design problem is carried out. The study includes multilayer perceptrons and radial basis functions networks. In [18] a grid-based hierarchical evolutionary algorithm hybridized with a radial basis function network is proposed also in different parts of aerodynamic design problems. There are also recent works discussing different aspect of hybridizing evolutionary algorithms and neural networks as metamodels for airfoil design [19-22]. Other authors have tested the performance of alternative evolutionary approaches such as particle swarm optimization [23]. There are also alternative methods applied to aerodynamic shape design, such as fuzzy logic approaches [24], multiobjective algorithms [25], works involving cokriging techniques [26], and papers that describe computation frameworks developed to enhance the design process [27-29].

This article is structured as follows: Section 2 presents the applied surrogate modelling strategy and Section 3 collects the numerical results. Finally, the main conclusions extracted from this study are summarized in Section 4.

2. Surrogate modelling strategy

This section briefly introduces the key components of the proposed surrogate model training and validation strategy, SVMs through the libSVM library, and the CFD tool to be used.

2.1 Support Vector Machines for regression (SVMr)

Support Vector Machines for regression (SVMr) are a powerful tool used on the machine learning field, and as a modelling tool in many regression problems in engineering. The SVMr can be solved as a convex optimization problem using kernel theory to face nonlinear problems. The SVMr consider not only the prediction error but also the generalization of the model. The SVMr consist of training a model with the form $y = w^T \Phi(x) + b$ given a set of parameters $C = \{(x_i, y_i), i = 1, 2, \dots, l\}$, to minimize a general risk function of the form:

$$R[f] = \frac{1}{2} \|w\|^2 + \frac{1}{2} C \sum_{i=1}^l L(y_i, f(x)) \quad (1)$$

where w controls the smoothness of the model, $\Phi(x)$ is a function of projection of the input space to the feature space, b is a parameter of bias, x_i is a feature vector of the input space with dimension N , y_i is the output value to be estimated and $L(y_i, f(x))$ is the loss function selected. In this paper, the L1-SVR (L1 support vector regression) is used, characterized by an ε -insensitive loss function

$$L(y_i, f(x)) = |y_i - f(x_i)|_\varepsilon \quad (2)$$

In order to train this model, the following optimization problem has to be solved

$$\min \left(\frac{1}{2} \|w\|^2 + \frac{1}{2} C \sum_{i=1}^l \xi_i + \xi_i^* \right) \quad (3)$$

subject to:

$$\begin{aligned} y_i - w^T \Phi(x) - b &\leq \varepsilon + \xi_i, i = 1, \dots, l \\ -y_i + w^T \Phi(x) + b &\leq \varepsilon + \xi_i^*, i = 1, \dots, l \\ \xi_i, \xi_i^* &\geq 0, i = 1, \dots, l \end{aligned} \quad (4)$$

To do this, a dual form is usually applied, obtained from the minimization of the Lagrange function that joins the function to minimize and the restrictions. The dual form is:

$$\max \left(-\frac{1}{2} \sum_{i,j=1}^l (\alpha_i + \alpha_i^*)(\alpha_j + \alpha_j^*) K(x_i + x_j) - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l y_i (\alpha_i + \alpha_i^*) \right) \quad (5)$$

subject to:

$$\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0; \alpha_i, \alpha_i^* \in [0, C] \quad (6)$$

In addition to the restrictions, also must be taken in account the Karush-Kuhn-Tucker conditions [30] and obtain the bias value. In the dual formulation we must emphasize the apparition of the kernel function $K(x_i, x_j)$, which is equivalent to the scalar product $\langle \Phi(x_i), \Phi(x_j) \rangle$. In our case, the kernel function is a Gaussian function:

$$K(x_i, x_j) = \exp(-\gamma \cdot \|x_i - x_j\|^2) \quad (7)$$

The final form of the regression model depends on the Lagrange multipliers α_i, α_i^* , following the expression:

$$f(x) \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (8)$$

In this way, the SVMr model depends on three parameters, ε , C and γ . ε controls the error margin permitted for the model, as can be seen in equations (4), (5), C controls the number of outliers allowed on the optimization of the function equation (4). Finally, γ determines the Gaussian variance for the kernel. Depending on the selection of these values, the model can have a different performance. To obtain the best SVM performance, a search of the most suitable combination of these three parameters must be carried out, usually by using cross validation techniques over the training set. To reduce the computational time of this process, different methods have been proposed in the literature to reduce the search space related to these parameters. In this case, it has been applied the one developed by Ortiz-García et al [30] which has proven to require pretty short search times.

2.2. LibSVM

LIBSVM is a library for support vector machines (SVM) [31] which has gained wide popularity in machine learning and many other areas. A typical use of LIBSVM involves two steps: first, training a data set to obtain a model and second, using the model to predict information of a testing data set. For SVC and SVR, LIBSVM can also output probability estimates. Many extensions of LIBSVM are available at libsvmtools.

2.3. Flow solver (DLR TAU)

Fluid flow computations are carried out with the TAU code [32], and the grids were generated with Centaur [33]. The TAU-Code solves the compressible, three-dimensional Reynolds-Averaged Navier-Stokes equations using a finite volume formulation. The TAU-Code is based on a hybrid unstructured-grid approach, which makes use of the advantages of semi-structured prismatic grids in the viscous shear layers near walls, and the flexibility in grid generation offered by tetrahedral grids in the surrounding flow volume. A dual-grid approach with an edge-based data structure is used in order to make the flow solver independent from the cell types used in the initial grid.

The TAU-Code consists of several different modules, including:

- The Grid Partitioner, which splits the primary grid into n number of subgrids for n processors.
- The Preprocessor module, which uses the information from the initial grid to create a dual-grid and secondly coarser grids for multi-grid.
- The Solver module, which performs the flow calculations on the dual-grid.
- The Postprocessing module, which is used to convert results to formats usable by popular visualization tools.

Together, all modules are available with python interfaces for computing complex application, e.g. unsteady cases, complete force polar curves or fluid-structure couplings in an automatic framework. Furthermore, it eases the usage on highly massive parallel computers to execute applications.

2.4. Surrogate model training and validation strategy

The surrogate model that estimates the aerodynamic coefficients for a given geometry over a range of different flow regimes and angles of attack is built using SVM regression techniques. To do this, we generate two databases, one for testing and the other for training. The testing library is generated only once at the beginning of the training process and stored. It contains the aerodynamic coefficients of the full range of Mach number and angle of attack considered for each geometry. The training database on the other hand must be computed such that the largest amount of features in the range of angles of attack and Mach numbers are captured. This is achieved implementing Latin Hypercube Sampling, which ensures a spread of the training parameters for any number of training cases. Therefore, one can obtain an arbitrarily long set of angles of attack and Mach numbers that are representative of the whole domain considered.

The following picture summarizes the proposed training and prediction approach.

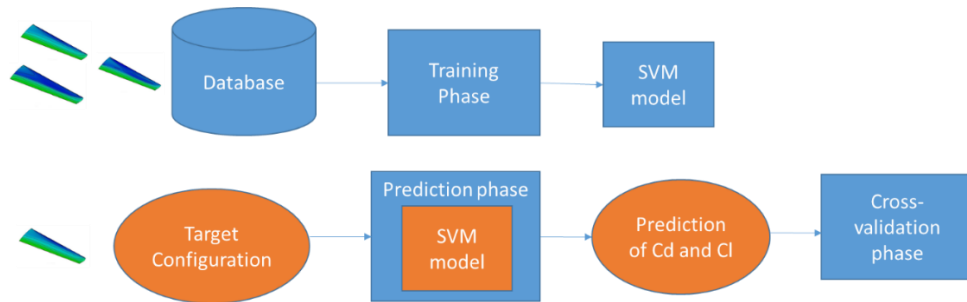


Figure 1: Proposed methodology for aerodynamic prediction using a support vector machine based surrogate model.

3. Numerical results

Three test cases have been considered based on three well-known geometries (Figure 2): the NACA0012 airfoil [34], the RAE2822 airfoil [35] and the DPW-W1 wing from the 3rd AIAA Drag Prediction Workshop [36].

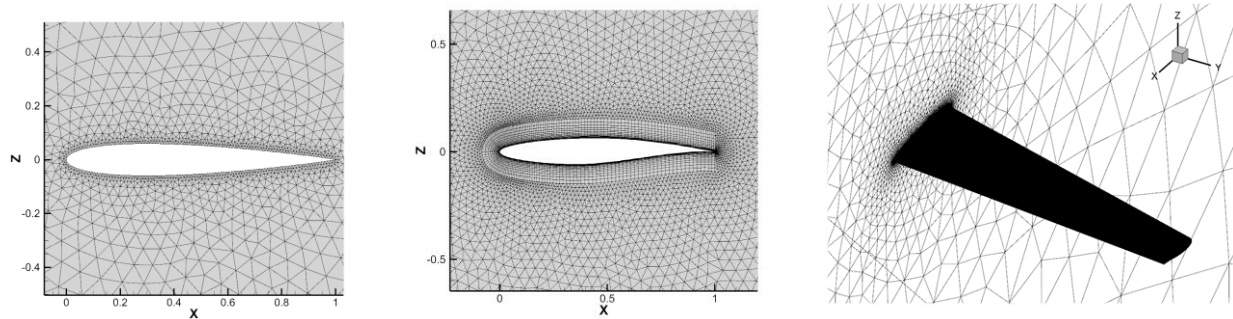


Figure 2: NACA0012, RAE2822 and DPW grids

3.1 Results on the NACA0012 test case

The proposed surrogate modelling strategy was tested in different databases for various training set sizes. In particular, as will be explained during this section, 25, 50 and 100 samples were used to build several models with different prediction accuracy. A triangular computational mesh with 5990 points (Figure 2) was used in the inviscid flow analysis.

3.1.1 Surrogate training with 25 samples

First, the model was trained using 25 samples, which were generated by LHS with Mach number ranging from 0.1 to 0.8 and AoA ranging from 0 to 17. The following figures show the results of the model prediction for Mach numbers 0.3, 0.5 and 0.7. For these results, 10 different sets of 25 samples were generated to train the model and the mean values were considered. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value.

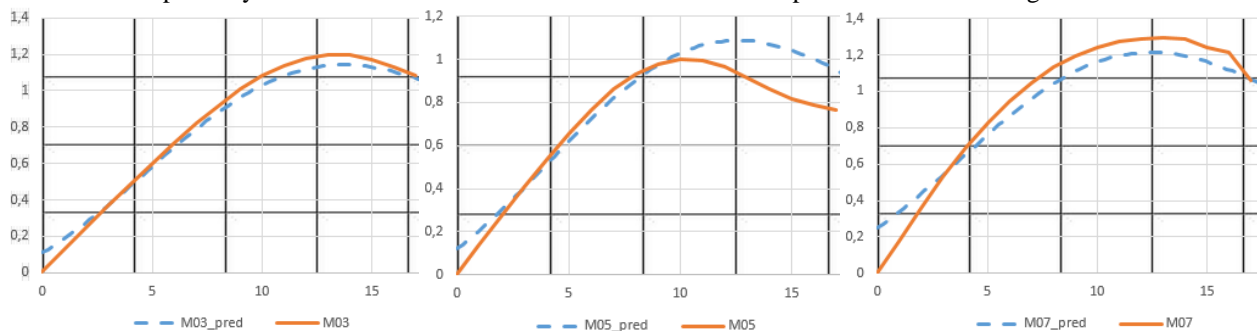


Figure 3: C_l model prediction vs. CFD C_l for Mach numbers 0.3, 0.5 and 0.7 (25 samples)

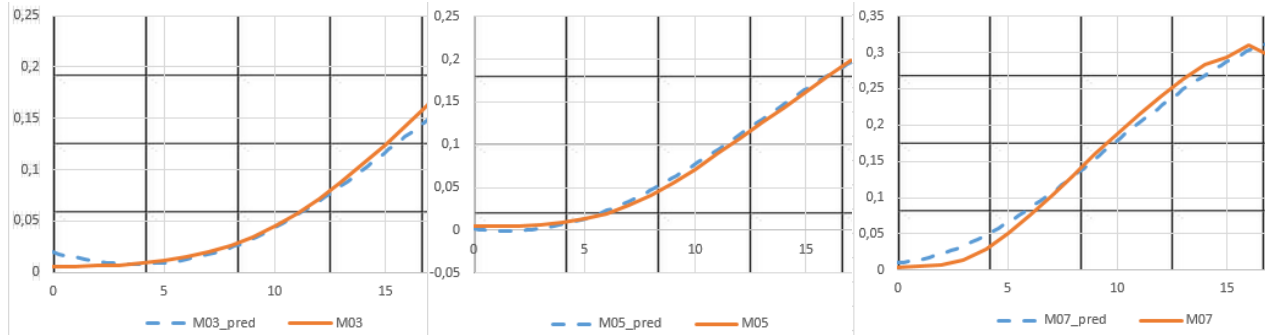


Figure 4: Cd model prediction vs. CFD Cd for Mach numbers 0.3, 0.5 and 0.7 (25 samples)

3.1.2 Surrogate training with 50 samples

Here, the model was trained using 50 samples, which were generated by LHS with Mach number ranging from 0.1 to 0.8 and AoA ranging from 0 to 17. The following figures show the results of the model prediction for Mach numbers 0.3, 0.5 and 0.7. For these results, 10 different sets of 50 samples were generated to train the model and the mean values were considered. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

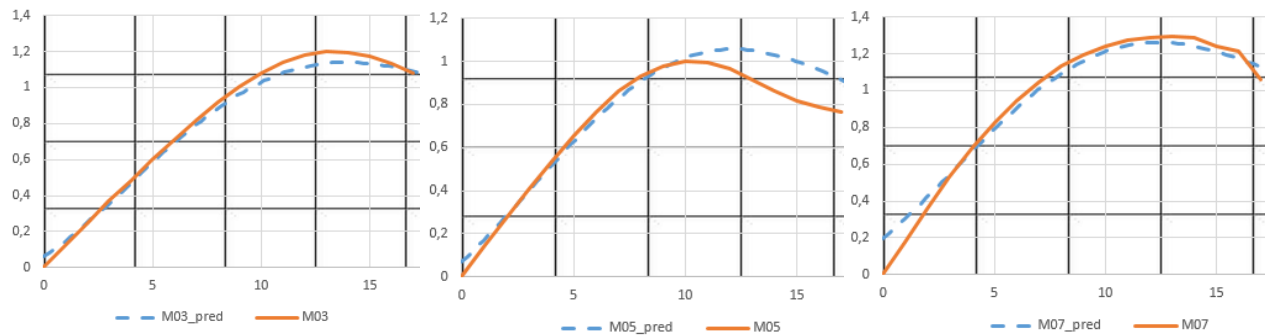


Figure 5: Cl model prediction vs. CFD Cl for Mach numbers 0.3, 0.5 and 0.7 (50 samples)

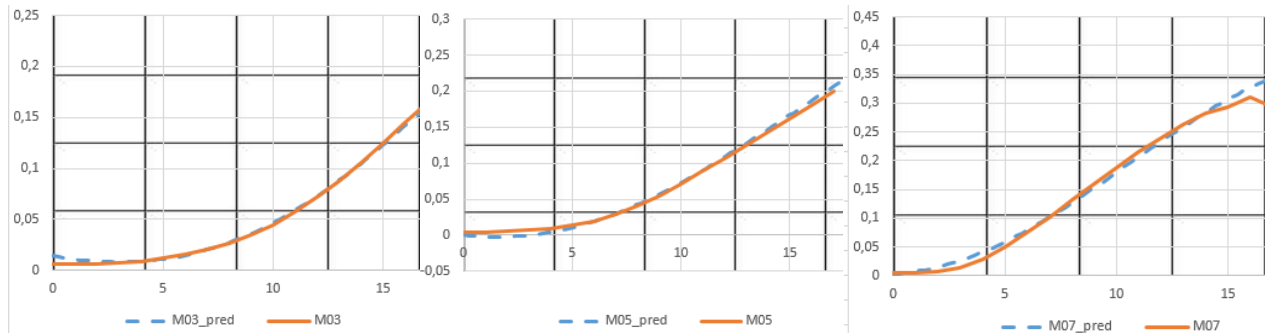


Figure 6: Cd model prediction vs. CFD Cd for Mach numbers 0.3, 0.5 and 0.7 (50 samples)

3.1.3 Surrogate training with 100 samples

The following figures show the results of the model prediction for Mach numbers 0.3, 0.5 and 0.6. For these results, 10 different sets of 100 samples were generated to train the model and again the mean values were considered. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

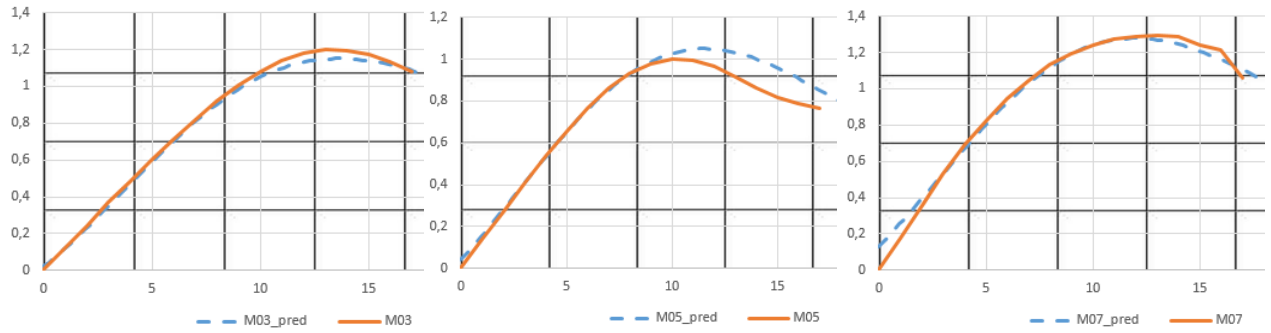


Figure 7: Cl model prediction vs. CFD Cl for Mach numbers 0.3, 0.5 and 0.7 (100 samples)

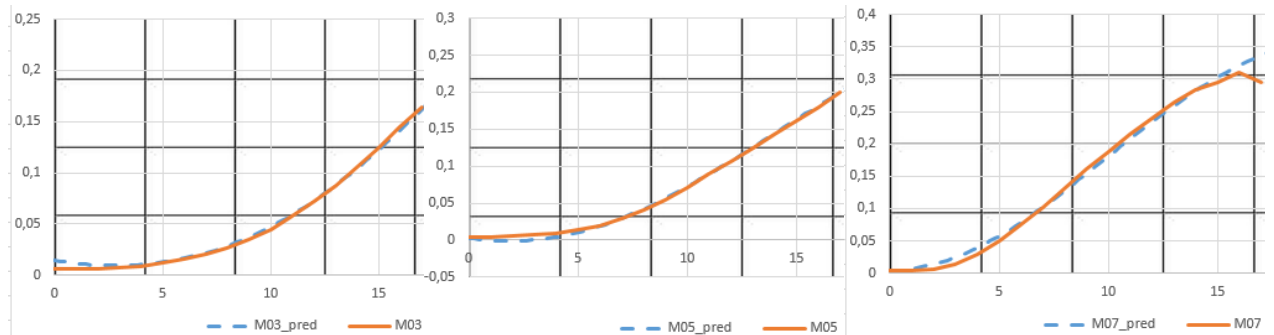


Figure 8: Cd model prediction vs. CFD Cd for Mach numbers 0.3, 0.5 and 0.7 (100 samples)

It can be observed, that, as expected, the bigger the training database is, the more accurate the surrogate is, especially in the case of the prediction of lift coefficient. However, in case of the drag, there is not significant differences in accuracy and a reduced number of samples in the training database could be used.

3.1.4 Surrogate error validation with a polar database

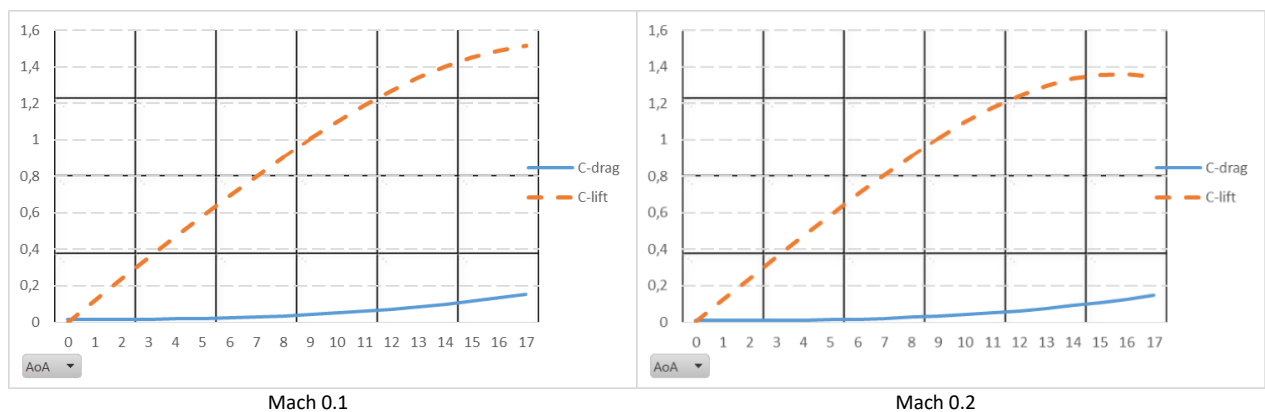
Finally, a polar on the NACA0012 test case was run. The ranges of AoA and Mach number are shown in the following table:

Table 1: AoA and Mach number ranges

	Minimum value	Maximum value	Increment
Mach	0.1	0.8	0.1
AoA	0	17	1

With these ranges, a database of 144 samples was generated. The computational time to generate this database what around 25 minutes in a single Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz core.

The following figures show the polar plots for different Mach numbers (X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively):



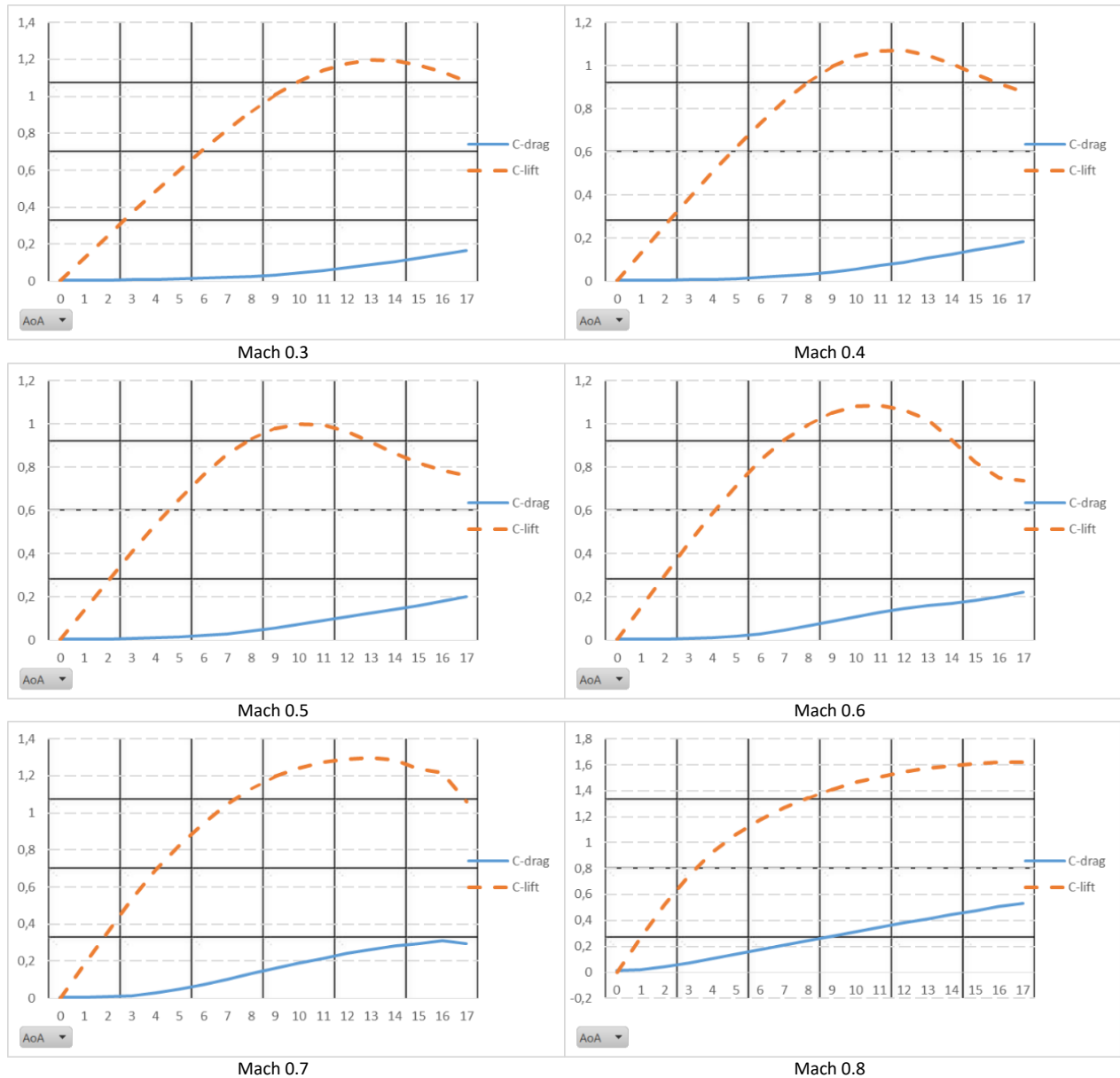
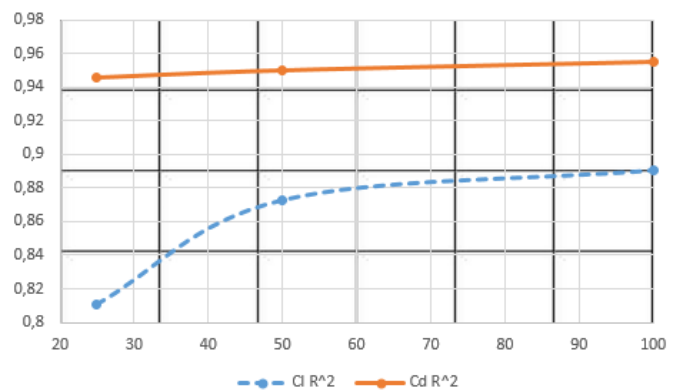


Figure 9: NACA0012 polar for Mach numbers ranging from 0.1 to 0.8

This database was used to compute the R^2 values of the surrogate model for different training set sizes (25, 50 and 100 training samples), and the values are displayed in the following figure:

N_{train}	$C_l R^2$	$C_d R^2$
25	0,8103	0,9461
50	0,8727	0,95
100	0,8906	0,9547

Figure 10: C_d and $C_l R^2$ values for different training set sizes

As commented before, it can be observed, that, as expected, the bigger the training database is, the more accurate the surrogate is, especially in the case of the prediction of lift coefficient, where R^2 values drastically increase when going from 25 samples to 50 samples. However, in case of the drag, there is not significant differences in accuracy and a reduced number of samples in the training database could be used.

3.2 Results on the RAE2822 test case

The surrogate model was tested on a polar database for a viscous RAE2822 airfoil using various training dataset sizes. Particularly, 10, 25, 50 and 100 samples were used to train the models which show different prediction accuracy. A hybrid computational mesh with 27874 points (Figure 2) was used in the analysis. The Reynolds number was 6.5 million.

3.2.1 Surrogate training with 10 samples

Initially, the model is trained with 10 samples generated using LHS with Mach number ranging from 0.1 to 0.9 and AoA from 0 to 15. The predicted and test polar curves are shown in the following figures for Mach numbers 0.3 and 0.6. These results are a product of 10 different training sets of 10 samples, on which the model is trained. The resulting predictions are then averaged for each AoA. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

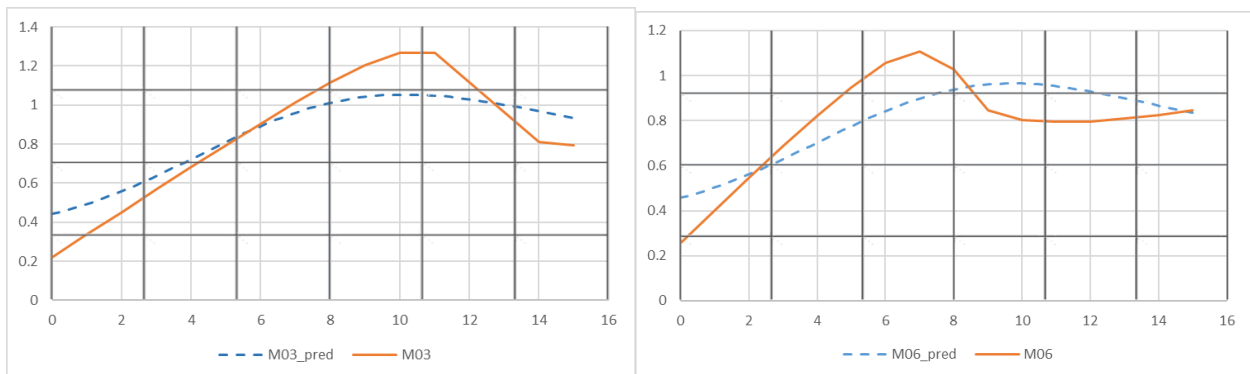


Figure 11: Cl model prediction vs. CFD Cl for Mach numbers 0.3 and 0.6 (10 samples)

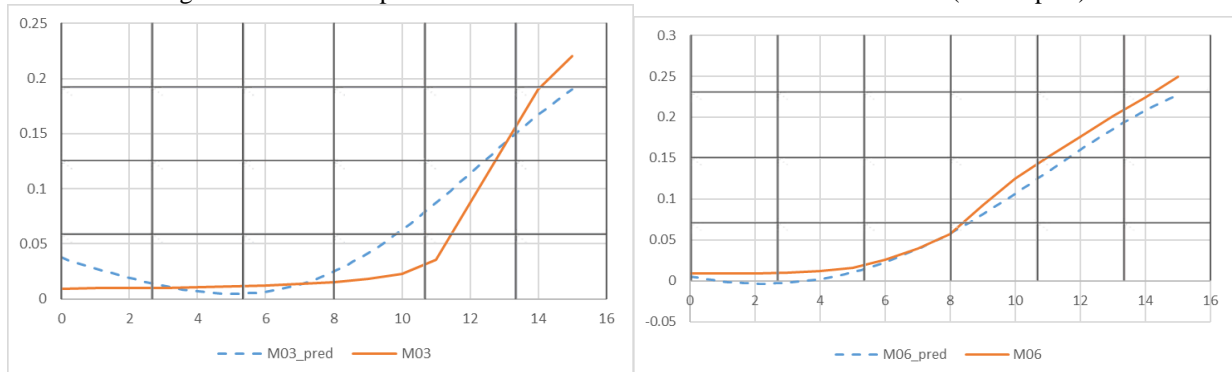


Figure 12: Cd model prediction vs. CFD Cd for Mach numbers 0.3 and 0.6 (10 samples)

3.2.2 Surrogate training with 25 samples

Now, the model is trained against 25 samples generated using LHS with Mach number ranging from 0.1 to 0.9 and AoA from 0 to 15. The predicted and test polar curves are shown in the following figures for Mach numbers 0.3 and 0.6. These results are a product of 10 different training sets of 25 samples, on which the model is trained. The resulting predictions are then averaged for each AoA. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

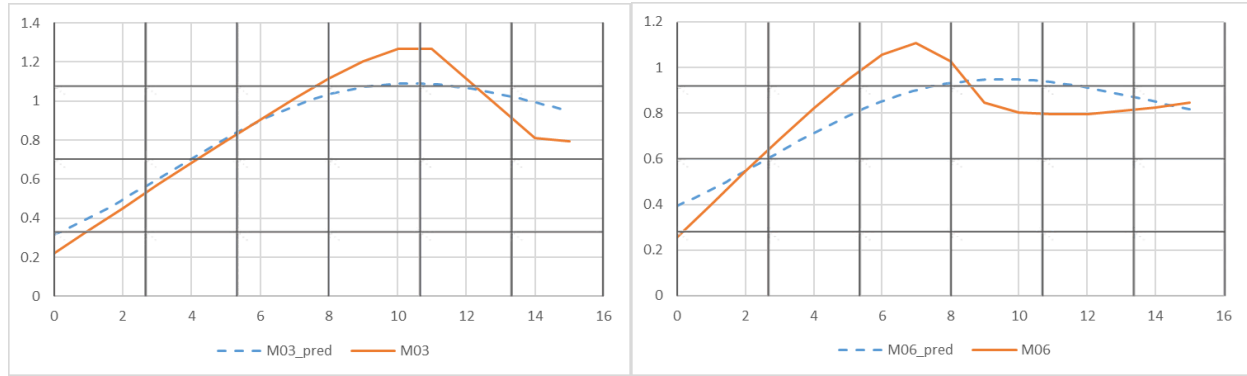


Figure 13: Cl model prediction vs. CFD Cl for Mach numbers 0.3 and 0.6 (25 samples)

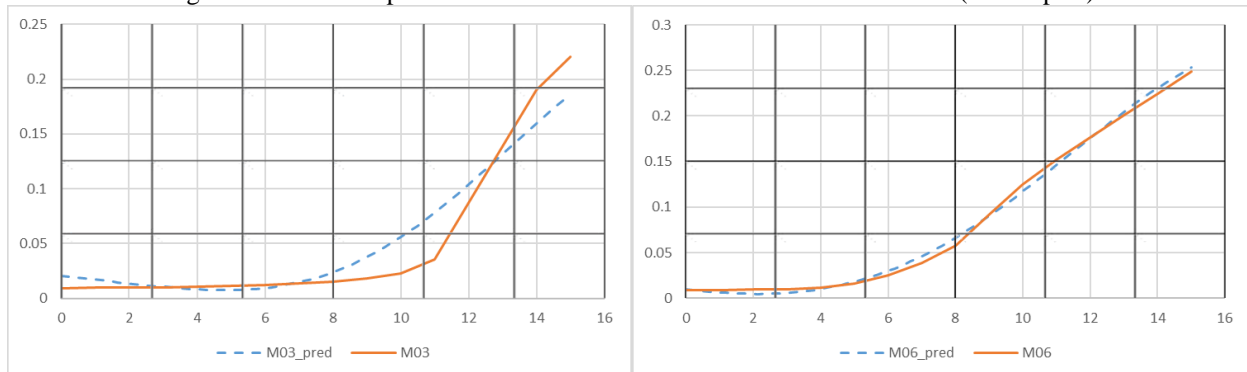


Figure 14: Cd model prediction vs. CFD Cd for Mach numbers 0.3 and 0.6 (25 samples)

In this case, the model is trained against 50 samples generated using LHS with Mach number ranging from 0.1 to 0.9 and AoA from 0 to 15. The predicted and test polar curves are shown in the following figures for Mach numbers 0.3 and 0.6. These results are a product of 10 different training sets of 50 samples, on which the model is trained. The resulting predictions are then averaged for each AoA. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

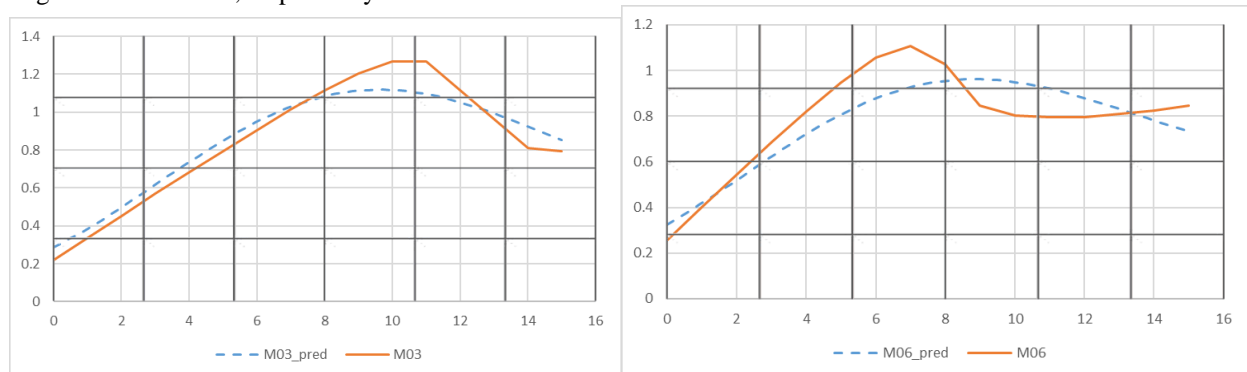


Figure 15: Cl model prediction vs. CFD Cl for Mach numbers 0.3 and 0.6 (50 samples)

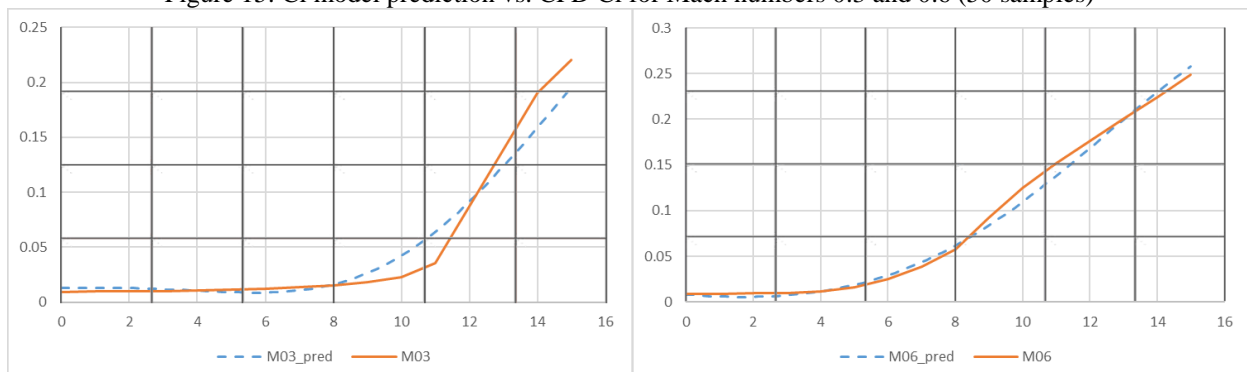


Figure 16: Cd model prediction vs. CFD Cd for Mach numbers 0.3 and 0.6 (50 samples)

3.2.3 Surrogate training with 100 samples

Initially, the model is trained against 100 samples generated using LHS with Mach number ranging from 0.1 to 0.9 and AoA from 0 to 15. The predicted and test polar curves are shown in the following figures for Mach numbers 0.3 and 0.6. These results are a product of 10 different training sets of 100 samples, on which the model is trained. The resulting predictions are then averaged for each AoA. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

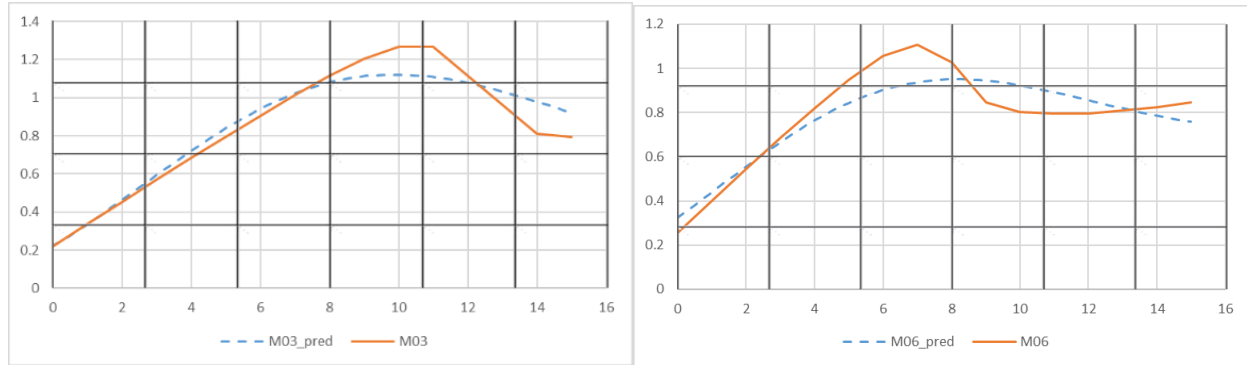


Figure 17: Cl model prediction vs. CFD Cl for Mach numbers 0.3 and 0.6 (100 samples)

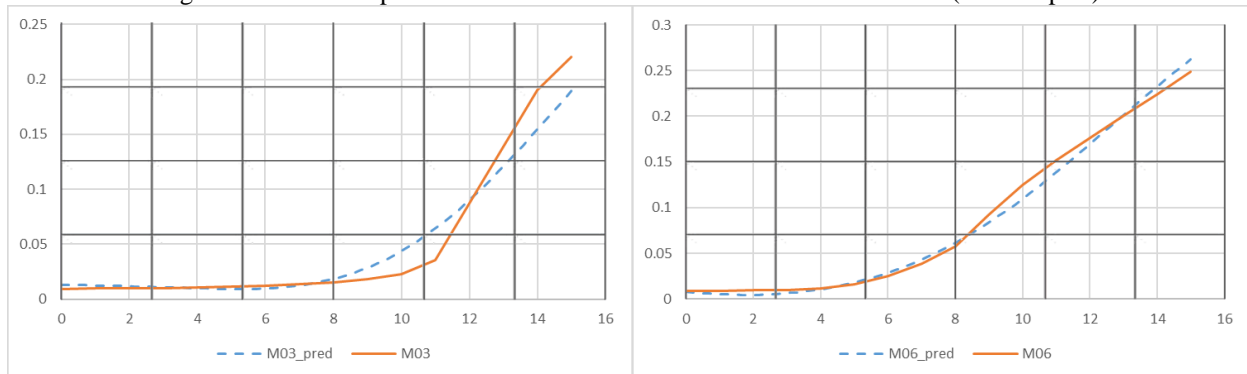


Figure 18: Cd model prediction vs. CFD Cd for Mach numbers 0.3 and 0.6 (100 samples)

As in the previous test case, the trend is that the bigger the training database is, the more accurate the surrogate is. In this case, it can be observed that the lift coefficient is reasonable well predicted but in the case of the drag, some differences occur, which suggests that it is necessary to include more samples in certain regions of the design space where flow discrepancies may appear.

3.2.4 Surrogate error validation with a polar database

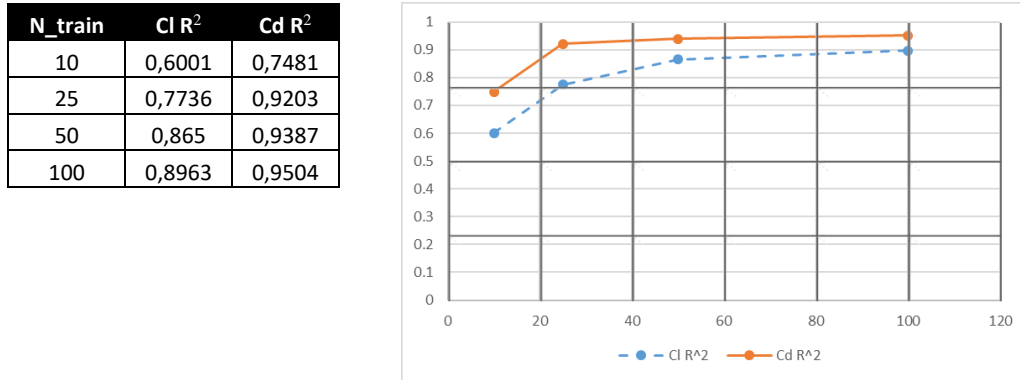
Finally, a polar on the RAE2822 test case was run. The ranges of AoA and Mach number are shown in the following table:

Table 2: AoA and Mach number ranges

	Minimum value	Maximum value	Increment
Mach	0.1	0.9	0.1
AoA	0	15	1

With these ranges, a database of 100 samples was generated. The computational time to generate this database was around 6 hours in 24 Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz cores.

This database was used to compute the R^2 values of the surrogate model for different training set sizes (10, 25, 50 and 100 training samples), and the values are displayed in the following table:

Figure 19: Cd and Cl R² values for different training set sizes

In this case, both lift and drag R² values drastically increase when going from 10 samples to 50 samples. However, no significant improvements are achieved by incrementing the database size to 100 samples.

3.3 Results on the DPW test case

As in the previous test cases, the proposed surrogate modelling strategy was tested in different databases for various training set sizes. In particular, as will be explained during this section, 10, 25, 50 and 100 samples were used to build several models with different prediction accuracy. A computational mesh with 427919 points (Figure 2) was used in the inviscid flow analysis.

3.3.1 Surrogate training with 10 samples

First, the model was trained using 10 samples, which were generated by LHS with Mach number ranging from 0.1 to 0.8 and AoA ranging from 0 to 15. The following figures show the results of the model prediction for Mach numbers 0.3, 0.5 and 0.6. For these results, 10 different sets of 10 samples were generated to train the model and the mean values were considered. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

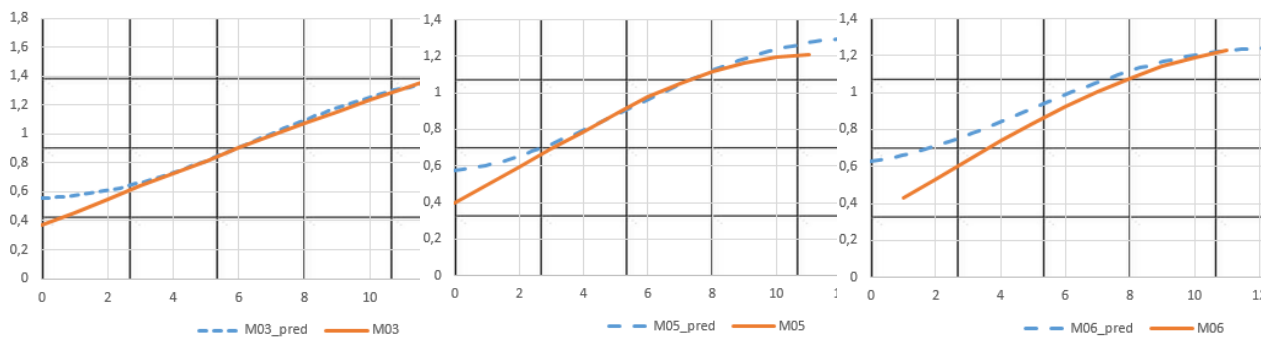


Figure 20: Cl model prediction vs. CFD Cl for Mach numbers 0.3, 0.5 and 0.6 (10 samples)

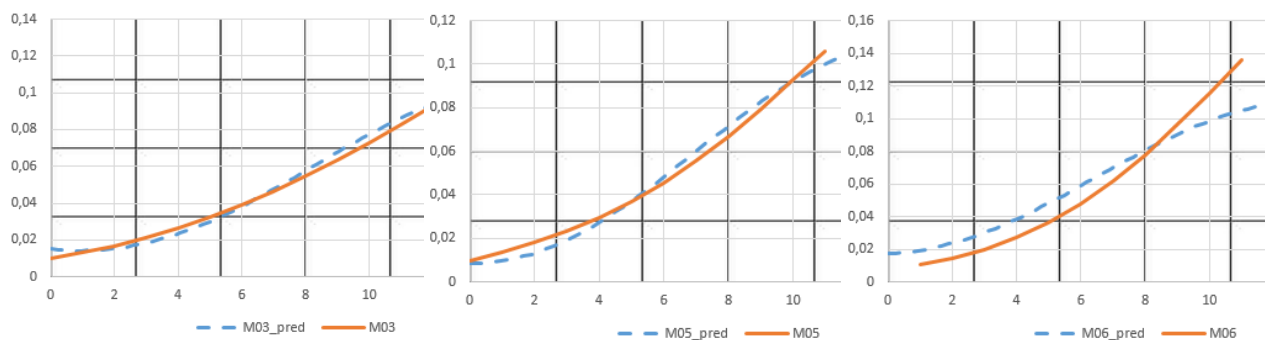


Figure 21: Cd model prediction vs. CFD Cd for Mach numbers 0.3, 0.5 and 0.6 (10 samples)

3.3.2 Surrogate training with 25 samples

The following figures show the results of the model prediction for Mach numbers 0.3, 0.5 and 0.6. For these results, 10 different sets of 25 samples were generated to train the model and the mean values were considered. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

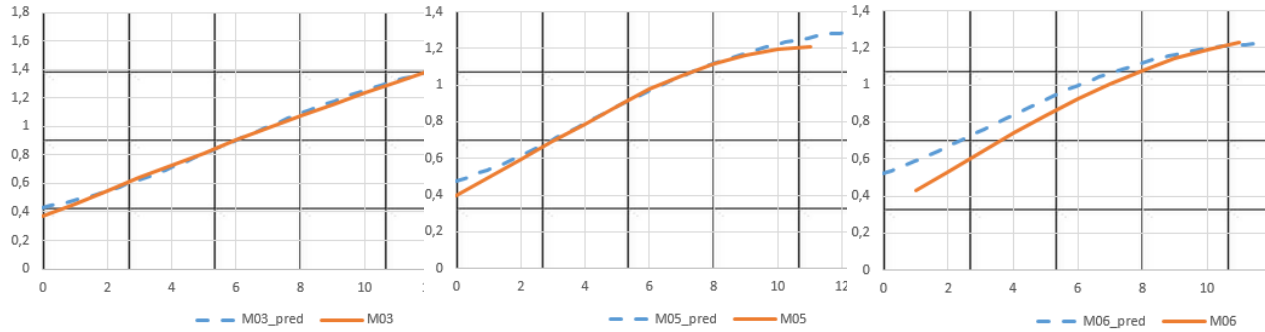


Figure 22: Cl model prediction vs. CFD Cl for Mach numbers 0.3, 0.5 and 0.6 (25 samples)

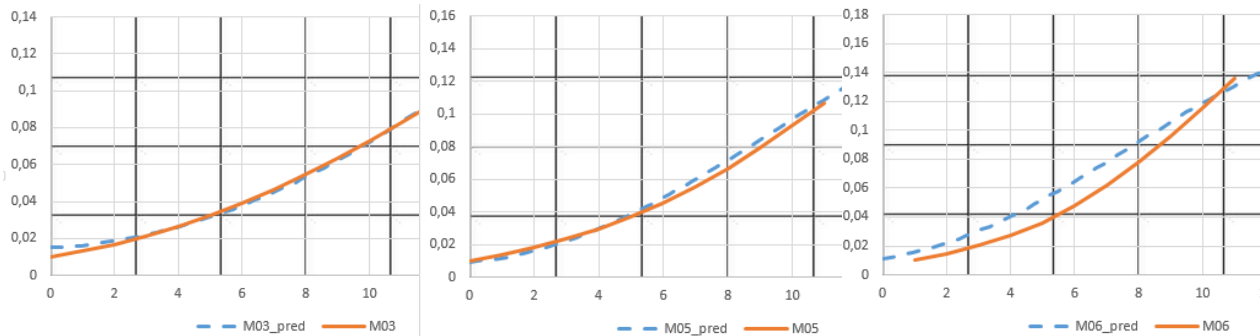


Figure 23: Cd model prediction vs. CFD Cd for Mach numbers 0.3, 0.5 and 0.6 (25 samples)

3.3.3 Surrogate training with 50 samples

The following figures show the results of the model prediction for Mach numbers 0.3, 0.5 and 0.6. For these results, 10 different sets of 50 samples were generated to train the model and the mean values were considered. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

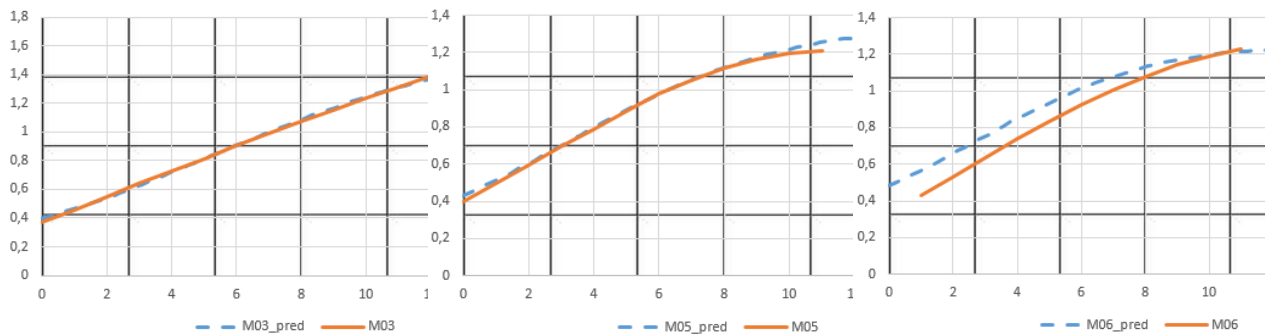


Figure 24: Cl model prediction vs. CFD Cl for Mach numbers 0.3, 0.5 and 0.6 (50 samples)

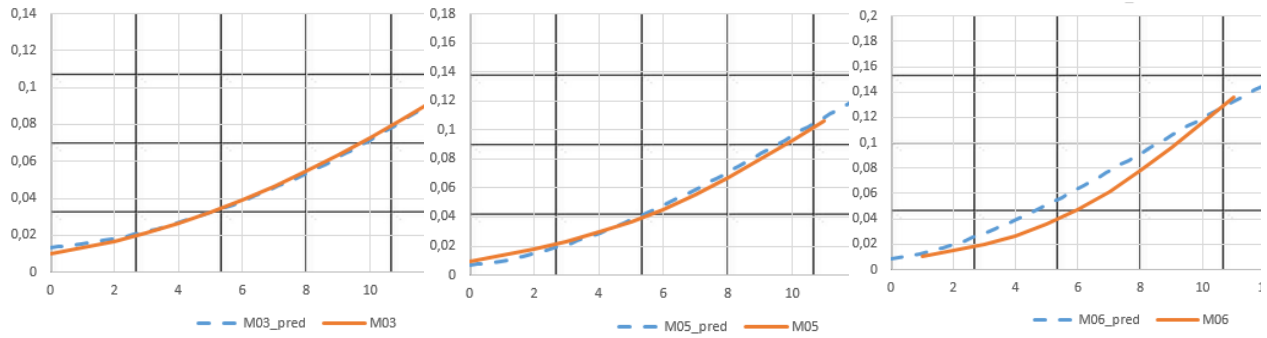


Figure 25: Cd model prediction vs. CFD Cd for Mach numbers 0.3, 0.5 and 0.6 (50 samples)

3.3.4 Surrogate training with 100 samples

The following figures show the results of the model prediction for Mach numbers 0.3, 0.5 and 0.6. For these results, 10 different sets of 100 samples were generated to train the model and the mean values were considered. These pictures show the lift and drag coefficients prediction against the lift and drag coefficients computed by the CFD tool. X axis is the AoA and the Y axis represents the lift or drag coefficient value, respectively.

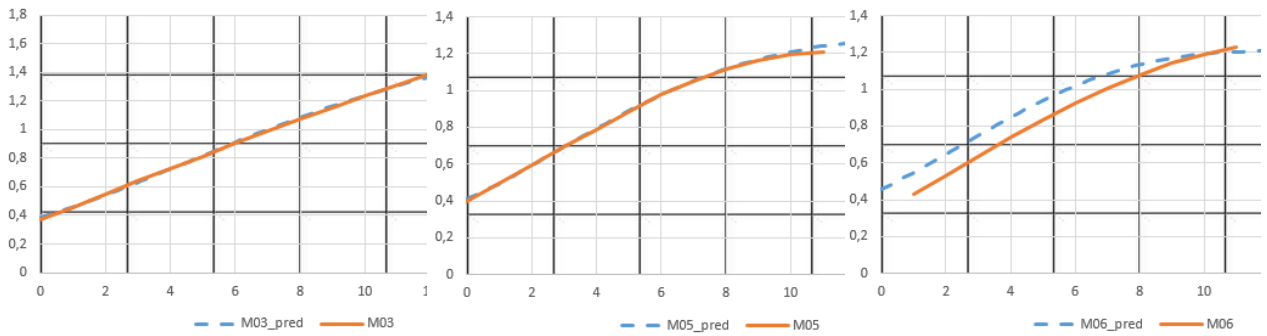


Figure 26: Cl model prediction vs. CFD Cl for Mach numbers 0.3, 0.5 and 0.6 (100 samples)

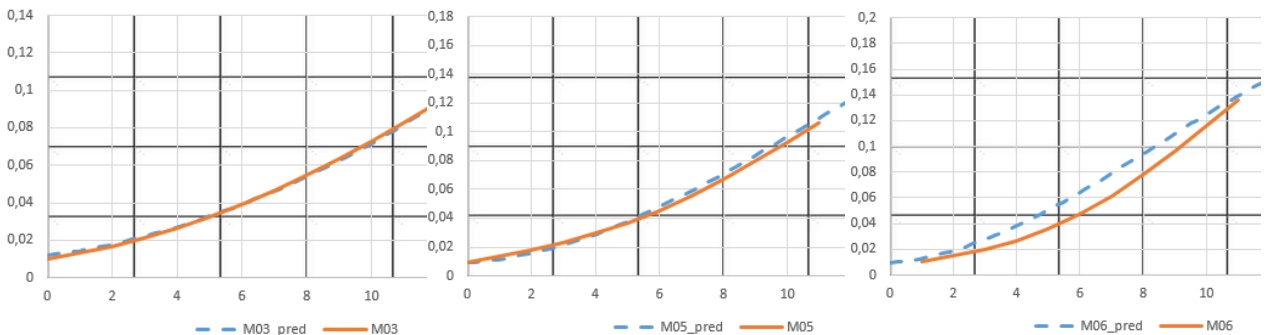


Figure 27: Cd model prediction vs. CFD Cd for Mach numbers 0.3, 0.5 and 0.6 (100 samples)

3.3.5 Surrogate error validation with a polar database

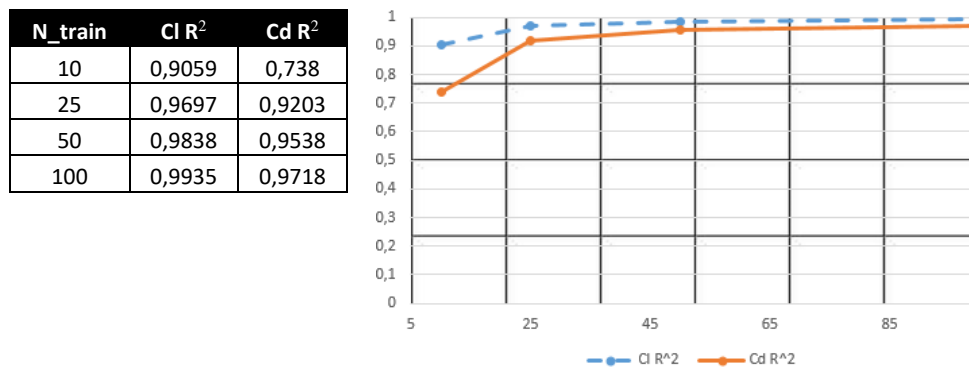
Finally, a polar on the DPW test case was run. The ranges of AoA and Mach number are shown in the following table:

Table 3: of AoA and Mach number ranges

	Minimum value	Maximum value	Increment
Mach	0.1	0.8	0.1
AoA	0	15	1

With these ranges, a database of 100 samples was generated. The computational time to generate this database was around 50 minutes in 24 Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz cores.

This database was used to compute the R^2 values of the surrogate model for different training set sizes (10, 25, 50 and 100 training samples), and the values are displayed in the following table:

Figure 28: Cd and Cl R² values for different training set sizes

This case follows the same trends as the previous one, both lift and drag R² values drastically increase when going from 10 samples to 50 samples. However, no significant improvements are achieved by incrementing the database size to 100 samples.

4. Conclusions and further work

This paper has presented a surrogate model application for the prediction of aerodynamic coefficients in aeronautical configurations, in particular, results on the 2D NACA0012 and RAE2822 airfoils and the 3D DPW wing have been shown. The results show how the surrogate model is able to predict aerodynamic data in these configurations with a reasonable level of accuracy, without the need of expensive computational fluid dynamics simulations.

The results show that, as expected, the bigger the training database is, the more accurate the surrogate is, especially in the case of the prediction of lift coefficient. However, in case of the drag, there is not as significant differences in accuracy and a reduced number of samples in the training database could be used. In all tested configurations, both lift and drag R² values drastically increase when going from 10 samples to 50 samples. However, no significant improvements are achieved by incrementing the database size to 100 samples.

Although this work is focused on CFD data, this approach is able to be used with wind tunnel and flight testing data instead of only CFD simulations. In addition, the integration of data with different fidelities to the model: CFD, wind tunnel, and flight data, could be possible based on the approach presented here. Thus, further research will be focused on this issue and other advance setting of the surrogate that could improve more its performance.

References

- [1] Dupuis, Romain, Jean-Christophe Jouhaud, and Pierre Sagaut. 2018. Aerodynamic Data Predictions for Transonic Flows via a Machine-Learning-based Surrogate Model. 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference.
- [2] Yondo, R., E. Andrés, and E. Valero. 2018. A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses. *Progress in Aerospace Sciences* 96: 23-61.
- [3] Hwang, J. T., and J.R.R.A. Martins. 2018. A fast-prediction surrogate model for large datasets. *Aerospace Science and Technology* 75: 74-87.
- [4] W.C. Carpenter, J.F.M. Barthelemy. 1992. A comparison of polynomial approximations and artificial neural nets as response surfaces. In: *A Collection of Technical Papers: The 33rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA, Dallas, TX, April 13–15, 1992, pp. 2474–2482.
- [5] E. Nikolaidis, L. Long, Q. Ling. 2000. Neural Networks and Response Surface Polynomials for Design of Vehicle Joints. *Computers & Structures* 75(6): 593-607.
- [6] N. Papila, W. Shyy, N. Fitz-Coy, R.T. Haftka. 1999. Assessment of neural net and polynomial-based techniques for aerodynamic applications. *AIAA Paper* 99-3167.
- [7] R. Jin, W. Chen, T.W. Simpson. 2000. Comparative studies of metamodeling techniques under multiple modeling criteria. *AIAA Paper* 2000-4801.
- [8] E. İçten, Z.K. Nagy, G.V. Reklaitis. 2015. Process control of a dropwise additive manufacturing system for pharmaceuticals using polynomial chaos expansion based surrogate model. *Comput. Chem. Eng.* 83: 221–231.
- [9] J. Steuben, J. Micholopoulos, A. Iliopoulos, C. Turner. 2015. Inverse characterization of composite materials via surrogate modeling. *Compos. Struct.* 132: 694–708.

- [10] X. Wu, Y. Zheng, B. Wu, et. al. 2016. Optimizing conjunctive use of surface water and groundwater for irrigation to address human-nature water conflicts: a surrogate modeling approach. *Agric. Water Manag.* 163: 380–392.
- [11] J.C. Pierson, A. Mortelliti, P.S. Barton, P.W. Lane, D.B. Lindenmayer. 2016. Evaluating the effectiveness of overstory cover as a surrogate for bird community diversity and population trends, *Ecol. Indic.* 61(2): 790–798.
- [12] M.K. Song, S.E. Ward, et al. 2015. Advance care planning and end-of-life decision making in dialysis: a randomized controlled trial targeting patients and their surrogates. *Am. J. Kidney Dis.* 66(5): 813–822.
- [13] K. Bollaerts, M. Snock, K. Simons, et al. 2014. Thyroid cancer incidence around the Belgian nuclear sites: surrogate exposure modelling. *Cancer Epidemiol.* 130: 12–26.
- [14] A.M. Osiceanu, L.E. Murao, D. Kollanur, et al. 2014. In vitro surrogate models to aid in the development of antivirals for the containment of foot-and-mouth disease outbreaks. *Antivir. Res.* 105: 59–63.
- [15] Y. Yin, P. Shang. 2015. Modified cross sample entropy and surrogate data analysis method for financial time series. *Phys. A, Stat. Mech. Appl.* 433: 17–25.
- [16] E. Andres, S. Salcedo-Sanz, F. Monge, A.M. Perez-Bellido. 2012. Efficient aerodynamic design through evolutionary programming and support vector regression algorithms. *Expert Syst. Appl.* 39: 10700–10708.
- [17] D. Viúdez-Moreiras, E. Andrés-Pérez, D. González-Juárez, M.J. Martín-Burgos. 2016. Performance comparison of Kriging and SVR surrogate models applied to the objective function prediction within aerodynamic shape optimization. In: *ECCOMAS CFD 2016*.
- [18] D. Viúdez-Moreiras, M. Martin, R. Abarca, E. Andrés-Pérez, J. Ponsín, F. Monge. 2017. Surrogate modeling for the main landing gear doors of an airbus passenger aircraft. *Aerospace Science and Technology* 68: 135–148
- [19] K. Giannakoglou, D. Papadimitriou, I. Karpolis. 2006. Aerodynamic shape design using evolutionary algorithms and new gradient-assisted metamodels. *Computational Methods in Applied Mechanics Engineering* 195: 6312–6329.
- [20] Liakopoulos, P., Karpolis, I., Giannakoglou, K. 2008. Grid enabled, hierarchical distributed metamodel-assisted evolutionary algorithms for aerodynamic shape optimization. *Future Generation Computer Systems* 24: 701–708.
- [21] Timnak, N., A. Jahangirian, and S. A. Seyyedsalehib. 2017. An optimum neural network for evolutionary aerodynamic shape design. *Scientia Iranica* 24(5): 2490-2500.
- [22] Chugh, Tinkle, et al. 2019. A survey on handling computationally expensive multiobjective optimization problems with evolutionary algorithms. *Soft Computing* 23: 1-30.
- [23] Kapsoulis, Dimitrios, et al. 2018. A PCA-assisted hybrid algorithm combining EAs and adjoint methods for CFD-based optimization. *Applied Soft Computing* 73: 520-529.
- [24] Maier, H. R., et al. 2019. Introductory overview: Optimization using evolutionary algorithms and other metaheuristics. *Environmental Modelling & Software* 114: 195-213.
- [25] Khurana, M., Winarto, H., & Sinha, A. 2009. Airfoil optimization by swarm algorithms with mutation and artificial neural networks. *AIAA Paper 2009-1278*.
- [26] Banal, Lemuel F., et al. 2017. Aircraft configuration development through surrogate-based robust optimization using a real-coded fuzzy-genetic algorithm. *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. IEEE.
- [27] Wang, Handing, Yaochu Jin, and J. Doherty. 2017. Committee-based active learning for surrogate-assisted particle swarm optimization of expensive problems. *IEEE transactions on cybernetics* 47(9): 2664-2677.
- [28] Poethke, B., S. Völker, and K. Vogeler. 2018. *Aerodynamic Optimization of Turbine Airfoils Using Multi-fidelity Surrogate Models*. International Conference on Engineering Optimization. Springer, Cham.
- [29] Kim, J. E., Rao, V. N., Koomullil, R. P., Ross, D. H., Soni, B. K., & Shih, A. M. 2009. Development of an efficient aerodynamic shape optimization framework. *Mathematics and Computers in Simulation* 79: 2373–2384.
- [30] E. Ortiz-García, S. Salcedo Sanz, Á. Pérez-Bellido and J. Portilla-Figueras. 2009. Improving the training time of support vector regression algorithms through novel hyper-parameters searchspace reductions. *Neurocomputing* 72: 3683–3691.
- [31] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2, article 27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [32] <http://tau.dlr.de/fileadmin/documents/meetings/2008/pdf/1st-day-theory/Tau-Overview.pdf>
- [33] <https://www.centaurosoft.com/grid-generator>
- [34] Abbott, I. H., and A. E. Von Doenhoff. 1959. *Theory of wing sections, including a summary of airfoil data*. Courier Corporation.
- [35] Cook, P. H., M. A. McDonald, and M. Firmin. 1979. *Wind Tunnel Measurements of the Mean Flow in the Turbulent Boundary Layer and Wake in the Region of the Trailing Edge of a Swept Wing at Subsonic Speeds*. RAE-TR-79062.
- [36] Epstein, B., et al. 2009. Comparative study of three-dimensional wing drag minimization by different optimization techniques. *J. Aircraft* 46(2): 526-541.