Implementation of an Automatic Multi-fidelity Scheme for Industrial Applications in the Cassidian SimServer Environment

Kurt Sermeus, Luca Nardin, and Kaare A. Sørensen

Cassidian, 85077 Manching, Germany

Abstract. A multi-fidelity simulation approach for the generation of aerodynamic data sets by computational fluid dynamics methods is discussed. Using Kriging-based surrogate modelling methods, data obtained a priori from a computationally less expensive low-fidelity model are combined with a fewer number of high-fidelity simulation data. An adaptive sampling method is used to iteratively select the high-fidelity data points in the parameter space and update the surrogate model. A description of the multi-fidelity simulation scheme is given and initial evaluations of the method for generating aerodynamic data on an aircraft configuration test case are discussed.

Keywords: Surrogate modelling, Kriging, multi-fidelity simulation.

1 Introduction

The generation of aerodynamic data sets, describing the aerodynamic forces and moments for determining the stability and control characteristics of an aircraft, requires significant amounts of wind tunnel testing. Due to model scaling and wind tunnel interference effects, flight test corrections to these data may be required for improving flight control laws. Handbook methods based on linear aerodynamic theories and empirical relations are commonly used in conceptual design, but are of very limited utility when confronted with the highly non-linear aerodynamic effects towards the edge of the flight envelope and with novel aircraft configurations. Computational fluid dynamics (CFD) methods, based on high-fidelity Reynolds-averaged Navier-Stokes (RANS) modelling and helped by tremendous advances in computer hardware performance, have shown an increasing capability for modelling the aerodynamics in the complete flight envelope, see e.g. [2]. The usage of CFD for generating aerodynamic data sets is however today still very limited by the vast number of data points that need to be computed and by the high computing costs per data point. Even when only a limited region of the flight envelope needs to be modelled, e.g. as part of a design process, the direct usage of high-fidelity CFD methods for generating aerodynamic data sets is prohibitively expensive. There are however also simplified simulation methods available, e.g. Euler or panel methods instead of Navier-Stokes methods, which model the same problem with a lower fidelity, but also have a significantly

N. Kroll et al. (eds.), *Computational Flight Testing*, NNFM 123, pp. 167–181.

Fluid Mechanics and Multidisciplinary Design 123,
DOI: 10.1007/978-3-642-38877-4_12,

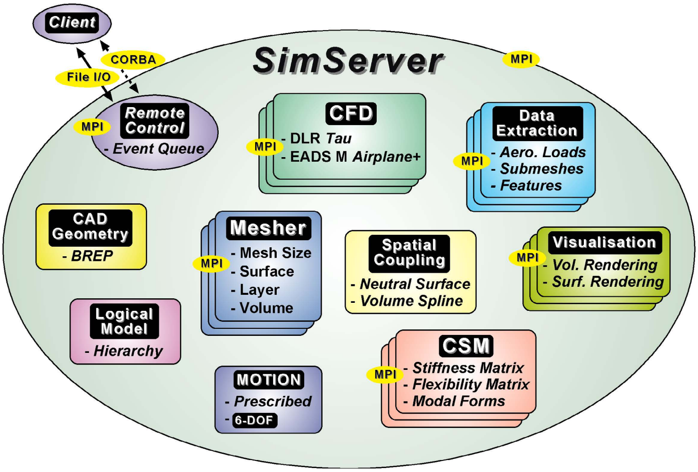


Fig. 1. Modular architecture of Cassidian *SimServer* simulation environment [8]

lower computational cost. Assuming that these lower fidelity methods model at least the global trend of the aerodynamic data in the parameter space somehow correctly, they can serve for guiding the interpolation between the sparse high-fidelity points. By a suitable combination based on Kriging interpolation, a multi-fidelity response surface can be constructed from few high-fidelity data points and a larger number of low-fidelity data points, so that given accuracy requirements are still met, while the overall computational cost is significantly reduced [2,3].

2 Multi-fidelity Simulation with Adaptive Sampling

A multi-fidelity simulation capability has been implemented in the Cassidian *SimServer* software suite through a coupling with the DLR developed *SMART Toolbox* [4], a software library providing advanced numerical methods for the generation of surrogate models based on simulation data of variable fidelity levels [2,5,6]. *SimServer* in itself is a simulation steering and ‘middleware’ software, which directly couples different components of the simulation tool chain within one completely parallelized environment [7,8]. Important components are the DLR *Tau* flow solver [2,9], Cassidian-proprietary preprocessing, Chimera grid hole cutting and six degree of freedom flight simulation methods [10,11]. Figure 1 sketches the modular architecture of the *SimServer* environment.

The *SMART Toolbox* [4] has been coupled as external library to *SimServer* in order to generate multi-fidelity surrogate models and corresponding error estimates of these models. The approach of this implementation is as follows. First a low-fidelity data set, which in principle covers the parameter range of interest with a relatively fine resolution of data points, is read by *SimServer* from a data

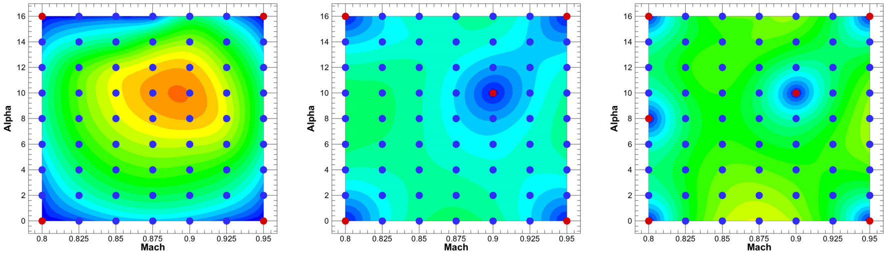
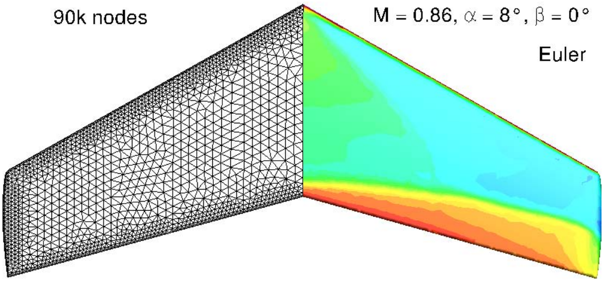


Fig. 2. Adaptive sampling in the (M, α) parameter space. The low-fidelity data (blue dots) have been previously computed in all defined grid points (full factorial). High-fidelity data (red dots) are iteratively selected in that grid point where the estimated model error (MSE) of the multi-fidelity surrogate model, shown by the colour contours, is the highest.

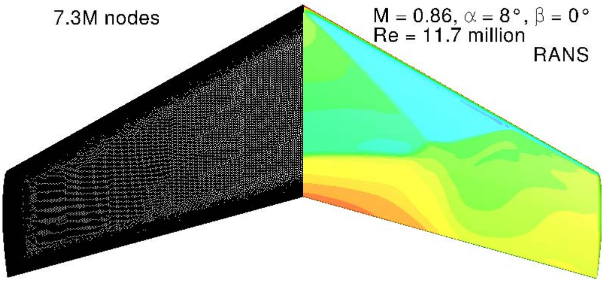
file and stored internally. Then a few initial high-fidelity computations are performed at pre-selected data points in the parameters space. From the low-fidelity and high-fidelity data points and using the *SMART Toolbox*, *SimServer* generates a Kriging-based multi-fidelity surrogate model for the complete parameter space as well as an error estimate for this surrogate model.

Two different multi-fidelity methods from the *SMART Toolbox* have been evaluated. The first is a correction-based method using an additive bridge function, which models the difference between the low-fidelity and the high-fidelity data using Kriging interpolation in the parameter space [6]. The second is the more recent hierarchical Kriging method [2], in which the low-fidelity data are used directly as trend model for the Kriging interpolation of the high-fidelity data. Both methods offer a mean squared error (MSE) estimation function for the surrogate model. The main advantage of the hierachical Kriging method is that in principle it more reliably approximates the real error than is the case with the correction-based Kriging method with bridge function. Except when indicated otherwise, all results reported herein have been obtained with the hierarchical Kriging method, using the 'thin plate spline' radial basis function and a linear regression function (universal Kriging).

The multi-fidelity error estimation is applied iteratively in order to select further data points in the parameter space (e.g. certain values of angle of attack, side slip angle, Mach number and control surface deflections), where the next high-fidelity CFD simulation is performed. After every simulation the surrogate model is updated using the last high-fidelity data and the next simulation point is selected. This adaptive sampling of the parameter space and updating of the surrogate model continues until the estimated error of the multi-fidelity surrogate model has fallen below a given threshold or until a given computational budget (CPU hours) is exhausted. The adaptive sampling of the high-fidelity model is illustrated in figure 2 for a simple example with two parameters (Mach number M and angle of attack α). Initially, low-fidelity data are computed in all points of a 7×9 grid, shown by the blue dots. High-fidelity data for the initialization of



(a) low-fidelity model: Euler, 5 minutes per data point



(b) high-fidelity model: Navier-Stokes, hours per data point

Fig. 3. Euler and Navier-Stokes simulation for the ONERA M6 wing at $M = 0.86$, $\alpha = 8°$, Reynolds number 11.7 million. Surface mesh and pressure distribution.

the surrogate model are computed in only a few points, shown by the red dots, e.g. at the corner points of the parameter space. Further high-fidelity points are then iteratively selected by the adaptive sampling method at the grid point where the MSE estimate of the model error is the largest.

The multi-fidelity implementation in *SimServer* is generic in the sense that low-fidelity data can be obtained from any simulation tool, e.g. a panel method. For aerodynamic applications of interest to Cassidian, it is anticipated that Euler simulations or Navier-Stokes simulations on coarse grids provide more effective low-fidelity models than panel methods.

Figure 3 shows the application of multi-fidelity simulation with *SimServer* for a simple example of the ONERA M6 wing, which has been computed using Euler (low-fidelity) and Navier-Stokes (high-fidelity) modelling. For the Euler simulation, a coarse grid with 90 thousand nodes was used. The Navier-Stokes simulation with the Spalart-Allmaras turbulence model used a fine grid with 7.3 million nodes. The aerodynamic data set, e.g. for the lift coefficient C_L in a given range of Mach number and angle of attack, is sufficiently well approximated with just a few Navier-Stokes high-fidelity simulations. Figure 4 shows the convergence of the estimated error and of the actual error of the surrogate model to the high-fidelity data, which have been computed for sake of comparison in all data points

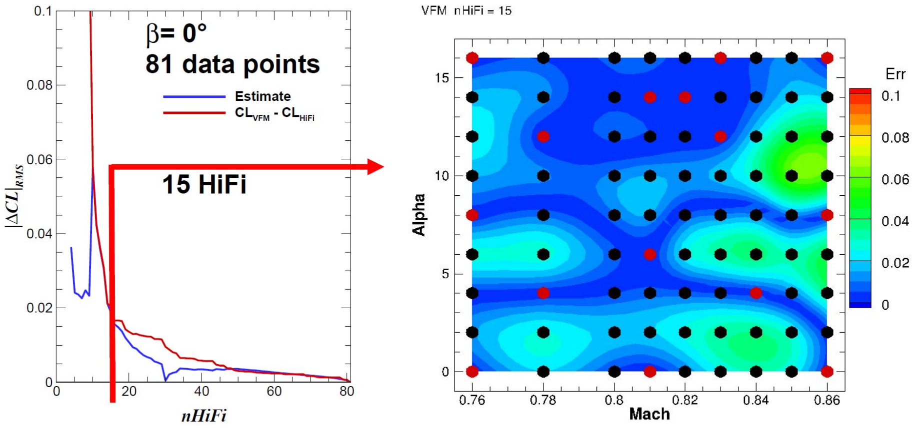


Fig. 4. Construction of (M, α) data set for the ONERA M6 wing using Euler / Navier-Stokes multi-fidelity simulation. Left: convergence of the actual error of the surrogate model with the number of high-fidelity points. Right: sampling in the (M, α) parameter space with 15 high-fidelity points and the corresponding MSE error estimate.

of the selected (M, α) range of the parameter space. For the initialization of the adaptive sampling, high-fidelity data were computed in the 4 corner points of the (M, α) range. After 11 iterations, i.e. with a total of 15 high-fidelity simulations on the grid of 81 data points, the error of the multi-fidelity surrogate model for C_L has dropped to a root mean square (RMS) average value of about 0.02. With a computing time for the *Tau* flow solver on 64 processors of 5 minutes per data point for the low-fidelity model and 5 hours for the high-fidelity model, the total computing time used by the multi-fidelity method is 82 hours. The high-fidelity simulation on all 81 data points of the (M, α) grid would require about 400 hours of computing. The cost savings obtained through the multi-fidelity method thus amount to about a factor 5.

3 Application to a UAV Aerodynamic Data Set

The multi-fidelity simulation method with *SimServer*, described in the previous section, has been tested on a generic Unmanned Aerial Vehicle (UAV) configuration, which is illustrated in figure 5. The aim of the test case is the computation of the aerodynamic data set in the following range of flight conditions:

- Mach number 0.80 to 0.95, with an interval of 0.025,
- angle of attack 0° to 26°, with an interval of 2°,
- zero side slip angle,
- Reynolds number of 45 million.

This problem has been computed with the DLR *Tau* code using three levels of modelling: Euler, Navier-Stokes on a coarse grid and Navier-Stokes on a fine grid.

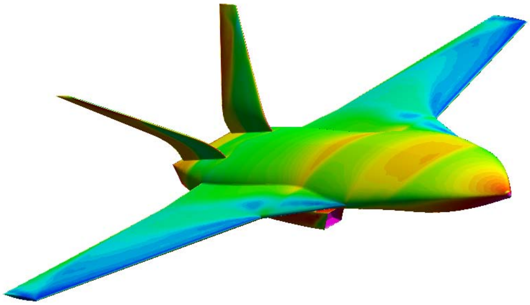


Fig. 5. Generic UAV configuration as test case for the adaptive sampling method with *Simserver* and the *SMART Toolbox*. Pressure distribution at Mach number 0.95 and 15° angle of attack; Navier-Stokes simulation on coarse grid.

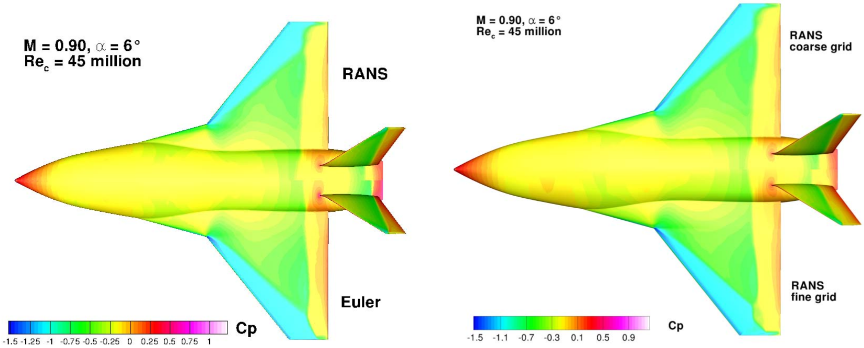
Table 1. Generic UAV test case. Grid size on half-model geometry and corresponding computing time (wall clock) with *Tau* code per data point.

Grid	Number of nodes	Computing time per data point, wall-clock time with 128 CPUs
Euler	1.6 million	0.25 hours
Navier-Stokes, coarse	3.0 million	1.75 hours
Navier-Stokes, fine	9.8 million	7.3 hours

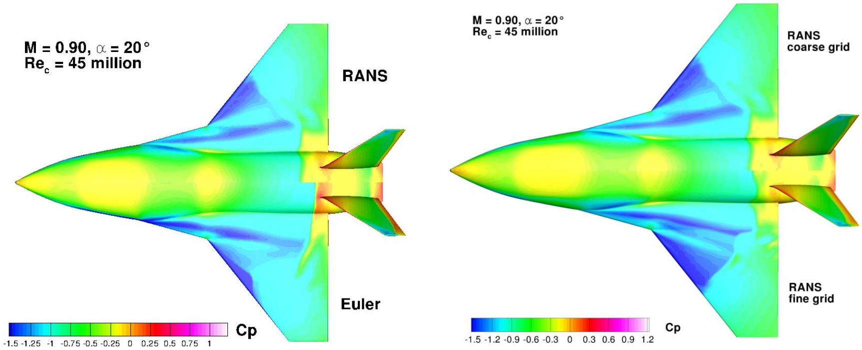
The size of the grids used for each of these simulations is listed in Table 1. The Navier-Stokes simulations have been performed using the Spalart-Allmaras turbulence model. The engine boundary conditions of the UAV model have been left at a constant setting for this exercise and correspond to cruise flight conditions.

Figure 6 shows the pressure distribution, computed using the three levels of modelling fidelity, for the data points at Mach number 0.90 and two different angles of attack, $\alpha = 6°$ and $\alpha = 20°$. While at the low-α case the pressure distributions are closely similar for the three models, larger differences can be seen for the high-α case. Especially the vortex structure on the wing is clearly different between the Euler and the Navier-Stokes simulations. The Euler simulation also shows a stronger shock at the body-tail intersection. These differences come of course as no surprise and it is to be expected that, especially at high-α conditions, Euler simulation provides a less accurate low-fidelity model than coarse grid Navier-Stokes simulation.

The construction of $C_L(M, \alpha)$ surrogate models has been evaluated using different adaptive sampling approaches. A first example, where the coarse grid Navier-Stokes simulation serves as high-fidelity model and the Euler simulation as low-fidelity model, is shown in Fig. 7. The convergence of the multi-fidelity surrogate model as a function of the number of high-fidelity data points is shown in Fig. 8. It is seen that with 20 high-fidelity data points, the average error $|\Delta C_L|_{RMS}$ of the surrogate model to the complete high-fidelity data, which for



(a) Mach number 0.90, angle of attack 6°.



(b) Mach number 0.90, angle of attack 20°.

Fig. 6. Comparison of pressure distributions computed with models of different fidelity. Left: coarse grid Navier-Stokes versus Euler; right: coarse versus fine grid Navier-Stokes.

comparison have been computed on all points of the (M, α) grid, drops below a threshold of 0.03. The four corner points of the (M, α) range are employed as fixed high-fidelity points.

4 Evaluation of Different Multi-fidelity Approaches

In the following sections, the influence of various elements in the adaptive sampling method on the error convergence of the multi-fidelity surrogate model is further examined. The test case is still the construction of the $C_L(M, \alpha)$ data set for the generic UAV, shown in Fig. 5, in the previously mentioned range of flight conditions. The high-fidelity model corresponds in this case to the fine grid Navier-Stokes simulation. The low-fidelity model is either the coarse grid Navier-Stokes simulation or the Euler simulation. The grid sizes and computing times per datTremel2005a point for each of the models are given in Table 1.

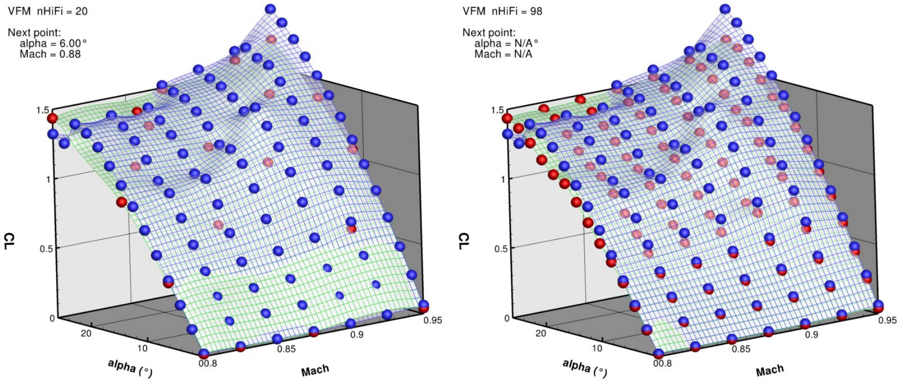
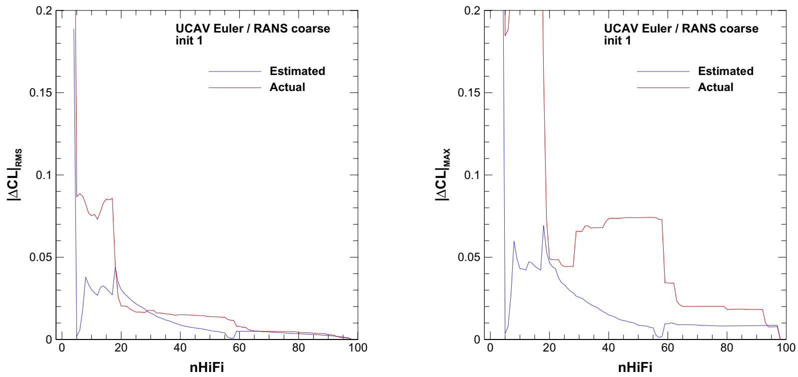


Fig. 7. Surrogate model for $C_L(M, \alpha)$ on generic UAV test case, based on Navier-Stokes simulation (coarse grid) as high-fidelity model and Euler simulation as low-fidelity model. The blue surface is the Kriging interpolation of the low-fidelity data points, indicated by the blue dots. The green surface is the multi-fidelity surrogate model, interpolating the high-fidelity data (red dots) and following the low-fidelity data. Left: 20 high-fidelity data points; right: high-fidelity simulation in all data points.

4.1 Euler Versus Navier-Stokes as Low-Fidelity Model

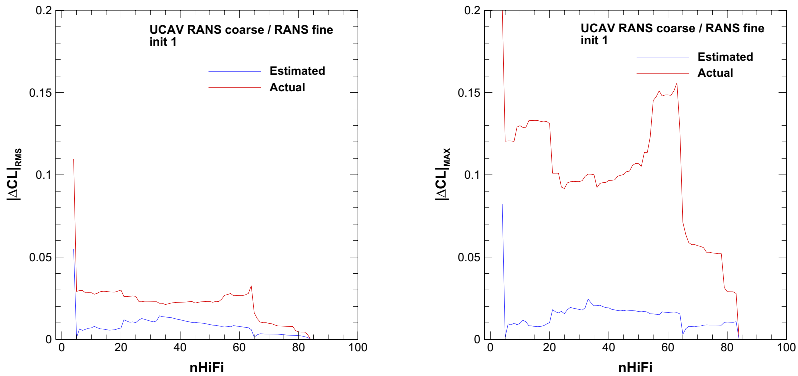
A key question from the practical point of view is which low-fidelity model leads overall to the most effective multi-fidelity strategy for generating an aerodynamic data set of given accuracy level. As the low-fidelity data points supposedly require a much lower computational effort than the high-fidelity data, it is assumed that the low-fidelity data are computed in each grid point within a complete range of the parameter space.

The convergence of the multi-fidelity surrogate model with the number of high-fidelity data points is shown in Fig. 9 and Fig. 10 for two approaches, using respectively Navier-Stokes and Euler simulation as low-fidelity model. In the first approach, the average error $|\Delta C_L|_{RMS}$ of the surrogate model with respect to the high-fidelity data, which for sake of comparison have been previously computed in all data points, drops after just 6 adaptive iterations, i.e. with in total 10 high-fidelity points, under a treshold of 0.03. In the second approach, using Euler simulation as low-fidelity model, about 25 high-fidelity simulations are required before the average error drops below the same threshold. The marked difference in convergence behaviour is attributed to the accuracy of the low-fidelity model. This can be understood from Fig. 11, which shows for the approach using Euler simulation as low-fidelity model the estimated as well as actual error of the surrogate model to the high-fidelity data. After adding the 26^{th} high-fidelity data point, the actual error of the surrogate model has its maximum around Mach number 0.82 and alpha 22°, while the estimated error has its maximum



(a) Average error. (b) Maximum error.

Fig. 8. $C_L(M, \alpha)$ surrogate model for UAV test case, based on Navier-Stokes simulation (coarse grid) as high-fidelity model and Euler simulation as low-fidelity model. Convergence of the model error with the number of high-fidelity points. Red: actual error of surrogate model to high-fidelity data; blue: estimated error (MSE) of the hierarchical Kriging interpolation.

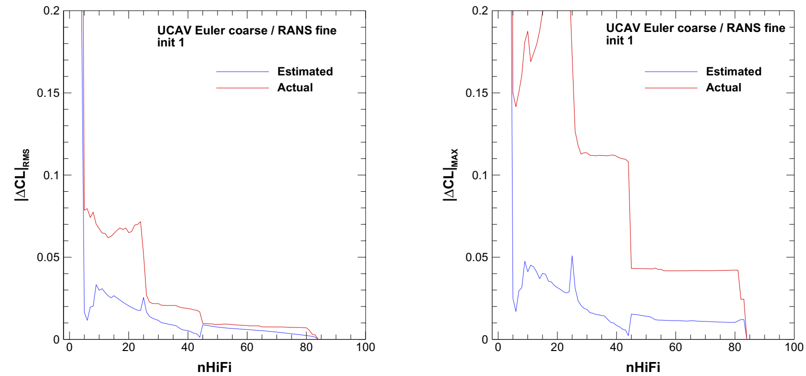


(a) Average error. (b) Maximum error.

Fig. 9. Convergence of $C_L(M, \alpha)$ surrogate model for UAV test case. High-fidelity model: fine grid Navier-Stokes; low-fidelity model: coarse grid Navier-Stokes.

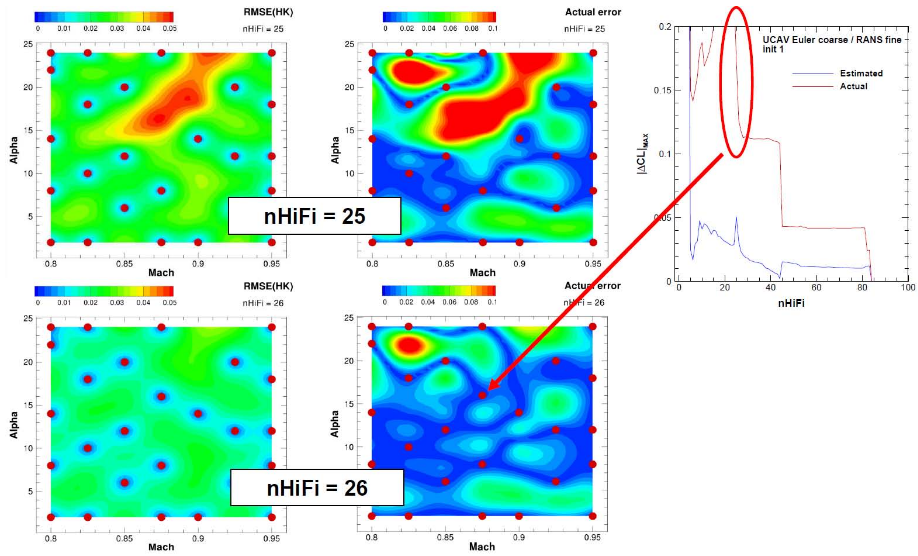
rather at a higher Mach number value. Because the estimated error differs so significantly from the actual error of the surrogate model, it takes a further 20 adaptive sampling iterations, before with the 45^{th} high-fidelity point the apparently important data point at Mach number 0.825 and alpha 22° is selected, which in one step reduces the maximum error of the surrogate model by half. This example shows that the convergence of the surrogate model can be dominated by just a few critical points in the parameter space. The underlying reason for this can be observed in Fig. 11b, where the comparison of the $C_L(M, \alpha)$ surfaces of the multi-fidelity surrogate model and of the low-fidelity Kriging interpolation shows, that precisely in this critical area of the parameter space the trend in the low-fidelity model is not representative of the high-fidelity model.

Figure 12 shows the convergence of the $C_L(M, \alpha)$ surrogate model as a function of the number of high-fidelity points as well as a function of the total computational cost, which is defined by the computing time of all low-fidelity points plus the computing time of the high-fidelity points up to the current sampling point. Although the multi-fidelity approach using Navier-Stokes simulation as low-fidelity model reaches with fewer high-fidelity points a comparable error level than the approach with Euler simulation as low-fidelity model, the computational time (budget) spent a priori in generating the low-fidelity data points is also much higher with the Navier-Stokes as with the Euler low-fidelity model. When the accuracy requirements for the surrogate model are relatively low, it is clear that the more simple and cheaper Euler simulation tends to be the better choice of low-fidelity model. For higher accuracy requirements on the surrogate model, using coarse grid Navier-Stokes simulation as low-fidelity model seems advantageous. However, the error convergence for the latter case exhibits in the present example a stagnation, giving little or no improvement until many more high-fidelity points are added. While the precise cause of the stagnation is unclear, it seems related to the reliability of the low-fidelity model as approximation to the high-fidelity model.

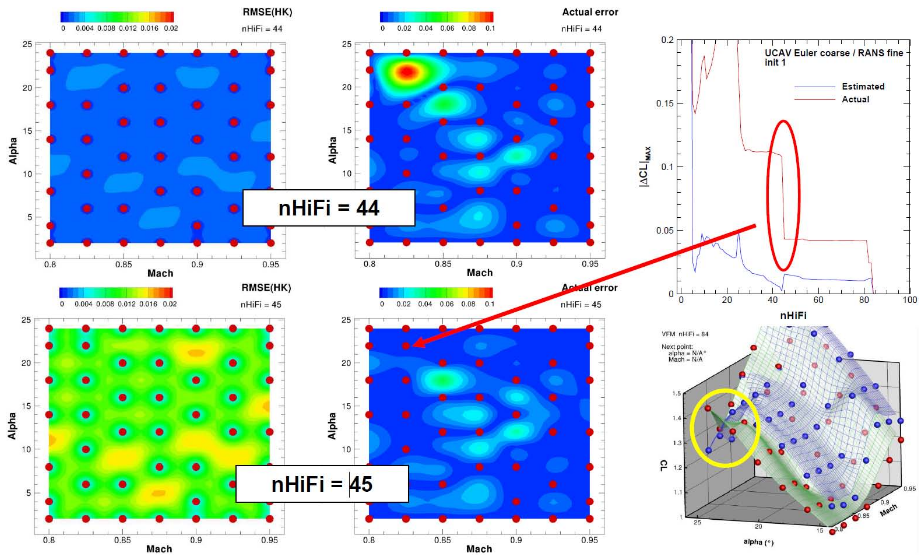


(a) Average error. (b) Maximum error.

Fig. 10. Convergence of $C_L(M, \alpha)$ surrogate model for UAV test case. High-fidelity model: fine grid Navier-Stokes; low-fidelity model: Euler simulation (coarse grid).

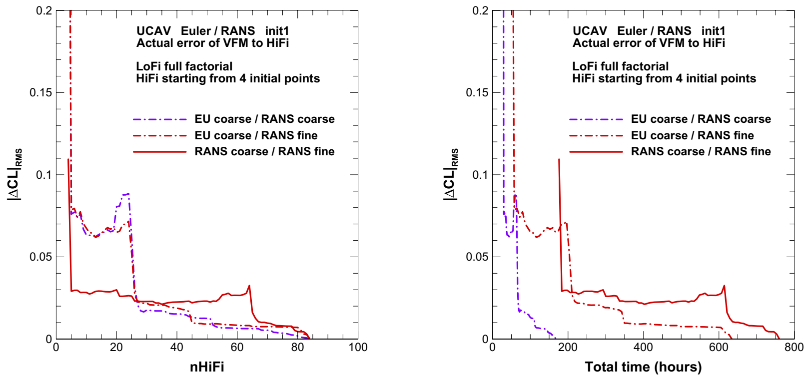


(a) Estimated (left) and actual error (right) at adding 26^{th} high-fidelity point.



(b) Estimated (left) and actual error (right) at adding 45^{th} high-fidelity point.

Fig. 11. Estimated error compared to actual error of the $C_L(M, \alpha)$ surrogate model for UAV test case. High-fidelity model: fine grid Navier-Stokes; low-fidelity model: Euler simulation (coarse grid).



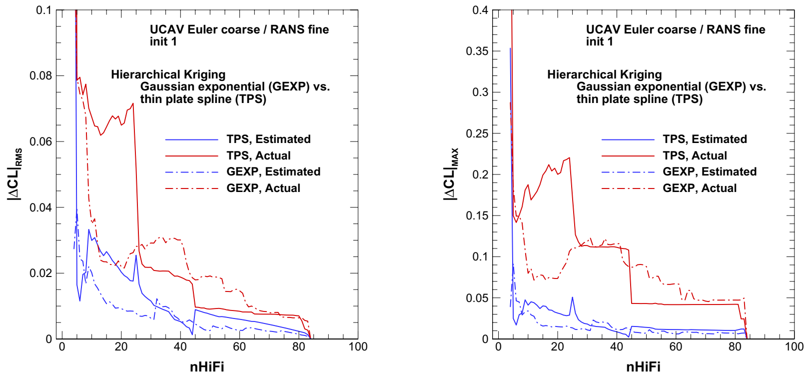
(a) As function of the total number of high-fidelity points.

(b) As function of the total computing time, including low-fidelity model generation.

Fig. 12. Convergence of $C_L(M, \alpha)$ surrogate model for UAV test case; average error on data points. Comparison between coarse grid Navier-Stokes and Euler simulation as low-fidelity model.

4.2 Gaussian Exponential Versus Thin Plate Spline

The convergence of the multi-fidelity surrogate model with the number of high-fidelity data points is shown in Fig. 13 for the approach using Euler simulation



(a) Average error.

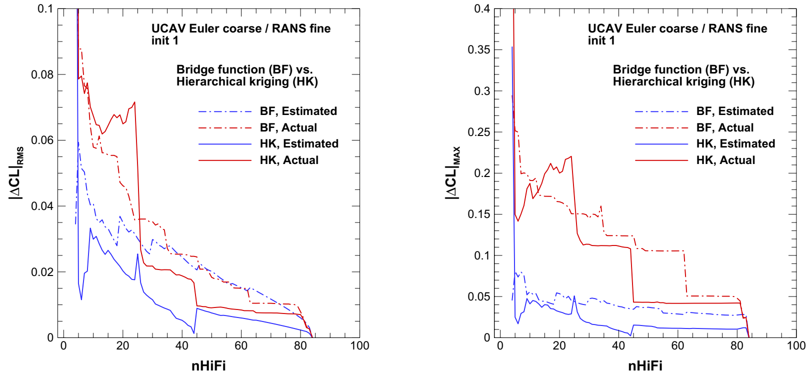
(b) Maximum error.

Fig. 13. Convergence of $C_L(M, \alpha)$ surrogate model for UAV test case; average error on data points. High-fidelity model: fine grid Navier-Stokes; low-fidelity model: Euler simulation (coarse grid). Comparison between Gaussian exponential and 'thin plate spline' Kriging correlation functions.

as low-fidelity model. The Hierarchical Kriging method is used, as in the previously discussed results, but now with two different choices for the Kriging correlation function, namely the Gaussian exponential and the 'thin plate spline' function [2]. No clear advantage of one correlation function over the other can be seen – however this conclusion is likely rather dependent on the specifics of the problem considered.

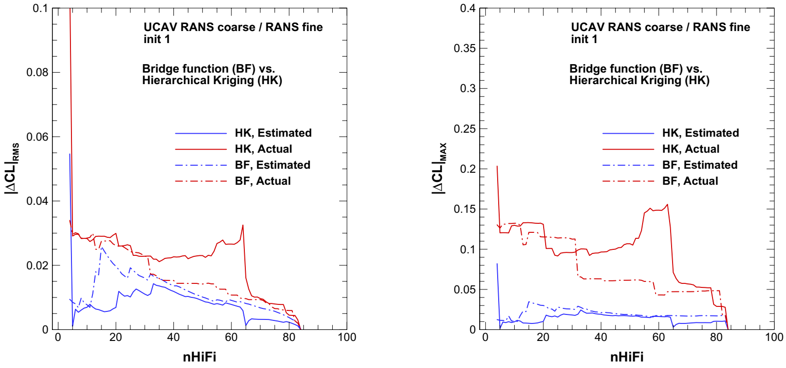
4.3 Kriging with Bridge Function Versus Hierarchical Kriging

In Fig. 14 and Fig. 15 the convergence of the surrogate model with the number of high-fidelity points is shown, respectively with Euler simulation and with coarse grid Navier-Stokes simulation as the low-fidelity model. For each of these two choices of the low-fidelity model a comparison is shown between the two different multi-fidelity methods, i.e. the correction based method using an additive bridge function and the hierarchical Kriging method. In both methods, the Kriging interpolation is based on the 'thin plate spline' correlation function. With coarse grid Navier-Stokes simulation as the low-fidelity model, the bridge function method appears – contrary to expectations – to be advantageous over the hierarchical Kriging method, in terms of a faster reduction of the surrogate model error. With Euler simulation as the low-fidelity model, the hierarchical Kriging method seems however overall advantageous. A fairer evaluation of the merits of the two multi-fidelity methods can of course only be made on the basis of several more test cases.



(a) Average error. (b) Maximum error.

Fig. 14. Convergence of $C_L(M, \alpha)$ surrogate model for UAV test case; average error on data points. High-fidelity model: fine grid Navier-Stokes; low-fidelity model: Euler simulation (coarse grid). Comparison between Kriging with bridge function (BF) and hierarchical Kriging (HK) multi-fidelity approaches.



(a) Average error. (b) Maximum error.

Fig. 15. Convergence of $C_L(M, \alpha)$ surrogate model for UAV test case; average error on data points. High-fidelity model: fine grid Navier-Stokes; low-fidelity model: coarse grid Navier-Stokes. Comparison between Kriging with bridge function (BF) and hierarchical Kriging (HK) multi-fidelity approaches.

5 Concluding Remarks and Prospects

A multi-fidelity scheme for the automated generation of aerodynamic data sets based on adaptive sampling of CFD simulation models of different fidelity levels has been implemented in the Cassidian *SimServer* software environment through a coupling with the DLR *SMART Toolbox*. The application of this automated multi-fidelity scheme has been tested for generating a still relatively small aerodynamic data set, dependent on Mach number and angle of attack, for a generic UAV aircraft configuration. Already on this relatively simple case, the multi-fidelity scheme has been found advantageous for the faster generation of aerodynamic data sets under typical accuracy requirements of the surrogate model.

Based on the current status of the multi-fidelity simulation scheme implemented in *SimServer* and the first applications tested, further developments that can be foreseen towards an industrial usage of such schemes are:

- first of all, building more experience in the application of multi-fidelity simulation for actual aerodynamic problems.
- simultaneous adaptive sampling for multiple output variables, e.g. lift and pitching moment, based on estimated errors relative to user-defined model tolerances for each of the variables,
- investigation of alternative adaptive sampling criteria, e.g. using adaptive gridding [12] or Generic Surrogate Modelling [13],
- development of criteria for the detection and elimination of 'bad' low-fidelity points,
- automation of flow solver restart and parameter changes in case of solver convergence problems,
- extension of the multi-fidelity scheme for the simulation of distributed aerodynamic loads.

References

1. Schwamborn, D., Gerhold, T., Heinrich, R.: The DLR TAU-code: Recent applications in research and industry. In: Proc. of ECCOMAS CFD Conference (2006)
2. Han, Z.-H., Görtz, S., Hain, R.: Hierarchical Kriging Model for Variable-Fidelity Surrogate Modeling. AIAA Journal 50(9), 1885–1896 (2012)
3. Da Ronch, A., Ghoreyshi, M., Badcock, K.J.: On the generation of flight dynamics aerodynamic tables by computational fluid dynamics. Progress in Aerospace Sciences 47, 597–620 (2011)
4. Han, Z.-H., Görtz, S.: Variable-fidelity modelling toolbox, User guide, Version 1.4.5., $C^2A^2S^2E$, DLR, Braunschweig (February 2010)
5. Han, Z.-H., Zimmermann, R., Görtz, S.: A New Cokriging Method for Variable-Fidelity Surrogate Modelling of Aerodynamic Data. AIAA Paper 2010–1225
6. Han, Z.-H., Görtz, S., Hain, R.: A variable-fidelity modeling method for aero-loads prediction. In: Dillmann, A., Heller, G., Klaas, M., Kreplin, H.-P., Nitsche, W., Schröder, W. (eds.) New Results in Numerical and Experimental Fluid Mechanics VII. NNFM, vol. 112, pp. 17–25. Springer, Heidelberg (2010)
7. Tremel, U., Deister, F., Sørensen, K.A., Rieger, H., Weatherill, N.P.: The SimServer – a parallel multidisciplinary simulation environment, DGLR 2003-099, Munich, Germany (2003)
8. Tremel, U., Deister, F., Sørensen, K.A., Rieger, H., Weatherill, N.P.: An object-oriented multidisciplinary simulation system – the SIMSERVER. In: Joubert, G.R., Nagel, W.E., Peters, F.J., Walters, W.V. (eds.) Parallel Computing: Software Technology, Algorithms, Architectures and Applications, Advances in Parallel Computing, pp. 331–338. Elsevier (2004)
9. Gerhold, T., Friedrich, O., Evans, J., Galle, M.: Calculation of Complex Three–Dimensional Configurations Employing the DLR–Tau Code. AIAA Paper 97–0167 (1997)
10. Sørensen, K.A., Tremel, U., Rieger, H., Hitzel, S.M.: Simulation of Manoeuvring Fighter Aircraft with the Unstructured Chimera Approach. AIAA Paper 2007–0123 (2007)
11. Tremel, U., Hitzel, S.M., Sørensen, K.A., Weatherill, N.P.: JDAM–Store Separation From an F/A–18C - An Application of the Multidisciplinary SimServer–System. AIAA Paper 2005–5222 (2005)
12. Rosenbaum, B., Schulz, V.: Comparing sampling strategies for aerodynamic Kriging surrogate models. ZAMM 92(11-12), 852–868 (2012)
13. Rosenbaum, B., Schulz, V.: Efficient response surface methods based on generic surrogate models. SISC (2012) (submitted to)