**EngOpt 2016 – 5rd International Conference on Engineering Optimization**
**Iguassu Falls,** Brazil, 19-23 June 2016.

# Integrated system to perform surrogate based aerodynamic optimisation for high-lift airfoil

**D. Di Pasquale†, F. Zhu‡, M. Cross*, M. Savill†, T. Kipouros†**

†Cranfield University, Cranfield, United Kingdom, {davide.dipasquale, savill, t.kipouros }@cranfield.ac.uk
*Airbus UK, Bristol, United Kingdom, murray.cross@airbus.com
‡Cranfield University, United Kingdom, feng.zhu.external@airbus.com

**1. Abstract**
This work deals with the aerodynamics optimisation of a generic two-dimensional three element high-lift configuration. Although the high-lift system is applied only during take-off and landing in the low speed phase of the flight the cost efficiency of the airplane is strongly influenced by it [1]. The ultimate goal of an aircraft high lift system design team is to define the simplest configuration which, for prescribed constraints, will meet the take-off, climb, and landing requirements usually expressed in terms of maximum L/D and/or maximum $C_L$. The ability of the calculation method to accurately predict changes in objective function value when gaps, overlaps and element deflections are varied is therefore critical. Despite advances in computer capacity, the enormous computational cost of running complex engineering simulations makes it impractical to rely exclusively on simulation for the purpose of design optimisation. To cut down the cost, surrogate models, also known as metamodels, are constructed from and then used in place of the actual simulation models. This work outlines the development of integrated systems to perform aerodynamics multi-objective optimisation for a three-element airfoil test case in high lift configuration, making use of surrogate models available in MACROS Generic Tools, which has been integrated in our design tool. Different metamodeling techniques have been compared based on multiple performance criteria. With MACROS is possible performing either optimisation of the model built with predefined training sample (GSO) or Iterative Surrogate-Based Optimization (SBO). In this first case the model is build independent from the optimisation and then use it as a black box in the optimisation process. In the second case is needed to provide the possibility to call CFD code from the optimisation process, and there is no need to build any model, it is being built internally during the optimisation process. Both approaches have been applied. A detailed analysis of the integrated design system, the methods as well as the optimisation results of the comparison between the techniques is provided.
**2. Keywords: High-Lift, Optimisation, Surrogate Model, Automation, Engineering Design**

## 3. Introduction

For an aircraft industry, in order to be competitive in today's global market, where aggressive weight targets, shortened development time scale and reduced costs are the primary objectives and constraints, a different approach for the design process is necessary to be introduced. Optimisation has become part of the design activity in many disciplines that are not only restricted to engineering. The motivation behind this inclusion is the need to produce economically relevant products or services with embedded quality. Improved production and design tools, as well as advancements in computational process technology, have assisted towards the consideration of optimisation methods in new developments and in different applications. In the industrial context, optimisation is usually associated with design and it means to identify the best solutions under certain circumstances.
The basic knowledge to understand the aerodynamics of high-lift devices is presented by A.M.O. Smith [2] in 1975. In his work titled "High-Lift Aerodynamics" he describes in detail the fundamental principles that dominate the aerodynamics of multi-element airfoils. More recently high-lift aerodynamics became a common research field for experimental [3] and numerical studies. The capabilities of numerical methods concerning the aerodynamic performance of high-lift devices are reviewed by Rumsey [4].
Most current methods for transport aircraft high lift system design rely on use of wind tunnel testing in conjunction with Computational Fluid Dynamics (CFD) analysis. These approaches do not allow varying the positions of the various elements, such as slats and flaps, in a systematic fashion.
Aircraft design, as many other engineering applications, is increasingly relying on computational power. Indeed, a strong effort has been done in the recent past to introduce potentially highly accurate analysis methods both in geometry and physics modelling. The main drawback is that they are computationally expensive. The solution of non-linear steady or unsteady aerodynamic flows by numerically solving the Navier-Stokes equations implies an amount of data storage, data handling and processor costs that may result very intensive, even when implemented on modern state-of-art computing platforms. This turns out to be an even bigger issue when used within parametric studies, automated search or optimisation loops which typically may require thousands analysis evaluations. Due to these obstacles, long running times and lack of analytic gradients, almost any optimisation method applied directly to the simulation will be slow. The techniques of optimisation in the recent years have not changed significantly but the areas in which they are applied have increased at considerable rate [5,6,7,8,9,10,11,12]. Successful use of optimisation requires prerequisites as a mathematical modelling of the design problem, knowledge of the computer software, and of the optimisation technique. The correct implementation of such an analysis hence involves the utilisation of the most advanced numerical simulation methods in a wide variety of disciplines. It is clear that from a technical point of view it represents a really challenging task. In this work will be explored how to achieve design convergence acceleration and what consequences this might have to the optimality level of the designed product, the airfoil high-lift in our case.
An adequate and general answer to optimisation based on long running and computationally intensive analysis lies in the exploitation of surrogate models. Over the last two decades, there has been an explosion in the ability of engineers to build numerical models to simulate how a complex product will perform. Moreover, the ability to quickly modify these simulation models to reflect design changes has also greatly increased and the potential for using optimisation techniques to improve engineering design is now higher

than ever before. Surrogate models are educated guesses as to what an engineering function might look like, based on a few points in space where it is possible affording to measure the function values. Recent advances in Surrogate-Based Optimisation (SBO) bring the promise of efficient global optimisation to reality [13]. A review of the state-of-the-art constructing surrogate models and their use in optimisation strategies is to be found in references [14, 15, 16]. Surrogate models may be usefully exploited through optimisation as they indeed try to provide answers in the gaps between the necessarily limited analysis runs that can be afforded with the available computing power. The basic idea is for the surrogate to act as a curve fit to the available data so that the results may be predicted without recourse to the use of the primary source, the computationally intensive simulation codes. The approach is based on the assumption that, once built, the surrogate will be many orders of magnitude faster than the primary source while still being usefully accurate when predicting away from known data points. This underlines the two key requirements of the approach: a significant speed increase in use and useful accuracy. Obviously these constitute two conflicting requirements and the compromise best suited to the application targeted will drive the choices set.

**4.High-Lift systems**

The purpose of high-lift devices comes from low-speed procedures as indicated in the FAR/CS-25 documents [17, 18]. These involve take-off, landing and go-around manoeuvres. Flaig et al. [19] explains that the unaltered, clean wing is based on optimum cruise conditions, since the larger part of the flight consists of cruising towards the destination. The aircraft high lift system designer is usually given a wing designed for cruise conditions. The maximum chord of the slat and/or flap(s) is usually dictated by the size of the wing box determined for structural and fuel capacity considerations. Therefore, very little leverage exists on the shape of these elements but more on the spacing with respect to each other (gap and overlap).

A broad range of different high-lift systems have been developed over the years, although the most widely used in civil aircraft is the multi-element wing. This configuration is typically composed of a leading-edge device that increases the stall angle of attack, and a trailing-edge device that produces an upward shift in the lift curve, see Figure 1.

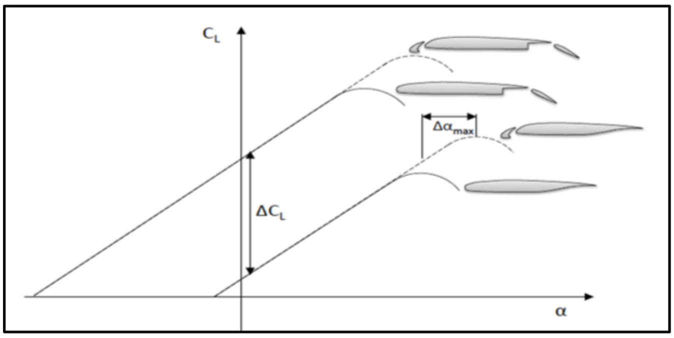



Figure 1: Effects of leading edge (increase of stall angle of attack) and trailing edge devices (shift upward) on lift curve [1]

The aerodynamics performance of a multi-element wing is strongly dependent on the interactions between the different elements. Compared with a single element airfoil, additionally complexity can be identified in the flow-field that develops around such configuration, as illustrated in Figure 2.

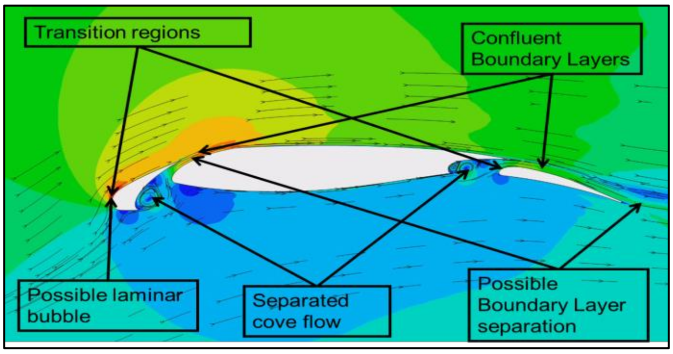



Figure 2: Visualisation of the flow-field for a multi element wing

In particular recirculation areas develop in the cove region of slat and main element, together with the mixing of the shear layers of the different elements. By changing the geometric spacing (slots) between the elements, the wakes of the corresponding boundary layers can merge, leading to a "confluent boundary layer". This deteriorates the stall characteristics of the airfoil. Therefore, the gap size must be properly balanced.

The optimisation of multi-element airfoils requires the relative position of the different elements to be varied. Therefore, a set of parameters have to be defined that uniquely defining the positioning of each element. Various solutions have been proposed in literature, although the most commonly used parameters, and in this work used, is the “gap-overlap definition”, see Figure 3.

The gap-overlap definition is related to the physical sensitivity of the flow to geometrical changes. Three variables are used for the definition of the deployment positions: “gap”, “overlap” and the “deflection angle” (θ). The gap is defined as the radius of the circle centred at the trailing edge of the preceding element and tangent to the following one. It is by definition always positive. The overlap is, as the name suggests, a measure of the elements overlapping, measured along the stowed configuration chord line. The overlap is defined by the chord-wise distance between the trailing edge point of the main element and the forward-most point on the following element. It is defined positive when the elements are overlapping, whereas a negative value indicates increased separation of the elements. Finally the deflection angle is the angle between the clean profile chord and the rotate chord fixed relative to the deployed element (slat or flap). However, a positive θ is associated to an increase in deflection angle of the elements, which correspond to a clockwise rotation for the flap element, and to a counter-clockwise rotation for the slat.

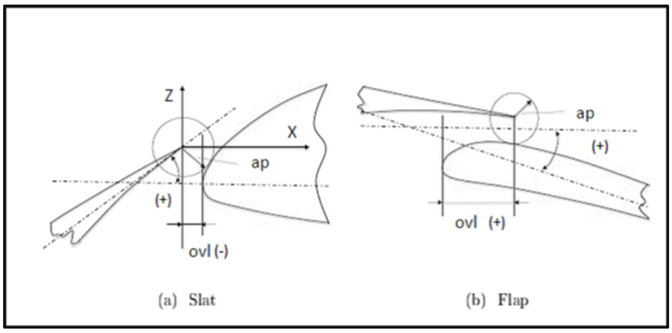



Figure 3: Gap-overlap definition for high-lift device deployments

The complexity of the underlying aerodynamics and the sensitivity of the airfoil performance to the value of gap and overlap make the determination of the optimum positions of the elements a challenging task.

**5. Optimisation brief overview**

The main aim of an optimisation process is to find an optimal geometry that fulfils the minimisation (or maximisation) of the objective functions. Therefore, the first task to be performed is to define some variables, called design variables in the optimisation process, allowing the parameterisation of the geometry. Hence, the first task to be executed is the parameterisation of the initial high-lift geometry. A second task is to define the design space where the optimisation is allowed to set the design variables values. If the design space is too narrow, the new geometry generated may not be good enough to provide any advantage upon the datum profile. Conversely, if the design space is too wide, the geometries generated could be unfeasible and problems will arise regarding other steps of the optimisation process, for instance in the CFD meshing, pre-processor and solver steps. In general, this step is one of the key bottlenecks of the production of an integrated automated optimisation process due to the difficulties of building a tool robust and flexible enough to provide a wide variety of new geometries, allowing for minimal changes, with the minimum amount of design variables.

The CFD solver represents the most computational and time demanding phase of the whole optimisation process. The level of accuracy required for the CFD solution is such that the physics of the analysed problem must be captured, in order not to mislead the optimisation process. However, it is not necessary to have a highly accurate solution if it comes at the cost of increased computational effort, either time or resources. In fact, the optimisation process will be successful as long as the correct trend in performance is identified.

The optimiser is the main module of the whole optimisation process and is crucial to the success or failure of producing an improved design. The performance of the process also relies on the ability of the optimiser to manage these types of problems. Assessing an optimisation cannot be approached without the use of “computer intelligence”, which is able to manage a high number of design variables and describe the problem, whilst evaluating the objective functions in order to improve them. In order for this assessment to be feasible, it is necessary to use an algorithm capable of managing the design variables in an efficient manner. Many optimisers have been developed throughout the years with the hope of finding a method able to solve any kind of problem. This is a weakness of the optimiser, as none of them are able to do that task perfectly. The reason for this is that the efficiency of one optimisation algorithm to solve a problem strongly depends on the nature of the problem itself. This is why algorithms are still currently under development.

**6. Methodology**

The main goal of this work is to enable deployments settings optimisation for multi-element airfoils at take-off and landing using surrogate models to accelerate the process.

The workflow of the newly developed automated and integrated surrogate optimisation approach is schematised in Figure 4.

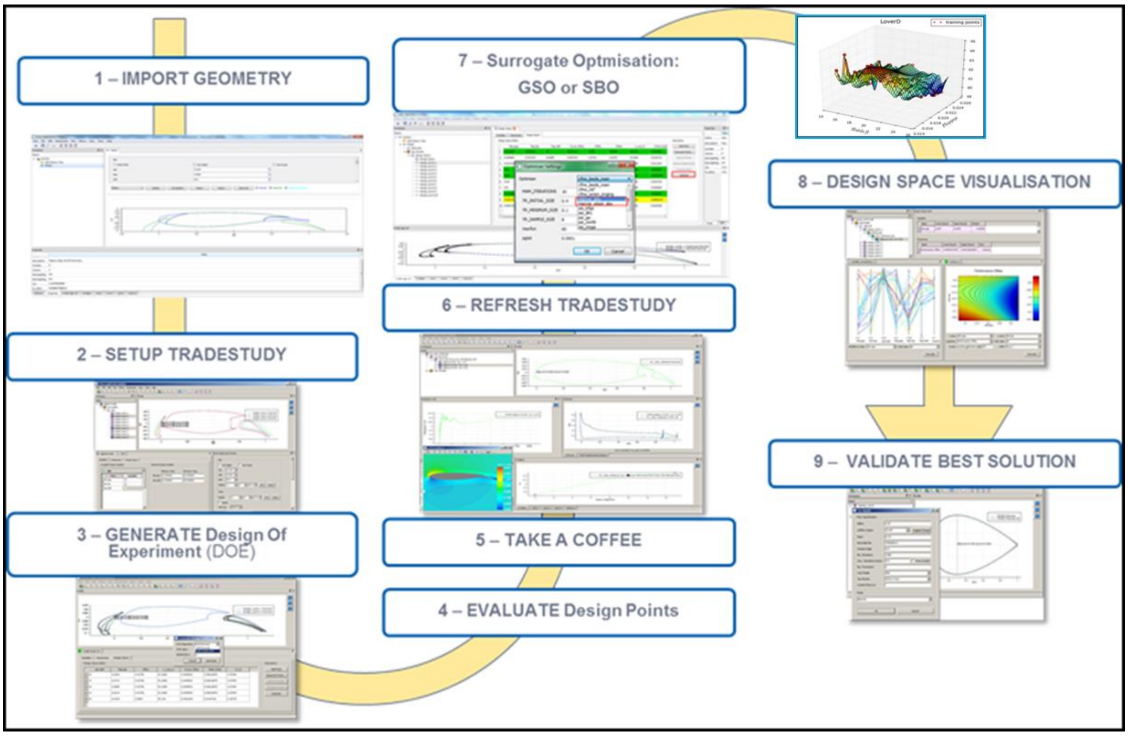



Figure 4: Schematic workflow of the integrated system to perform section HL optimisation using surrogate models

First the baseline geometry is imported in the tool, for which to facilitate the data exchange with other design tools several file formats can be read in as well as exported. Moreover, when considering multi-elements airfoils, the geometry can be imported in either the stowed or pre-deployed configuration. The next step is to set-up the trade study, specifying the flow conditions, i.e. Mach number, angle of attack range and the RANS code settings. A set of design variables will be presented to the user, who will have to choose an appropriate subset and define the range of variation for each variable. Finally the selection of a method for generating a set of different deployment settings configurations is selected. In this particular work the Latin Hypercube Sampling (LHS) as Design of Experiment (DoE) algorithm has been selected, given that this method provides good uniformity and flexibility on the size of the sample. A background process is then started, in which these input files are then transferred to a High Parallel Computing (HPC) cluster queuing system, and RANS simulations are performed for each design point, using TAU 2D (DLR RANS code), making use of the SST turbulence model, after that the meshes are automatically generated. Within the integrated system the mesh “re-generation” approach is used, meaning that a new mesh is generated for the flow field evaluation of each design identified. In addition, feasibility checks are carried out to exclude geometry intersections and low mesh quality from the process. At the completion of all the converged simulations on the HPC cluster, data can be retrieved in the local machine and optimisation process can be set-up. In order to perform single and multi-objective optimisation MACROS Generic Tool has been integrated in our design tool.

6.2 MACROS_GT

MACROS_GT is a is a high performance C++ core platform that offers a set of software tools for process integration, predictive modelling, data mining and multidisciplinary optimisation. MACROS_GT is developed by DATADVANCE, which is a joint venture company between Airbus Group and a group of private investors specialised in the development of predictive modelling and multi-objective optimisation software. It is a powerful toolkit for predictive modelling, data analysis and optimisation. It provides state-of-the-art algorithms for optimisation, approximation, dimension reduction, design of experiments, and sensitivity analysis, including both well-known and unique modern methods. It has a Python interface easy to adopt in the existing engineering development process. MACRO allows to reduce design time and cost thanks to design optimisation technology based on the synergy between data analysis and numerical optimisation. Generic Tool for Optimisation (GTOpt), which is an essential part of MACROS predictive modelling and optimisation toolbox is a software package for multi- and single-objective nonlinear optimisation. Optimisation problems of this kind arise in almost all engineering and/or scientific applications. GTOpt implements variety of modern methods, however, most of the technical details is hidden from the end user. Using MACROS is possible to perform either optimisation of the model built with predefined training sample, using the GTApprox module, which is a software package for construction of approximations fitting user-provided training data, the so called Global Surrogate Optimisation (GSO), or an Iterative Surrogate-Based Optimization (SBO).

The main difference is that in a GSO, GT Approx constructs surrogate model that is expected to be as accurate as possible in all input domain and minimizes average prediction error for the model, and clearly the solution of optimisation problems using a prebuild surrogate model, essentially depends on the training sample. In this case the model is built independently from the optimisation process, within our integrated design tool, and then used as a black box in the optimisation process.

Making use of the SBO approach is needed to provide the possibility to call the source CFD code from the optimisation process, and

there is no need to build any model, it is being built internally during the optimisation process.
During an optimisation process there is no need to have good accuracy everywhere. It is needed best precision near optimum points to locate them and in other places it is necessary a model to be accurate enough just to indicate in which regions the optimiser should dig for optimum (or in which regions it is not known the function behaviour). This logic is implemented in the SBO by means of special nested multi-resolution surrogate model inside MACROS GT Opt and corresponding adaptive DoE (Design of Experiment) process where new samples are sequentially inserted to improve the prediction accuracy, as illustrated in Figure 5.

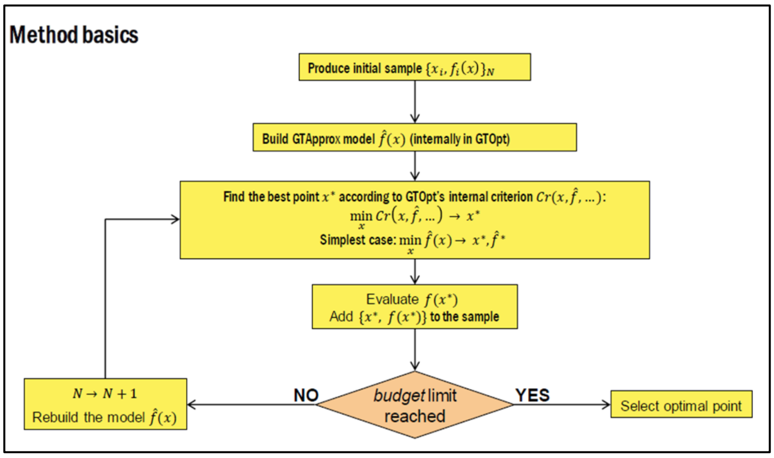



Figure 5: Scheme of the Surrogate Based Optimization (SBO) method

The goal of criterion function is to find

$$x^* = f(x^*) < min_i f_i \tag{1}$$

Using only min$x$ $(x)$ $x^*$ as criterion is not efficient, because it leads to localisation of the search procedure. The criterion has to account for model and error:

$$[\hat{f}(x), \hat{\sigma}(x)] \tag{2}$$

GT Opt assumes Gaussian process behaviour so we can estimate CDF of improving

$$\Phi\left[\frac{f^* - \hat{f}(x)}{\hat{\sigma}(x)}\right] \tag{3}$$

Where $f^*$ is the function's target value and $\Phi$ is error function.
For more information is advisable to look at the DATADVANCE technical reference documentation that can be found on their website [20].

**7. Test Case**
The high lift airfoil under investigation is a section cut of a confidential three element wing. It is shown dimensionless in Figure 6.

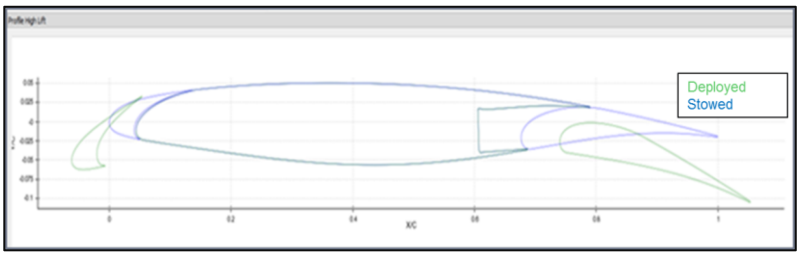



Figure 6: Airfoil geometry

The CFD analyses have been performed under the following flow condition: $M_\infty$=0.2, $Re$=6·10$^6$ for an entire polar in which the angle of attack ranges from 0 to 23 degree. A hybrid mesh approach has been selected for the automated meshing procedure. It is structured in the near wall regions, keeping a y+ < 0.5 all over the surface from the first cell away from the wall, in order to have a good resolution of the boundary layer and unstructured in the reaming parts of the domain. Furthermore, a quad-dominant mesh has been preferred to a pure triangle unstructured one. The mesh is shown in Figure 7.

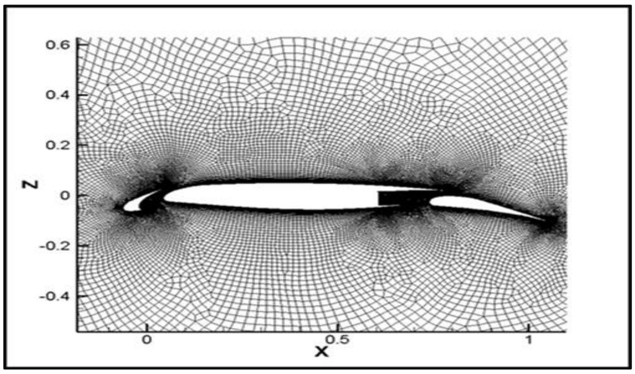



Figure 7: Mesh around the airfoil

It is made up of nearly 350000 cells, and the far-field boundary is a circle with a radius of 100 times the stowed single element chord. A pressure far-field boundary condition has been applied to the external domain and a no-slip condition at the airfoil surface.
The optimisation aims at improving the aerodynamic performance of the configuration varying the deployment settings of the lift devices. The gap, overlap and deflection angle of each element are used as design variables, adding up a total of 6 parameters. In order to define the design space, i.e. the range of variability of the design variables, many different constraints should be considered. One of the most important class of constraints is represented by kinematics used to deploy the high-lift devices. This aspect has an important influence on limiting the relative positions of slat and flap in respect of the main element. Although, in the current study these constraints are not taking into account, the design space is defined, for both slat and flap, keeping these limitations in mind. (see Table 1).

Table 1: Design variables, their datum value and range of variation

| Design Variables | Minimum Value | Datum | Maximum Value |
|---|---|---|---|
| Slat gap | 0.015 | 0.02 | 0.025 |
| Slat overlap | 0.005 | 0.01 | 0.015 |
| Slat deflection | 15.0 | 20.0 | 25.0 |
| Flap gap | 0.015 | 0.02 | 0.025 |
| Flap overlap | 0.025 | 0.05 | 0.075 |
| Flap deflection | 10.0 | 15.0 | 20.0 |

Specifically, an optimisation study has been conducted using the Global surrogate optimisation approach making use of different surrogate model technique and varying the number of train sample to build the model and tested on variable number of design parameters. For these optimisation studies two objective functions have considered. Specifically the lift over drag (L/D) and the max lift coefficient ($Cl_{max}$). Additionally, a Surrogate-Based single objective optimisation approach has been applied and compared the results with only one of the GSO approach performed.

**8. Overview of the approximation techniques**
The principle features of the five metamodeling techniques compared in this study are described in the following sections.

8.1 Response Surface Model quadratic (RMSq)
It is a kind of linear regression model with several approaches to regression coefficients estimation. RSM can be either linear or quadratic with respect to input variables. The quadratic RSM has been used in this study. It is a very robust and fast technique with a wide applicability in terms of the space dimensions and amount of the training data.

8.2 High Dimensional Approximation (HDA)
It is an advanced algorithm of division of the training set into the proper training and validating parts. A non-linear, self-organizing technique for construction of approximation using linear decomposition in functions from functional dictionary consisting of both linear and non-linear base functions. Particular example of such decomposition is so-called two-layer perceptron with nonlinear (sigmoid) activation function. However, the structure and algorithm of HDA is completely different from that of two-layer perceptron and contains the following features:

- an advanced algorithm of division of the training set into the proper training and validating parts;
- different types of base functions, including sigmoid and Gaussian;
- adaptive selection of the type and number of base functions and approximation's complexity;
- an innovative algorithm of initialization of base functions' parameters;
- an adaptive regularization algorithm for controlling the generalization ability of the approximation;

In short the HDA algorithm works as follows:

1. The training set is devised into the proper training and validating parts.

2. Functional dictionary with different types of base functions, including sigmoid and Gaussian functions, is initialized.
3. The number, type of base functions from the functional dictionary and approximation's complexity are adaptively selected. Thus a basic approximator is initialized.
4. The basic approximator is trained using an adaptive strategy controlling the type and parameters of used optimization algorithms. Specially elaborated adaptive regularization is used to control the generalization ability of the basic approximator.
5. Post-processing of the basic approximator's structure is done in order to remove redundant base functions.

For more details, see [21].

8.3 Gaussian Process (GP)

A flexible non-linear, non-parametric technique for constructing of approximation by modelling training data sample as a realization of an infinite dimensional Gaussian Distribution fully defined by a mean function and a covariance function [22, 23]. Approximation is based on a posteriori mean (for the given training data) of the considered Gaussian Process with a priori selected covariance function. GP contains the following features:

flexible functional class for modelling covariance function;

- an innovative algorithm of initialization of parameters of the covariance function;
- an adaptive strategy controlling the parameters of used optimization algorithm;
- an adaptive regularization algorithm for controlling the generalization ability of the approximation;
- Post-processing of the results to remove redundant features in the approximation.

In short the GP algorithm works as follows:

1. Parameters of the covariance function are initialized;
2. Covariance model of the data generation process is identified by maximizing likelihood of the training data.
3. Post-processing of approximator's structure is done.

See [24] for details.

8.4 High Dimensional Approximation combined with Gaussian Processes (HDAGP)

It is a flexible non-linear approximation technique, with a wide applicability in terms of space dimensions. HDAGP extends the ability of GP to deal with spatially inhomogeneous functions, functions with discontinuities. However, HDAGP approximation is the slowest method compared to HDA method and GP method.

8.5 Mixture of Approximations (MoA)

Because, any approximation algorithm has natural restrictions on training sample size due to limited computational resources

Suppose training sample size is huge and there is no enough memory to handle it. Possible solution is to use sub-sample. However, this solution is bad because it does not use all given information. A better solution is to decompose the design space into sub-regions, extract from the initial sample the corresponding sub-samples, for each sub-sample construct local approximation and then "glue" these approximations together into one Surrogate Model, as shown in Figure 8 This is the idea behind this method, which is to decompose the design space into sub-domains and construct in each of them valid approximation. It could be useful for discontinuous function.

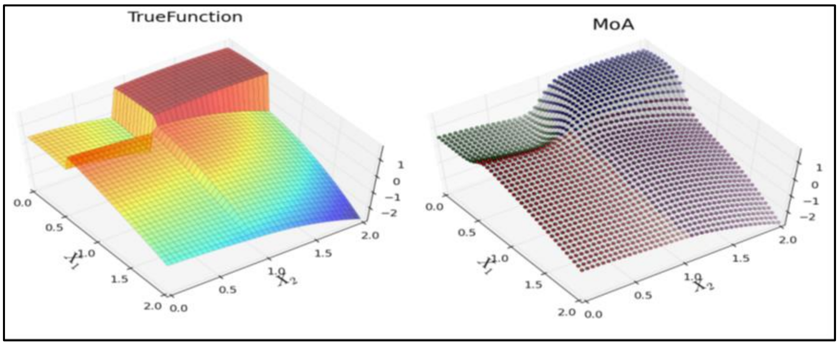



Figure 8: Example of surrogate model built with the MoA technique

**9. Performance Metrics**

There are various commonly used performance metrics for approximation models that are given in Ref. [25]. In the engineering design, the cross-validation method is currently a popular method for model validation Cross-validation is actually a measurement for degrees of insensitivity of a metamodel to lost information at its data points. Cross-validation is a well-known and well-established way of statistical assessment of the algorithm's efficiency on the given training set, but it should be stressed, however, that it does not directly estimate the predictive power of the main approximation $\hat{f}$ outside the training set S. Rather, the purpose of cross-validation is to assess the efficiency of the approximation algorithm on various subsets of the available data, assuming that the conclusions can be extended to the (unavailable) observations from the total design space.

Suppose that a training set ($X_N$ ; $Y_N$) represents an unknown response function $Y = f(X)$, and $\hat{f}$ is an approximation of $f$ constructed from this training data. An important property of the approximation is its predictive power $Q^2$ which is understood as the agreement between $f$ and $\hat{f}$ on points X not belonging to the training set ($X_j$ ; $Y_j$). $Q^2$ is a statistical measure of how close the data are to the fitted regression line. The larger the value of $Q^2$ more accurate is the metamodel.

$$Q^2 = 1 - \sum_{j=1}^{m} (Y_j - \hat{f}(X_j))^2 \Big/ \sum_{j=1}^{m} (Y_j - \bar{Y})^2 \qquad (4)$$

A closely related concept is accuracy of the approximation, which is understood as the deviation $|\hat{f} - f|$ on the design space of interest. Accuracy is often measured statistically, e.g., standard accuracy measures are the Mean Absolute Error (MAE) and root-mean-squared error (RMS). The MAE here is the mean of the half-normal distribution (i.e., the average of the positive subset of a population of normally distributed errors with zero mean), which provides an understanding of the maximum local deviation of the model estimation from the actual output.

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{f}(X_i) - \bar{f}(X_i) \right| \qquad (5)$$

The RMSE is a quadratic scoring rule, which measures the average magnitude of the error, which is a global error measure and provides an understanding of the model accuracy over the entire design domain.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left| \hat{f}(X_i) - \bar{f}(X_i) \right|^2} \qquad (6)$$

Both the MAE and RMSE can range from 0 to ∞. They are negatively oriented scores: lower values are better.

**10. Results**

Several optimisation problems on same test case have been performed varying the number of design variables and training set size (100, 250, 500) for model construction, making use of five different techniques, but due to page constraint the results of a GSO multi-objective optimisation problem using 6 DVs with surrogate trained with 250 design points making use of HDAGP, which is the one that predicted best results has shown. Figure 9 shows the model accuracy comparison in terms of MAE and RMSE and $Q^2$

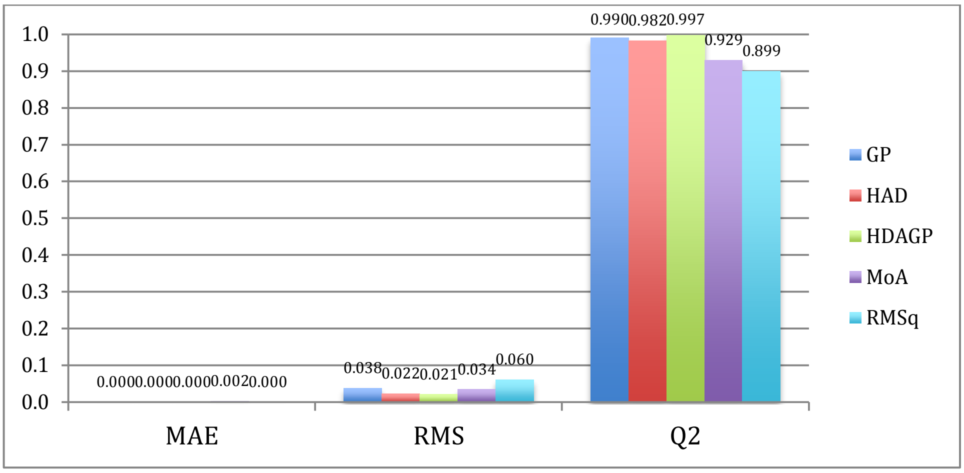



Figure 9: Model accuracy comparison.

Figure 10 shows a scatter plot for the built HDAGP surrogate obtained using the validation set. In this case predictions are close to the true values, so the obtained model seems to be accurate enough.

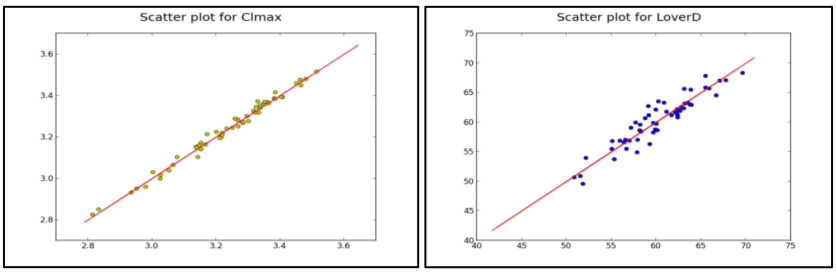



Figure 10: Scatter plots obtained with a separate test sample

The result of the optimisation in term of Pareto front is shown in Figure 11.

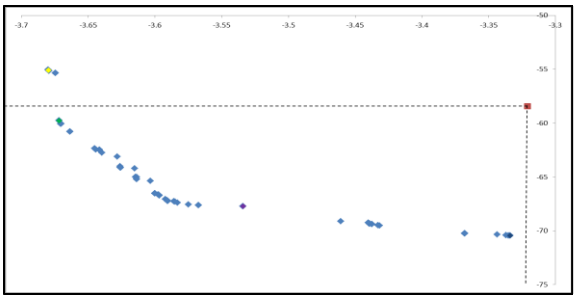


Figure 11: Pareto front

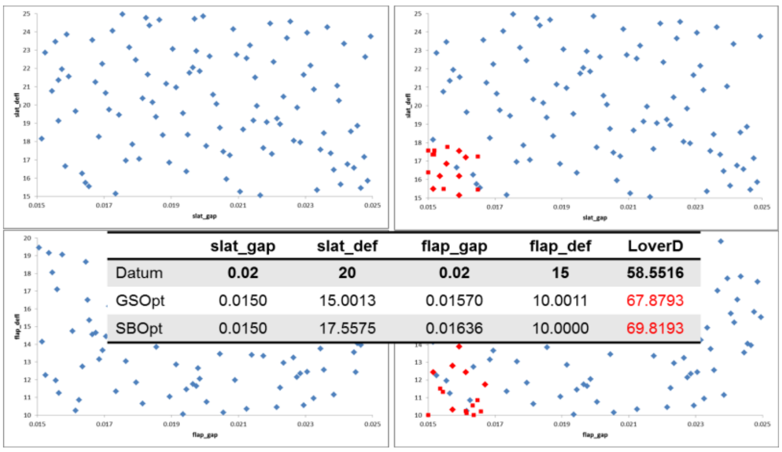
The Pareto front obtained has shown a substantial improvement of the objective function values. Table 2 show the validation of some points belonging to the Pareto front against CFD analysis and it possible to see that the results are generally in good agreement, although the optimisation results using the surrogate slightly over predict both the objective functions.

Table 2: Design variables, their datum value and range of variation

| Design Variables | L/D (SM) | L/D (CFD) | Cl (SM) | Cl (CFD) |
|---|---|---|---|---|
| Datum | 58.5516 | 57.4378 | 3.3366 | 3.3375 |
| Max Cl | 62.8399 | 61.9875 | 3.6798 | 3.6548 |
| Compromise | 67.7175 | 65.9148 | 3.5339 | 3.5287 |
| Max L/D | 70.4384 | 68.8793 | 3.3439 | 3.3245 |
| Maximum Cl | 55.7703 | 54.8762 | 3.6717 | 3.6571 |

During this exercise, when using 100 or 500 design points to train the surrogate no better results have been found. In the first case is most probably due to the fact that they are not sufficient for the generation of an enough accurate model due to the non-linearity of the physics behind. In the second case as opposite the detrimental of results is due to of an overfitting or overtraining phenomenon, which consist in getting the approximation to be very accurate on the training set, at the cost of excessively increasing approximation's complexity which leads to a less robust behaviour and ultimately lower accuracy on points not belonging to the training set, especially when we are dealing with noisy data.

A comparison between the two different optimisation approaches, GSO vs. SBO has been performed. In this case just 4DVs are considered and a single objective taken into account. Starting with a model built on 90 points 15 new points have been automatically generated to improve the Surrogate. Looking at the design space, see Figure 12, is clear that these points are added in a particular sub-region where the optimum is probably to be and clearly looking at the LoverD results there is an improvement of the Objective function.



| | slat_gap | slat_def | flap_gap | flap_def | LoverD |
|---|---|---|---|---|---|
| Datum | 0.02 | 20 | 0.02 | 15 | 58.5516 |
| GSOpt | 0.0150 | 15.0013 | 0.01570 | 10.0011 | 67.8793 |
| SBOpt | 0.0150 | 17.5575 | 0.01636 | 10.0000 | 69.8193 |

Figure 12: Comparison between the two different optimisation approaches, GSO vs. SBO

**11. Conclusion**

An integrated optimisation approaches have been developed and has demonstrated to be seamless and robust. The surrogate optimisation architecture is preferred to a RANS-in-the-loop one. Firstly, the optimisation process is de-coupled from the RANS execution, reducing the risk of failure during the process. At the same time the surrogate optimisation approach presents also some drawbacks. The model is not guaranteed to be accurate over the whole design space, especially if the problem tackled is highly non-linear. Besides, the number of samples required for constructing an accurate surrogate model increases rapidly with the number of design variables. Clearly, the choice of the optimisation architecture has a significant influence on both the solution time and the final design. As for the different approximation techniques:

**RMS** is usually rather crude, and the approximation can hardly be significantly improved by adding new training data. In addition, the number of regression terms can increase rapidly with increasing of dimension, if quadratic RSM is used. An advantage of RSM is that it can smooth out the various scales of numerical noise in the data while captures the global trend of the variation.

**HDA** is a flexible non-linear approximation technique, with a wide applicability in terms of space dimensions. HDA can be used with large training sets (more than 1000 points). However, HDA is not well-suited for the case of very small sample sizes. Significant training time is often necessary to obtain accurate approximation for large training samples. HDA can be applied to noisy data.

**GP** usually demonstrates accurate behaviour in the case of small and medium sample sizes. GP is not well-suited for modelling spatially inhomogeneous functions, functions with discontinuities. GP is a resource-intensive method in terms of ram capacity, therefore large samples are not supported, although large dimension are supported.

**HDAGP** is a resource-intensive method in terms of ram capacity; therefore large samples are not supported. It typically provides less smooth approximation, but it has given the best results for almost all cases investigated.

**MoA** has not proved its benefits, but it is due to the nature of the objective functions considered that are not so much discontinuous.

Models from GT Approx and internal models from GT Opt solve different problems. That is why it is advisable to use GTOpt SBO-optimisation with direct calling of CFD code for optimisation purpose, and surrogate model constructed with GT Approx to study function behaviour or make predictions over all input domain.

To conclude, automated design process is very attractive for commercial aircraft industry as it greatly reduces the development period, and the optimisation approach described could be used as a low-cost, low-order approximation for aerodynamic shape optimisation in an industrial context, given its ability to produce good results in a limited amount of time.

## 12. References


1. C P Van Dam, The aerodynamic design of multi-element high-lift systems for transport airplanes, Progress in Aerospace Sciences, Vol. 38, No. 2, 2002, 101–144.
2. A M O Smith, High-Lift Aerodynamics, Journal of Aircraft, Vol. 12, No. 6, 1975, 501–530.
3. D Reckzeh, and H Hansen, High Reynolds-number wind tunnel testing for the design of Airbus high-lift wings, New Results in Numerical and Experimental Fluid Mechanics V, edited by H Rath, C Holze, H J Heinemann, R Henke, and H Honlinger, Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Springer, 2006.
4. C L Rumsey and S X Ying, Prediction of high-lift: review of present CFD capability, Progress in Aerospace Sciences, Vol. 38, No. 2, 2002, 145–180.
5. R P Henderson, J R R A Martins, and R E Perez, Aircraft Conceptual Design for Optimal Environmental Performance, The Aeronautical Journal, Vol. 116, No. 1175, 2012,1–22.
6. J J Alonso, and M R Colonno, Multidisciplinary Optimization with Applications to Sonic-Boom Minimization, Annual Review of Fluid Mechanics, Vol. 44, No. 1, 2012, 505–526.
7. P Geyer, Component-Oriented Decomposition for Multidisciplinary Design Optimization in Building Design, Advanced Engineering Informatics, Vol. 23, No. 1, 2009, 12–31.
8. R Enblom, Two-Level Numerical Optimization of Ride Comfort in Railway Vehicles, Journal of Rail and Rapid Transit, Vol. 220, No. 1, 2006, 1–11.
9. B Potsaid, Y Bellouard, and J T-Y Wen, A Multidisciplinary Design and Optimization Methodology for the Adaptive Scanning Optical Microscope (ASOM), Proceedings of the SPIE, Vol. 6289, 2006, 62890L1–62890L12.
10. B Glaz, P P Friedmann, and L Liu, Helicopter Vibration Reduction Throughout the Entire Flight Envelope Using Surrogate-Based Optimization, Journal of the American Helicopter Society, Vol. 54, No. 1, 2009, 12007-1–12007-15.
11. G Cai, J Fang, Y Zheng, X Tong, J Chen, and J Wang, Optimization of System Parameters for Liquid Rocket Engines with Gas-Generator Cycles, Journal of Propulsion and Power, Vol. 26, No. 1, 2010,113–119.
12. J T Hwang, D Y Lee, J Cutler, J., and J R R A Martins, Large- Scale MDO of a Small Satellite Using a Novel Framework for the Solution of Coupled Systems and Their Derivatives, Proceedings of the 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Boston, MA, April 2013.
13. I L Cenaero, Complements on Surrogate Based Optimization for Engineering Design, RTO- EN-AVT-167 5 – 167.
14. J Forrester and A J Keane, Recent advances in surrogate-based optimization, Progress in Aerospace Sciences, Volume 45, Issues 1-3, January-April, 2009, 50-79.
15. N V Queipo, R T Haftka, W Shyy, T Goel, R Vaidyanathan, P K Tucker, Surrogate-based analysis and optimization, Progress in Aerospace Sciences, 2005 Volume 41, Issue 1, 1-28.
16. A Forrester, A Sobester, A Keane, Engineering Design via Surrogate Modelling A Practical Guide, John Wiley & Sons Ltd, 2008.
17. Anon., FAA Federal Aviation Regulations (FAR-25), 2013.
18. Anon., EASA Certification Specifications (CS-25), 2013.
19. A Flaig and R Hilbig, High-lift design for large civil aircraft, in AGARD-CP-515, 1993.
20. https://www.datadvance.net/product/macros/documentation.html
21. M Belyaev and E Burnaev, Approximation of a multidimensional dependency based on a linear expansion in a dictionary of parametric functions. Informatics and its Applications, 7(3), 2013.
22. N A C Cressie, Statistics for Spatial Data. Wiley, 1993.
23. C E Rasmussen and C K I Williams. Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning). The MIT Press, 2005.
24. E Burnaev, A Zaytsev, M Panov, P Prihodko, and Y Yanovich, Modeling of non-stationary covariance function of gaussian process using decomposition in dictionary of nonlinear functions. In Proceedings of the conference Information Technology and Systems, Gelendzhik, Russia 2011, 357-362.
25. S. Shan and G.G. Wang, Survey of Modeling and Optimization Strategies to Solve High-Dimensional Design Problems With Computationally-Expensive Black-Box Functions, Struct. Multidiscip. Optim, 41(2), 2010, pp. 219–241.