MULTI-FIDELITY GAUSSIAN PROCESS MODEL FOR CFD AND WIND TUNNEL DATA FUSION

Rubén Conde Arenzana, Andrés F. López-Lopera, Sylvain Mouton, Nathalie Bartoli, Thierry Lefebvre

AeroBest 2021
Lisboa, 21-23 July 2021

MULTI-FIDELITY GAUSSIAN PROCESS MODEL FOR CFD AND WIND TUNNEL DATA FUSION

Rubén Conde Arenzana[1], Andrés F. López-Lopera[1], Sylvain Mouton[2], Nathalie Bartoli[1] and Thierry Lefebvre[1]

1: ONERA/DTIS, Université de Toulouse, F-31055 Toulouse, France
2: ONERA/DS, F-31410 Mauzac, France

Abstract. *In aerodynamic design, both Computational Fluid Dynamics (CFD) simulations and wind tunnel (WT) experiments deliver datasets that are complementary in terms of uncertainties and density of information. In many situations it is desirable to use methods that combine all available information while accounting for the strengths of both sources. In this paper, we combine the sources in a multi-fidelity Gaussian process (GP) model to foretell the aerodynamic forces and moments. The resulting model accounts for input-dependent error measurements in both experimental and simulated data. To validate the model, we construct a proper multi-source aerodynamic database containing CFD and WT data. This database is based on the NASA Common Reference Model. The CFD simulations are based on Reynolds-averaged Navier-Stokes (RANS) equations. We demonstrate in numerical settings that the suggested multi-fidelity GP framework outperforms the single-fidelity one in terms of prediction accuracy at the highest level of fidelity (i.e. WT data). The resulting model also allows to reconstruct common aerodynamic profiles (e.g. representations of longitudinal forces) with uncertainties.*

Keywords: CFD, wind tunnel tests, multi-fidelity databases, error measurements, surrogate models, Gaussian processes

1 INTRODUCTION

Nowadays, aerodynamic data for aerospace vehicles are obtained from different sources such as in-flight tests, wind tunnel (WT) experiments and several types of numerical simulations [1–4]. All those sources provide information about the same flying aircraft, but with different characteristics in terms of accuracy, cost, cycle-time and availability [2]. In principle, data collected from flight testing are the most representative of the reality, but are limited in scope by cost and safety issues. They are also unavailable during most of the design phase, although data from a previous similar aircraft can actually be useful. On the other hand, numerical simulations, in particular Computational Fluid Dynamics (CFD), can be exploited even in early design phases, especially if one bears in mind their significant advances during the last decades [3]. However, their accuracy is still limited by issues such as turbulence modelling and discretization errors, and their ability to produce thousands of flight conditions is also limited by available computational power. Finally, WT testing lies somewhat in between.

Although these sources of data were sometimes seen as competitors in the past, the growing consensus today is that they can be used in a complementary manner for the foreseeable future [2]. Indeed, their characteristics and the aerodynamic data they generate are very different, and, in many instances, complementary. For example, flow simulations are able to provide a complete mapping of the local pressure on the aircraft surface, but generally for a limited number of flow conditions due to the high recurring computational cost involved for each new simulation. Quite the opposite is observed in WT data: the number of pressure measurements is limited by the available room in the model, but any new flow condition is easily obtained by changing the tunnel conditions. In this context, appropriate data fusion methods are needed to consider these datasets in a unified framework aiming at fully leveraging the knowledge gathered by the different sources.

For aerodynamic data fusion, due to the cost of experimental tests including both in-flight and WT tests, many of the previous works focus on simplistic models (e.g. such the vortex lattice model, VLM), and/or CFD simulations [5–7]. It is often assumed that Reynolds-averaged Navier-Stokes (RANS) simulations accurately describe aerodynamic phenomena at a tractable computational cost (in the order of minutes or hours depending on the mesh resolution and machine capacity). On the other hand, since VLM-based frameworks are faster but less accurate, they can provide complementary information that may reinforce aerodynamic databases [5, 6]. As a result, it is possible to meet the desired requirements in terms of reliability but with the advantage of achieving a non-negligible reduction of resources. There are only a few works that exploit WT tests as another source data [see, e.g., 8–12], however they commonly consider a low amount of data (less than 100 points).

Efforts in this direction have been made [6, 8–10], and in particular, surrogate models based on Gaussian processes (GPs) are drawing attention in the aerodynamic community due to its versatility and its ability to quantify uncertainties [11–17]. For example, in [11], a multi-task GP regression model (also known as co-Kriging) has been proposed, accounting for both CFD simulations (Euler and RANS equations) and a notional WT database. In [14], a multi-fidelity GP framework based on the auto-regressive model proposed by [18] has been further investigated for (Bayesian) optimization purposes. As shown by [11, 14], GPs provide promising results that can be exploited for decision tasks when considering multiple aerodynamic data sources. In fusing such datasets, uncertainty quantification is

Figure 1: Large Reference Model (LRM) used in the S1-WT at ONERA.

playing a key role. Quantifying and propagating uncertainties is a difficult topic, both in CFD modeling [see, e.g., 9] and in WT experiments [see, e.g., 19].

Our contributions here are threefold. We first provide a large aerodynamic database that features the implementation of WT data and CFD simulations based on RANS equations. Second, we set up simplified uncertainty models of the CFD and WT data, aiming at defining realistic values for the input varying uncertainties of force coefficients. Finally, we further investigate a multi-fidelity GP framework that exploits both WT and CFD data, while accounting for input-varying uncertainties (heteroscedastic case). The suggested model is based on the auto-regressive GP scheme proposed by [18].

This paper is organized as follows: Section 2 presents the construction of the multi-source database, explaining how WT and CFD data are collected. Section 3 delves into the WT and CFD uncertainty models. The heteroscedastic multi-fidelity GP model is detailed in Section 4. The results are discussed in Section 5. Lastly, Section 6 summarizes the conclusions and potential future works that could complement this study.

2 CONSTRUCTION OF THE MULTI-SOURCE DATABASE

The construction of a proper database is an essential component to develop a data-driven aerodynamic model. In addition, different data acquisition schemes can lead to different features in terms of reliability and lead time, which are usually inversely related. For this purpose, a multi-source aerodynamic database based on the NASA Common Reference Model (CRM) [20] is proposed. One of the main contributions here is the use of WT data provided by ONERA which will be considered as the high-fidelity data. Figure 1 shows one of the models used in the WT experiments. The low-fidelity data are obtained from numerical CFD simulations based on RANS equations. Further details are given in Sections 2.1 and 2.2.

The aerodynamic database contains three inputs: the Mach number (M), the Reynolds number (Re), and the angle of attack (α [deg]). Figure 2 shows pairwise histograms for those input parameters. The outputs are the drag coefficient (C_x), the lift coefficient (C_z), and the pitch moment coefficient (C_M). 4D scatter plots of those coefficients are shown in Fig.3. Section 2.3 describes the methodology considered to construct the design of experiments (DoE) used to run the simulations and to choose the experimental points that will be taken into account to train the model.

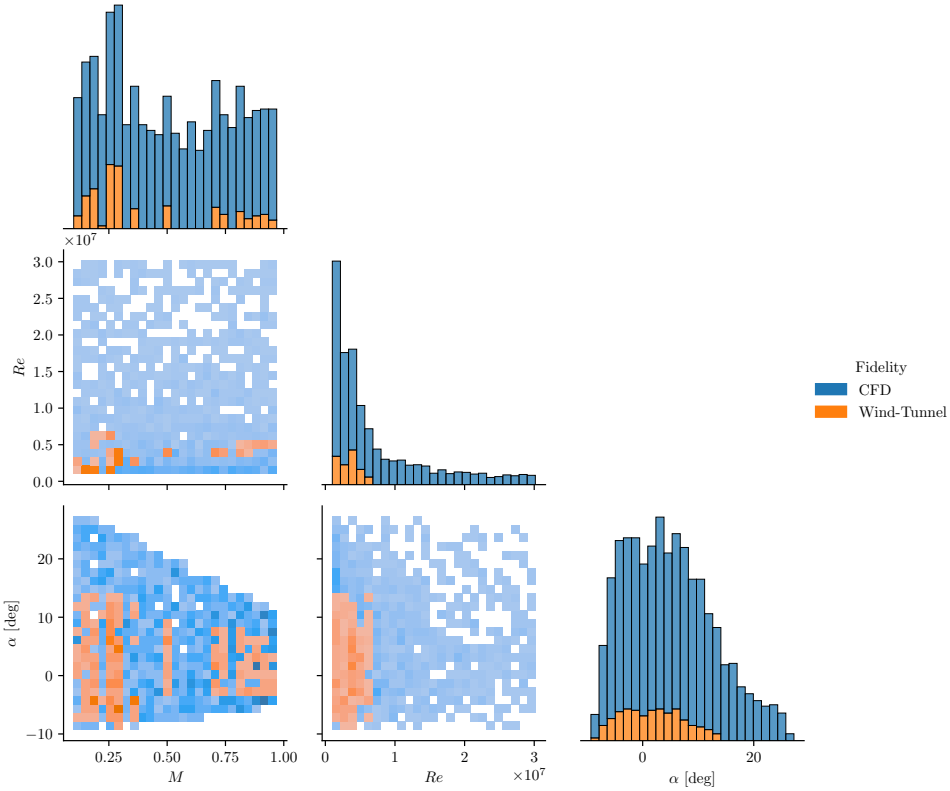


Figure 2: Pairwise histograms of the aerodynamic inputs α, Re and M. Results are shown for both (orange) WT and (blue) CFD data. The panels in the diagonal show 1D marginal histograms, while the other ones show 2D histograms of pairs of inputs. For the 2D illustrations, darker pixels represent a bigger concentration of points.

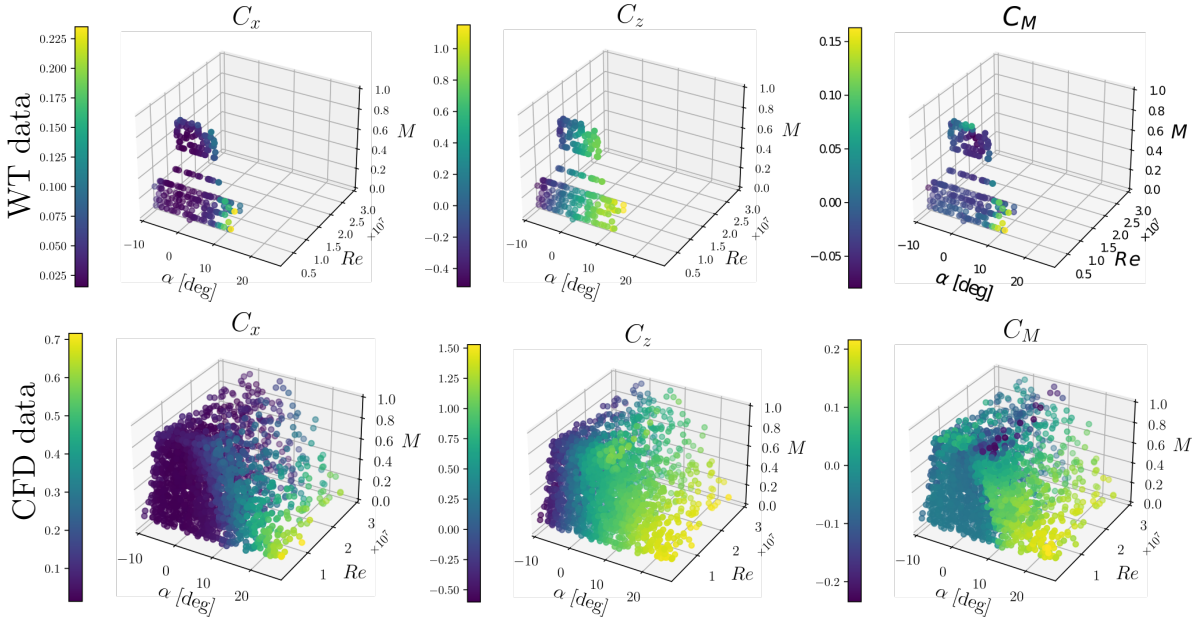


Figure 3: 4D scatter plots for the multi-fidelity database generated in Section 2. 3D axes correspond to the aerodynamic input parameters (α, Re and M), and the colorbar shows the magnitude of the output coefficient (either C_x, C_z or C_M). Results are shown for both (top) WT and (bottom) CFD data.

2.1 Wind tunnel (WT) data

Experimental testing is performed in two different ONERA WT centers: the F1-WT, which focuses on high Re values and $M \in [0.05, 0.36]$ [21]; and the S1MA-WT, which focuses on transonic speeds [22] (see Fig.1). The CRM shape is used to build two Large Reference Models (LRMs): the first one for S1-WT with a wing span of 3.5m (scale 1/16.835) and the second one for F1-WT with a wing span of 3m (scale 1/19.5). Aeroelastic deformations are considered in the design to produce a shape under load that is comparable to previous CRM tests in NTF [23] and ETW [24]. In the database construction, we use the wing-body configuration without tail planes to be consistent with the CFD model. The boundary layer transition is forced on the wing and fuselage. The WT models are equipped with hundreds of pressure taps, but for the present study only the aerodynamic forces measured by internal balances are considered. The measurements are corrected from the effect of the WT walls thanks to potential flow theory [25], which has been recently verified to work well even under transonic conditions [26]. They are also corrected from support effect, using CFD for the S1 database [27], and by performing a dummy sting test for the F1 database. The extent of WT data is limited by the operating envelope of the WT and some other experimental considerations, such as the avoidance of excessive loads on the model support or force balance.

By collecting data from experimental tests performed in the past with the aforementioned configuration, we construct a WT database composed by 5473 samples. Due to the cost of WT tests, in further developments (e.g. in the construction of the DoE, Section 2.3), we use historical WT events instead of performing new tests.

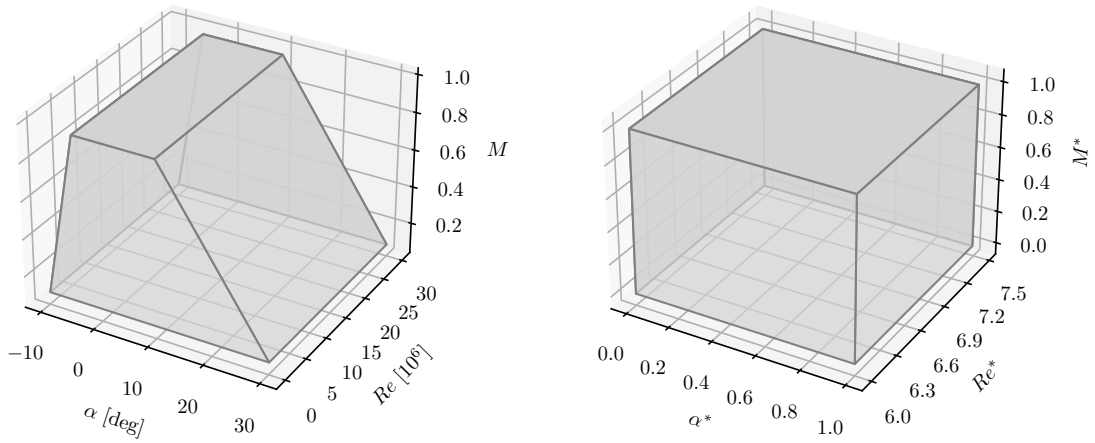
2.2 Numerical CFD data

The RANS equations are the most widespread way to model turbulent flows, and even though if its calculation is usually related to long-time computations, the use of a coarse grid allows the extraction of a large set of points in a suitable time. RANS simulations here are carried out with the elsA software developed at ONERA [28]. In our study, we use a mesh, namely "L1", composed of 63.9×10^3 cells (approximately), which can be retrieved in the DPW5 website.[1] The one equation Spalart-Allmaras turbulence model is used, resulting in computations that are similar to the ones performed in [29]. With this configuration, seven simulations (executed in parallel) are obtained in about 35 min on a Intel® Xeon® E5-2680 v4, 14 cores, 2.4 GHz.

2.3 Design of experiments (DoE) via Latin hypercube sampling

For the construction of a DoE that properly covers the input domain, we consider a nested construction based on a maximin Latin hypercube sampling (LHS). The nested assumption is contemplated to count on an efficient implementation of the surrogate model described in Section 4. We must note that standard LHS are usually defined for hypercube domains [30]. However, this is not our case since we need to account for physical constraints. For example, experiments with simultaneously high values of M and α will lead to high-speed deep stall situations that are far outside of the flight envelope of this aircraft, and are therefore not interesting to include in the database.

[1] Drag Prediction Workshop 5 : `https://aiaa-dpw.larc.nasa.gov/Workshop5/workshop5.html`



(a) Original input space (b) Transformed input space

Figure 4: Aerodynamic input space used for the construction of the DoE via LHS.

By considering the different physical constraints based on [31] and on expert knowledge, it is possible to define the input space in Fig.4a with maximal ranges defined in Tab.1. Such space can be transformed into a cube (see Fig.4b) by using the parametric transformation given by

$$M^* = \frac{M - 0.1}{0.87}, \qquad \alpha^* = \frac{\alpha + 10 - 5M}{40 - 25M^*}, \qquad Re^* = \log_{10}(Re) \tag{1}$$

With this new parametrization (M^*, α^*, Re^*), we can then propose an optimal nested DoE via LHS. To do so, we use the dedicated LHS function `NestedLHS` from the Python toolbox SMT [32], based on the enhanced stochastic evolutionary algorithm proposed in [33]. This function returns an optimal DoE per each level of fidelity while preserving a nested construction. After defining the nested DoE, we then map the generated design points into the original input space using the expressions below:

$$M = 0.1 + 0.87M^*, \qquad \alpha = -10 + 40\alpha^* + 5M^* - 25\alpha^* M^*, \qquad Re = 10^{Re^*} \tag{2}$$

In our study, we generate a nested DoE with 250 and 1949 design points for the WT and CFD databases, respectively (see Fig.2). We must note that, to benefit from WT tests performed in the past, we match the resulting DoE of the highest fidelity level with respect to the WT database described in Section 2.1. This procedure is based on a nearest-neighbor scheme where a generated design point is replaced by the nearest WT input configuration. Although such procedure breaks down the Latin hypercube's properties, the resulting DoE still covers the input space properly (see Fig.2).

2.4 Database overview

As shown in Tab.1, CFD data are taken in a parametric space larger than the one of WT data. This is chosen with the aim of stretching out the study to the entire flight envelope of the aircraft. This is meant to illustrate the complementary nature of the two datasets in hands, one being more accurate but covering only a limited domain of the parametric space, and the other one being used to expand that knowledge, especially in terms of α and Re values.

Table 1: Parameters of the CRM aerodynamic database. For each data source, the ranges of the inputs are shown along with the number of samples that are part of the database.

	Wind Tunnel	elsA CFD
Mach Number (M)	[0.01, 0.95]	[0.01, 0.97]
Reynolds Number (Re) [1×10^6]	[0.77, 6.60]	[0.1, 30.0]
Angle of Attack (α) [deg]	[-9.3, 20.5]	[-10.0, 30.0]
Number of samples	250	1949

The capabilities of the RANS simulations are here pushed to their limits, and sometimes beyond, with cases that do not converge properly, especially at high AoA (see Section 3.2). This kind of data would probably be of little use to actually design the aircraft, but are regarded as sufficient for the purpose of demonstrating the data fusion process.

3 ERROR PROPAGATION ANALYSIS

In assembling data, it is helpful, and sometimes necessary, to know the uncertainty of samples. We now present how the experimental and numerical uncertainties are modelled.

3.1 WT data

A full analysis of experimental uncertainties is not immediately available for the WT tests. Indeed, such analysis has always been a complex subject [see, e.g., 19, 34]. Nevertheless, in the framework of this work, a simplified uncertainty model of the experimental dataset is set up to define realistic values for the uncertainties of C_x, C_z and C_M. This analysis does not consider uncertainties in Mach number, and ignores or overly simplifies a number of sources such as the manufacturing accuracy of the model. Some sources of uncertainty (e.g. on wall corrections) are given by an expert estimate that is appropriate for most of the database, but whose validity is more doubtful beyond stall for example. In spite of its deficiencies, this uncertainty model is regarded as rich enough for using it in the data fusion process that is the topic of the present work.

To be consistent with the surrogate model described in Section 4, all the aerodynamic parameters considered here are regarded as random Gaussian variables. We denote $\tau^2_{\mathcal{X}}$ as the variance parameter of the variable $\mathcal{X}$. In Appendix 8, we provide further details about the uncertainty model.

As an example, the drag coefficient is obtained from balance and dynamic pressure measurement as $C_x = \dfrac{A\cos\alpha + N\sin\alpha}{qS_{\text{ref}}}$, where A and N are the axial and normal aerodynamic forces in model axes (respectively), q is the dynamic pressure, S_{ref} and L_{ref} are the reference length and surface area (respectively) of the model. Consequently, the variance of the error on C_x is given by

$$\tau^2_{C_x} = \left(\frac{\cos\alpha}{qS_{\text{ref}}}\right)^2 \tau^2_A + \left(\frac{\sin\alpha}{qS_{\text{ref}}}\right)^2 \tau^2_N + C_x^2 \left(\frac{\tau_q}{q}\right)^2 + C_z^2 \tau^2_\alpha. \tag{3}$$

From Eq.(3), we observe that $\tau^2_{C_x}$ varies with α. Moreover, as detailed in Appendix 8, τ^2_q and τ^2_α also depend on the values of M and C_z, which itself also depends on M, Re and α. This implies that the resulting variances $\tau^2_{C_x}$, $\tau^2_{C_z}$ and $\tau^2_{C_M}$ are input-varying as

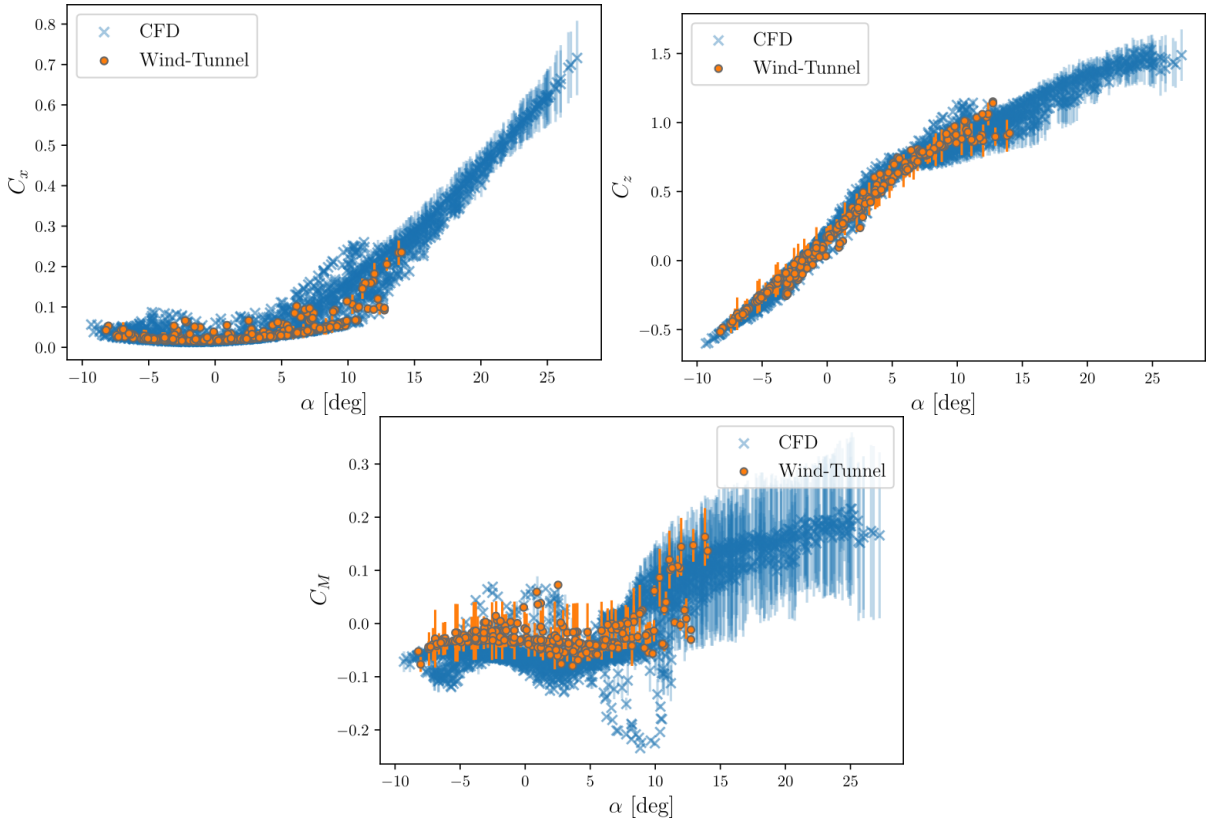


Figure 5: 1D profiles of the force coefficients C_x, C_z and C_M with respect to α [deg]. Panels show: the WT data (orange points), CFD data (blue crosses) and the associated one standard-deviation error intervals (vertical bars). To improve the visibility, four standard-deviation error intervals (orange bars) are displayed for the WT data.

we illustrate in Fig.5. There, panels show 1D profiles of the C_x, C_z, C_M as functions of α. One of the driving factors for this matter of fact is that the force balance is chosen to withstand the maximum forces expected during the test, which makes it oversized for the lowest forces to measure, which occur when the dynamic pressure q is low. In Appendix 8, we also write the expressions of variance of the errors for C_z and C_M.

3.2 CFD simulations

Defining uncertainty on the outcome of CFD simulations is even more difficult than for experiments, since both physical modeling, discretization, and solution errors play a role [35]. In particular, the uncertainty associated with turbulence modelling is an active field of research [36]. A practical solution being not available for the simulations dealt with in the present work, it is decided to use the convergence history of the force coefficient as an indication of simulation uncertainty. For that purpose, the standard deviation of the force coefficients over the last 300 iterations is used. Figure 6 shows an example of an improperly converged CFD simulation (properly converged CFD results are shown in Appendix 7), where the flow solver exhibits limit-cycle oscillations, due to the massive flow separation on the wing upper surface. This is of course a crude approximation of simulation uncertainty, and certainly not a good practice as underlined by [37]. This is contemplated however to demonstrate the ability of the GPs to handle this kind of uncertainty, until better estimates are available.

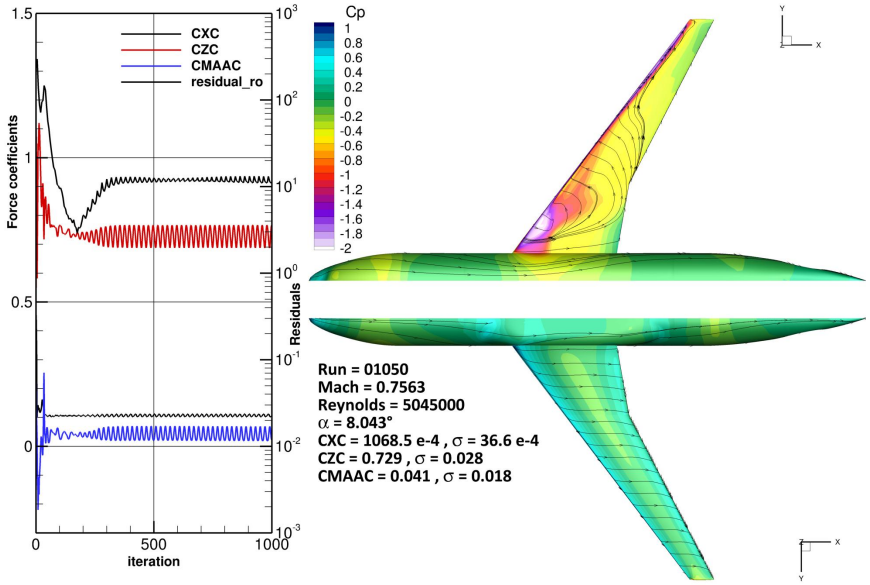


Figure 6: Example of a CFD simulation from the database constructed in Section 2, at $M = 0.7563$, $Re = 0.1 \times 10^6$ and $\alpha = 8.043$ [deg], where limit-cycle oscillations are encountered. On the left panel, convergence plots (solid lines) after 1000 iterations are shown for C_x (CXC, in black), C_z (CZC, in red) and C_M (CMAAC, in blue). On the right panel, the pressure coefficient distribution is depicted for the CRM model.

4 MULTI-FIDELITY GAUSSIAN PROCESS MODEL

In this paper, we study a multi-fidelity surrogate model based on Gaussian processes (GPs). More precisely, we consider the framework proposed by [18]. To account for the error measurements proposed in Section 3, we adapt the GP model in [18] for the case of input-varying additive Gaussian noises (heteroscedastic case).

4.1 Gaussian processes

A GP is a collection of random variables, where the resulting joint distribution of those variables is Gaussian [38]. Let $\{Y(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{D}\}$ be a GP in $\mathbb{R}$ with compact input space $\mathcal{D} \subset \mathbb{R}^d$, e.g. $\mathcal{D} = [0,1]^d$. Then, Y is completely defined by its mean function $m : \mathcal{D} \to \mathbb{R}$ and covariance function (kernel) $k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$, i.e.

$$Y \sim \mathcal{GP}(m, k).$$

For ease of notation, we focus on centered GPs, i.e. $m(\cdot) = 0$, but equations can be generalized for non-centered cases [see, e.g., 38]. An example of a valid (stationary) covariance function is the squared exponential (SE) kernel given by

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma^2 \exp\left(-\sum_{i=1}^{d} \theta_i (x_i - x_i')^2 \right), \tag{4}$$

with $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$, $\boldsymbol{x} = (x_1, \ldots, x_d)$, $\boldsymbol{x} = (x_1', \ldots, x_d')$, σ^2 the variance parameter and $\theta_1, \ldots, \theta_d$ the inverse length-scale parameters.

In regression tasks, we aim at fitting a GP to a training dataset $D = (\boldsymbol{x}_i, f(\boldsymbol{x}_i))_{1 \le i \le n}$, which relies in the computation of the conditional distribution given by

$$Y | \{Y(\boldsymbol{x}_1) = f(\boldsymbol{x}_1), \ldots, Y(\boldsymbol{x}_n) = f(\boldsymbol{x}_n)\} \sim \mathcal{GP}(\mu_n, c_n), \tag{5}$$

where the conditional mean function is given by $\mu_n(\boldsymbol{x}) = \boldsymbol{k}_n^\top(\boldsymbol{x})\boldsymbol{K}_n^{-1}\boldsymbol{y}_n$, with an observation vector $\boldsymbol{y}_n = [y_1 := f(\boldsymbol{x}_1), \dots, y_n := f(\boldsymbol{x}_n)]^\top$; the conditional covariance function is given by $c_n(\boldsymbol{x}, \boldsymbol{x}') = k(\boldsymbol{x}, \boldsymbol{x}') - \boldsymbol{k}_n^\top(\boldsymbol{x})\boldsymbol{K}_n^{-1}\boldsymbol{k}_n(\boldsymbol{x}')$, with a cross-covariance vector $\boldsymbol{k}_n^\top(\boldsymbol{x}) = (k(\boldsymbol{x}, \boldsymbol{x}_i))_{1\leq i\leq n}$ and a covariance matrix $\boldsymbol{K}_n = (k(\boldsymbol{x}_i, \boldsymbol{x}_j))_{1\leq i,j\leq n}$. The conditional process in Eq.(5) is then used as a surrogate model to approximate the target function f. We must note that the kernel $k := k_{\sigma,\theta_1,\dots,\theta_d}$ is parametrized by $\sigma, \theta_1, \dots, \theta_d$ (also known as hyperparameters) that can be estimated via maximum likelihood [38].

4.2 Extension to the multi-fidelity framework

Since the aerodynamic database considers only two data sources (WT and CFD data), here we focus on a multi-fidelity GP framework with two levels of fidelity. The extension to more than two levels of fidelity can be obtained recursively as it is shown by [18, 39]. We denote Y_e and Y_s as the processes related to the WT *experiments* and CFD *simulations*, respectively. As suggested by [18], we consider the auto-regressive model given by

$$Y_e(\boldsymbol{x}) = \rho(\boldsymbol{x})Y_s(\boldsymbol{x}) + \nu(\boldsymbol{x}), \tag{6}$$

where $\rho : \mathcal{D} \to \mathbb{R}$ is a scale factor between Y_s and Y_e, and $\nu : \mathcal{D} \to \mathbb{R}$ the discrepancy function tasked with capturing the differences between both fidelity levels beyond scaling. The scale factor ρ can be a scalar parameter as suggested in [40]. As proposed in [18], we assume $\rho(\boldsymbol{x}) = \sum_{i=1}^q g_i(\boldsymbol{x})\beta_i$ where $g_1, \dots, g_q$ are regression functions (e.g. constant, linear, quadratic), and $\beta_1, \dots, \beta_q$ are their corresponding weights. We must note that predictors of the latter weights are obtained by best linear unbiased estimators [18, 38].

Let Y_s and ν be two independent (centered) GPs given by $Y_s \sim \mathcal{GP}(0, k_s)$ and $\nu \sim \mathcal{GP}(0, k_\nu)$ with covariance functions k_s and k_ν. Due to the linearity of Eq.(6), we can show that Y_e is also (centered) GP-distributed with covariance function given by

$$k_e(\boldsymbol{x}_i, \boldsymbol{x}_j) := \operatorname{cov}\{Y_e(\boldsymbol{x}_i), Y_e(\boldsymbol{x}_j)\} = \rho(\boldsymbol{x}_i)\rho(\boldsymbol{x}_j)k_s(\boldsymbol{x}_i, \boldsymbol{x}_j) + k_\nu(\boldsymbol{x}_i, \boldsymbol{x}_j). \tag{7}$$

According to our aerodynamic application, we aim at predicting profiles of Y_e using both WT data $(\boldsymbol{x}_{e,i}, y_{e,i})_{1\leq i\leq n_e}$ and CFD simulations $(\boldsymbol{x}_{s,i}, y_{s,i})_{1\leq i\leq n_s}$. Therefore, we need to compute the conditional distribution of $Y_e|\{Y_s(\boldsymbol{x}_{s,1}) = y_{s,1}, \dots, Y_s(\boldsymbol{x}_{s,n_s}) = y_{s,n_s}, Y_e(\boldsymbol{x}_{e,1}) = y_{e,1}, \dots, Y_e(\boldsymbol{x}_{e,n_e}) = y_{e,n_e}\}$ which is also GP-distributed [18]. Note that this construction allows to exploit CFD data in order to improve the predictability of the GP model Y_e. We refer to [18] for a further discussion about the predictive formulas of the multi-fidelity GP framework.

In further developments, in order to benefit from efficient computations due to the nested structure of the aerodynamic databases in Section 2, i.e. $(\boldsymbol{x}_{e,i})_{1\leq i\leq n_e} \subseteq (\boldsymbol{x}_{s,j})_{1\leq j\leq n_s}$, we consider the recursive GP formulation proposed in [41].

4.3 Consideration of input-varying noisy observations

To account for additive input-varying noisy observations, we can consider the system:

$$Y_s^{\text{noise}}(\boldsymbol{x}_i) = Y_s(\boldsymbol{x}_i) + \varepsilon_{s,i}, \tag{8}$$

$$Y_e^{\text{noise}}(\boldsymbol{x}_j) = \rho(\boldsymbol{x}_j)Y_s(\boldsymbol{x}_j) + \nu(\boldsymbol{x}_j) + \varepsilon_{e,j}, \tag{9}$$

where $\varepsilon_{s,i} \sim \mathcal{N}\left(0, \tau_{s,i}^2\right)$ and $\varepsilon_{e,j} \sim \mathcal{N}\left(0, \tau_{e,j}^2\right)$, for all $i = 1, \dots, n_s$ and $j = 1, \dots, n_e$. Here we assume that the additive Gaussian noises are independent and identically distributed,

and that they are independent of Y_s and ν. Then we have that Y_s^{noise} and Y_e^{noise} are (centered) GP-distributed with covariance functions given by

$$\widetilde{k}_s(\boldsymbol{x}_i, \boldsymbol{x}_j) = k_s(\boldsymbol{x}_i, \boldsymbol{x}_j) + \tau_{s,i}^2 \delta_{\boldsymbol{x}_i}(\boldsymbol{x}_j), \tag{10}$$

$$\widetilde{k}_e(\boldsymbol{x}_i, \boldsymbol{x}_j) = \rho(\boldsymbol{x}_i)\rho(\boldsymbol{x}_j)k_s(\boldsymbol{x}_i, \boldsymbol{x}_j) + k_\nu(\boldsymbol{x}_i, \boldsymbol{x}_j) + \tau_{e,i}^2 \delta_{\boldsymbol{x}_i}(\boldsymbol{x}_j), \tag{11}$$

with $\tau_{s,i}^2$ and $\tau_{e,i}^2$ the noise variances at the input $\boldsymbol{x}_i$, and $\delta_{\boldsymbol{x}_i}(\boldsymbol{x}_j)$ the Dirac delta function that is equal to one if $\boldsymbol{x}_i = \boldsymbol{x}_j$, and zero otherwise. Equation (11) follows a similar structure than the one in Eq.(7) but with one additional terms $\tau_{e,i}^2 \delta_{\boldsymbol{x}_i}(\boldsymbol{x}_j)$ that models the input-varying noises. Since Y_s^{noise} and Y_e^{noise} are GP-distributed, we can establish the formulas for computing the conditional Gaussian distribution $Y_e | \{Y_s(\boldsymbol{x}_{s,1}) + \varepsilon_{s,1} = y_{s,1}, \ldots, Y_s(\boldsymbol{x}_{s,n_s}) + \varepsilon_{s,n_s} = y_{s,n_s}, Y_e(\boldsymbol{x}_{e,1}) + \varepsilon_{e,1} = y_{e,1}, \ldots, Y_e(\boldsymbol{x}_{e,n_e}) + \varepsilon_{e,n_e} = y_{e,n_e}\}$. Those formulas follow the same structure as the ones for the noise-free case but replacing $k_s(\boldsymbol{x}, \boldsymbol{x}')$ to $\widetilde{k}_s(\boldsymbol{x}, \boldsymbol{x}')$ and $k_e(\boldsymbol{x}, \boldsymbol{x}')$ to $\widetilde{k}_e(\boldsymbol{x}, \boldsymbol{x}')$. Since the Gaussian noises are mutually independent, independent of Y_s and ν, then the cross-covariance terms will remain unchangeable:

$$\begin{aligned}
\operatorname{cov}\left\{Y_e^{\text{noise}}(\boldsymbol{x}_i), Y_s^{\text{noise}}(\boldsymbol{x}_j)\right\} &= \operatorname{cov}\left\{\rho(\boldsymbol{x}_i)Y_s(\boldsymbol{x}_i) + \nu(\boldsymbol{x}_i) + \varepsilon_{e,i},\ Y_s(\boldsymbol{x}_j) + \varepsilon_{s,j}\right\} \\
&= \operatorname{cov}\left\{\rho(\boldsymbol{x}_i)Y_s(\boldsymbol{x}_i) + \nu(\boldsymbol{x}_i),\ Y_s(\boldsymbol{x}_j)\right\} \\
&= \operatorname{cov}\left\{Y_e(\boldsymbol{x}_i), Y_s(\boldsymbol{x}_j)\right\}.
\end{aligned}$$

In our application, the noise variances $\varepsilon_{s,1}, \ldots, \varepsilon_{s,n_s}$ and $\varepsilon_{e,1}, \ldots, \varepsilon_{e,n_e}$ are known and they are defined in Section 3. However, we must note that they can also be estimated via sensible estimators [42] or via maximum likelihood [43]. In those cases, we need to account for repetitions of observation, i.e., for a fixed set of aerodynamic inputs $\boldsymbol{x}$, we have to execute n_{e,r_e} times the same WT test and keep the outcomes $Y_e^{(1)}(\boldsymbol{x}), \ldots, Y_e^{(n_{e,r_e})}(\boldsymbol{x})$. Taking into account the cost (in both time and resources) of experimental WT tests, we can easily conclude that the approaches in [42, 43] become expensive. This drawback is mitigated by considering directly the error propagation analysis described in Section 3.

4.4 1D numerical illustration

Python codes of the resulting heteroscedastic GP framework described in Section 4.3 are available in the *Surrogate Modeling Toolbox* (SMT) [32], an open-source toolbox developed by ONERA, ISAE-SUPAERO, University of Michigan and NASA Glenn.

Figure 7 shows 1D predictions under heteroscedastic assumptions either considering an independent GP model for each level of fidelity or the multi-fidelity GP framework.[2] We consider the one dimensional "benchmark" problem proposed in [44] and used in some other works related to multi-fidelity surrogate models [14, 45]. The target functions are:

$$f_s(x) = (6x - 2)^2 \sin(2[6x - 2]), \quad \text{(low-fidelity function)}$$

and

$$f_e(x) = 0.5 f_s(x) + 10(x - 0.5) - 5, \quad \text{(high-fidelity function)}$$

for $x \in [0, 1]$. As a nested DoE, we proposed 12 and 4 equidistant low-fidelity and high-fidelity design points, respectively. For illustration, the associated error measurements

[2]The example in Fig. 7 can be reproduced using the Python-based Jupyter notebook available in SMT: `https://github.com/SMTorg/smt/blob/master/tutorial/SMT_MFK_Noise.ipynb`

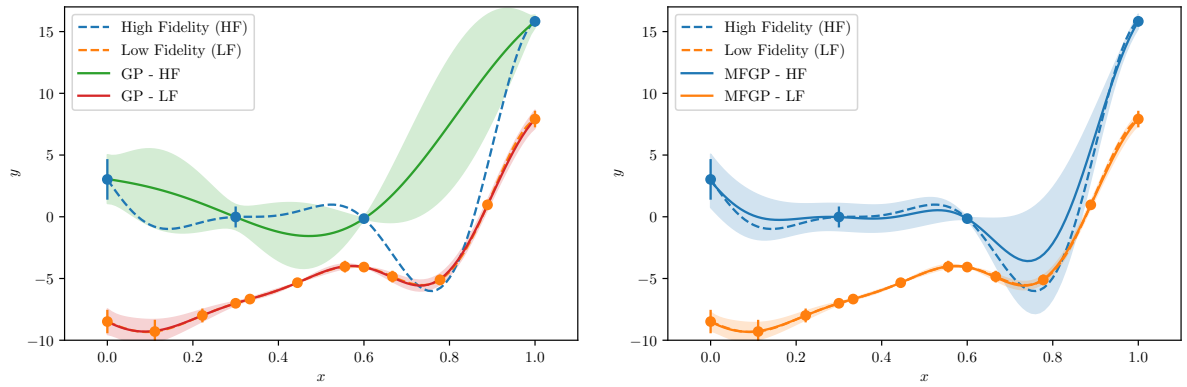


Figure 7: Predictions considering (left) an independent GP per level of fidelity or (right) a multi-fidelity GP (MFGP). Both schemes account for the input-varying noise variances described in Section 4.4. Each panel shows: the target low-fidelity and high-fidelity functions (in orange and blue dashed lines, respectively), the design points (dots) with the corresponding 3 standard-deviation error bars, the resulting predictions (solid lines) and 3 standard-deviation confidence intervals (light areas) provided by the GP frameworks.

are manually fixed and are given by

$$\tau_s^2 = (10^{-1}, 3.5\times10^{-2}, 10^{-3}, 3\times10^{-8}, 2.5\times10^{-2}, 2.5\times10^{-2}, \\ 2\times10^{-2}, 1.5\times10^{-2}, 10^{-1}, 5\times10^{-3}, 2.5\times10^{-6}, 5\times10^{-2}),$$

and $\tau_e^2 = (3\times10^{-1}, 8\times10^{-2}, 2\times10^{-3}, 2\times10^{-2})$.

For the GP models, we consider SE kernels (see Eq.(4)), with covariance parameters estimated via maximum likelihood.[3] In Fig.7, we observe that the noise variance of both independent GPs and the multi-fidelity GP properly cover the error measurements (vertical bars) at the design points (dots), however, prediction at the high-fidelity level f_e is significantly improved by considering the multi-fidelity GP framework, with a better approximation where data are scarce (e.g. around $x = 0.8$).

5 RESULTS

5.1 Assessment of the multi-fidelity GP model

We here assess the performance of the multi-fidelity GP model considering different percentage of training WT data. The training data are randomly chosen from the nested database generated in Section 2. We test three different GP models for each coefficient (C_x, C_z and C_M): two independent single-fidelity GPs exploiting data either from CFD simulations (namely GP-CFD) or WT experiments (namely GP-WT), and a multi-fidelity GP (MFGP) that exploits both data sources. All the models are trained using the 100% of the CFD dataset (1949 data points). For the models accounting for WT data (250 data points), we consider different WT training dataset, i.e. 10%, 30%, 50%, 70% and 90%. Matérn 5/2 kernels are considered as covariance functions for the GP priors:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma^2 \prod_{i=1}^{3} \left(1 + \sqrt{5}\theta_i|x_i - x_i'| + \frac{5}{3}\theta_i^2(x_i - x_i')^2\right) \exp\left(-\sqrt{5}\theta_i|x_i - x_i'|\right),$$

[3] We considered as initial covariance parameters $\sigma_e^2 = \sigma_s^2 = 1$, $\theta_e = 0.1$ and $\theta_s = 0.5$.

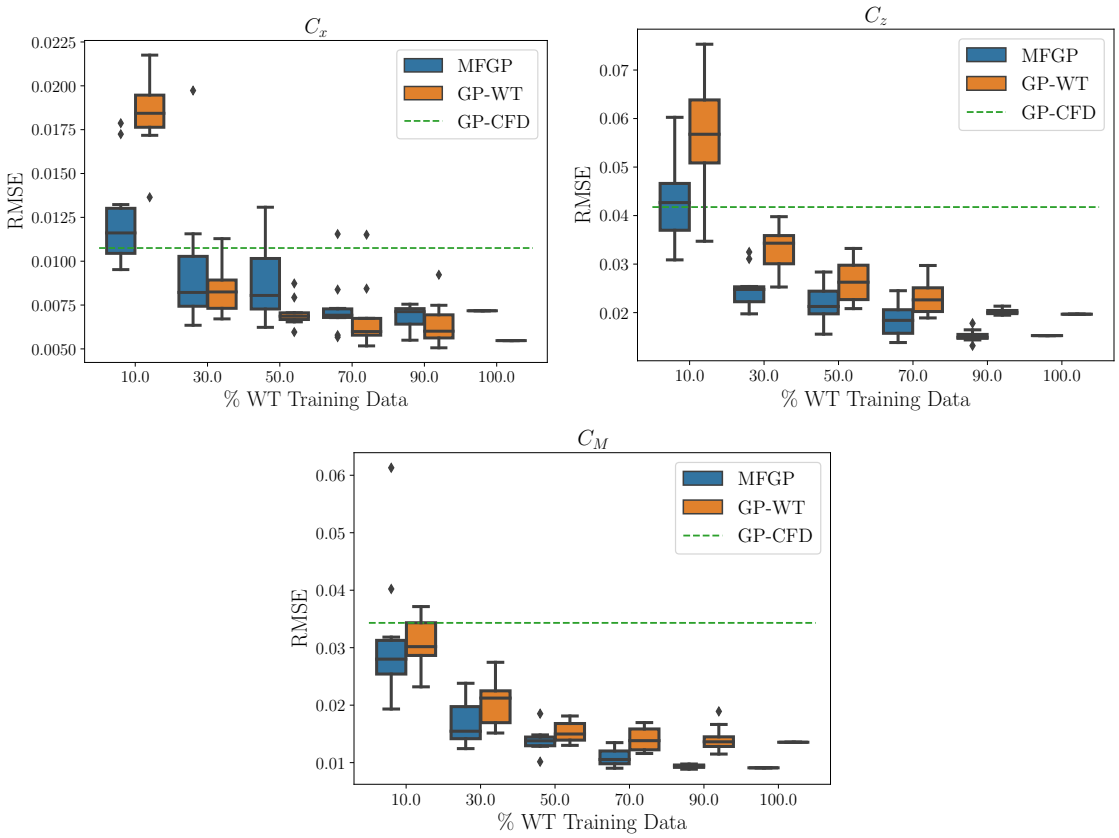


Figure 8: RMSE results for three different GP models for each coefficient (C_x, C_z and C_M): two independent single-fidelity GPs exploiting data either from CFD simulations (GP-CDF, green line) or WT experiments (GP-WT, orange), and a multi-fidelity GP that exploits both data sources (MFGP, blue). All the models are trained using the 100% of the CFD data (1949 data points). For the models accounting for WT data (250 data points), we consider different WT training dataset. Boxplots are computed over ten different random replicates.

with $\boldsymbol{x} = (x_1, x_2, x_3)$.[4] The hyperparameters $(\sigma, \theta_1, \theta_2, \theta_3)$ are estimated via maximum likelihood. As mean functions of the priors, we assume linear trends for all the forces coefficients, i.e. $m_{\text{linear}}(\boldsymbol{x}) = a_0 + a_1\boldsymbol{x}$, except for C_x where we consider a quadratic trend $m_{\text{quadratic}}(\boldsymbol{x}) = a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2$. The coefficients a_0, a_1, a_2 are obtained via best linear unbiased estimation within a universal GP framework [38]. Those trends are defined according to expert knowledge. For the testing step, the predictions of the resulting models are assessed on a validation WT database containing a thousand test points. The test WT DoE is constructed via LHS as discussed in Section 2.3.

Figure 8 shows the root mean square error (RMSE)[5] results for ten different random replicates. The RMSE values obtained when considering 100% of the WT data for training the GP models are also displayed. From C_z and C_M profiles, we observe that the MFGP outperformed the RMSE results of the single-fidelity implementations GP-CFD and GP-

[4]Other types of kernels (e.g. SE kernels) were tested but the Matérn 5/2 class resulted in more accurate results.

[5]RMSE $= \sqrt{\frac{1}{n_{\text{test}}}\sum_{i=1}^{n_{\text{test}}}(y_i - \mu_i)^2}$ with $y_1, \dots, ny_{n_{\text{test}}}$ the test observations and $\mu_1, \dots, \mu_{n_{\text{test}}}$ the corresponding GP predictions.

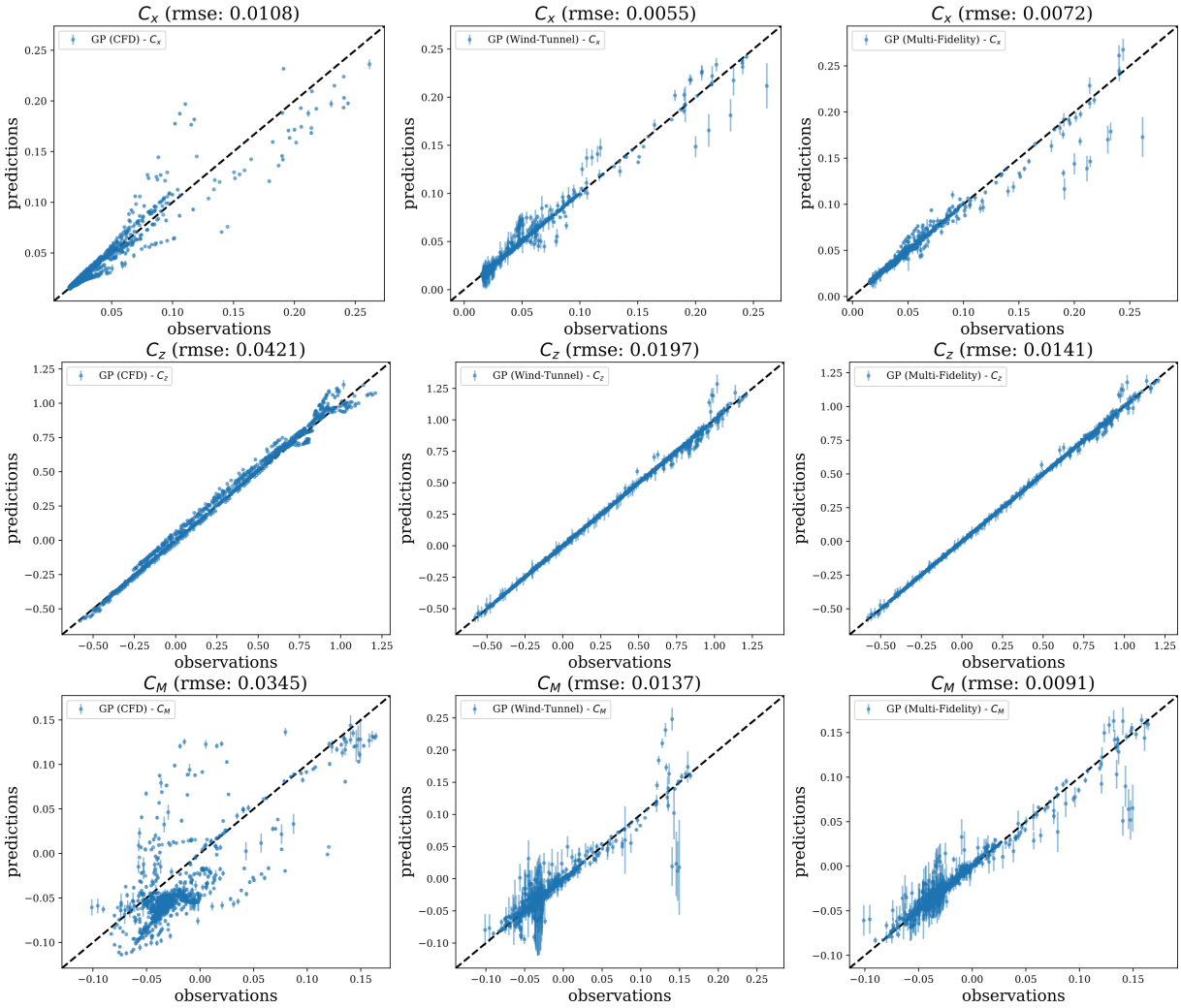


Figure 9: Scatter plots of the test data vs predictions obtained by two independent GPs exploiting data from CFD simulations (left) or WT tests (middle), and a multi-fidelity GP that exploits both types of data (right). Results are shown for C_x (top), C_z (middle) and C_M (bottom). The vertical bars represent the one standard-deviation conditional predictive errors led by the models. RMSE indicators are shown on top of each panel.

WT, leading to significant improvements when WT training sets are small (e.g. using less than 30% of the WT data). For the C_x coefficient, although we note smaller RMSE values when considering only 10% of the WT data, the quality of predictions provided by the MFGP is degraded when the number of WT data points increases. This drawback is produced due to WT data are scarce for $C_x > 0.2$, and therefore, predictions rely mostly on the biased CFD simulations. This bias is observed in Fig.5, and is also reflected by the Pearson correlation coefficient. By computing the Pearson coefficient only over the C_x observations associated to the 250 design points shared by both WT and CFD data, and for $C_x > 0.2$, it results in a value of 0.863 (compared to 0.934 for the case where $C_x \leq 0.2$). To mitigate this issue, we can add additional WT data with $C_x > 0.2$ aiming at better learning the discrepancy function ν.

In Fig.9, we show scatter plots of the test data vs predictions obtained by the proposed GP models accounting for the 100% of WT data. Observe that clearer improvements are

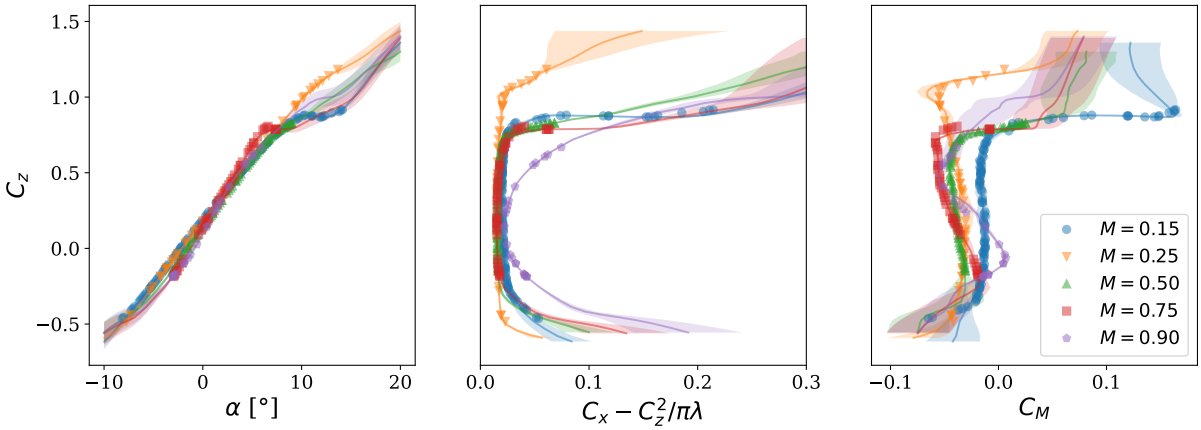


Figure 10: Multi-fidelity GP predictions of common polar of longitudinal forces used in the analysis of WT experiments. For $Re = 7 \times 10^6$ and different values of M, each panel shows: the nearest available test points (dots), the conditional GP mean functions (solid lines) and one standard-deviation predictive confidence intervals (light areas). Confidence intervals are considered for the horizontal-axes except for the plot with respect to α where uncertainty is displayed vertically.

obtained for high-values of the lift coefficient C_z. For C_x and C_M, we note that significant mispredictions are obtained for high values. Due to data are scarce for $C_x > 0.2$ and $C_M > 0.15$, and the high variability of the C_M (see Fig.3 and Fig.5), we can conclude that the MFGP model could not learn such behaviours from the training dataset. Note also that those mispredictions are also encountered when considering either the GP-CFD or GP-WT. As pointed out in the results from Fig.8, this drawback can be mitigated by considering additional training points for $C_x > 0.2$ and $C_M > 0.15$. The poor predictions of the GP-CFD models are justified by the existing bias between the CFD and WT data.

5.2 Representation of longitudinal forces

As discussed in Section 4, one of the main benefits of considering (multi-fidelity) GP models relies in the associated confidence intervals of predictions. Next, we show how those confidence intervals can be exploited in the analysis of WT experiments, more precisely, in the construction of polar representation of longitudinal forces. In Fig.10, for a fixed Reynolds number $Re = 7 \times 10^6$ and for different values of Mach number M, we show the profiles of C_z with respect to α, $C_x - C_z^2/(\pi\lambda)$, with $\lambda = 9$ the LRM model wing aspect ratio,[6] and C_M. Confidence intervals are considered for the horizontal-axes except for the profile with respect to α where the uncertainty is displayed for the vertical-axe. We observe that the resulting multi-fidelity GP model commonly leads to reasonable aerodynamic profiles, where higher uncertainties are obtained in regions where data are not available. We must note that those uncertainties can be reduced by adding new WT training points on those regions. This may motivate the proposition of adaptive DoE aiming at reducing uncertainty in multi-fidelity GP models [46, 47]. We do not consider such improvement in this paper but it can be considered in further developments.

[6]Since $C_z^2/(\pi\lambda)$ is a crude estimate of lift-induced drag (from Prandtl lifting-line theory), then the term $C_x - C_z^2/(\pi\lambda)$ is an estimate of the drag other than lift induced, sometimes called profile drag.

6 CONCLUSION

In this paper, several contributions were addressed in the field of aerodynamic data fusion. First, we constructed a large multi-source aerodynamic database based on the NASA Common Research Model. The database contains 250 and 1949 data points, from wind-tunnel tests and CFD simulations (respectively), covering the entire flight envelope of the model. Second, we set up a simplified uncertainty model of the experimental WT data aiming at defining realistic values for the input varying uncertainties of force coefficients. Finally, we adapted a multi-fidelity Gaussian process (GP) framework to account for input-varying additive noises (heteroscedastic case) in all the levels of fidelity. We demonstrated, in both a synthetic example and a real-world application, that the resulting multi-fidelity GP outperformed single-fidelity ones in terms of predictive capability.

The work presented in this paper can be improved in different ways. A further investigation may be contemplated with the increase in the number of inputs and outputs. We may consider the slideslip angle as an input to allow the prediction of the lateral aerodynamic forces and moments of the aircraft. We can also consider additional data sources, such as numerical simulation based on VLM or Euler equations. Those sources will be considered as lower or intermediate fidelity levels in the multi-fidelity GP framework. Finally, to improve the predictability of the GP model, and looking for uncertainty reduction, adaptive design of experiments can be further investigated accounting for multi-fidelity schemes.

ACKNOWLEDGMENTS

This work was supported by ONERA internal research project dedicated to multidisciplinary design optimization, namely MUFIN. We would like to acknowledge Stefan Goertz and Philipp Bekemeyer (The German Aerospace Center DLR) for the further discussions within the collaboration DLR-ONERA AI4aerospace, ONERA-DLR Joint Virtual Center in AI for Aerospace Engineering. We also thanks the role of the wind tunnel teams in Modane-Avrieux and Le Fauga-Mauzac test centers (both in France) in creating the WT database, especially Aurélia Cartieri.

REFERENCES

[1] F. Stoliker. Introduction to flight test engineering. AGARDograph AG-300, pages 1–456. RTO/NATO, 2005. ISBN 92-837-1126-2.

[2] M. Malik and D. Bushnell. Role of computational fluid dynamics and wind tunnels in aeronautics R and D. NASA/TP–2012-217602, pages 1–63, 2012.

[3] F. Johnson, E. Tinoco, and N. Yu. 30 years of development and application of CFD at Boeing commercial airplanes, Seattle. *Computers & Fluids*, 34:1115–1151, 2005.

[4] J. Slotnick, A. Khodadoust, J. Alonso, D. Darmofal, W. Gropp, E. Lurie, and D. Mavripilis. CFD vision 2030 study: A path to revolutionary computational aerosciences. NASA/CR-2014-218178, pages 1–51. NASA, 2014.

[5] D. Singh, A. Antoniadis, P. Tsoutsanis, H. Shin, A. Tsourdos, S. Mathekga, and K. Jenkins. A multi-fidelity approach for aerodynamic performance computations of formation flight. *Aerospace*, 5:66, 2018.

[6] Z. Zhang, N. Bartoli, A. Jungo, W. Lammen, E. Baalbergen, and M. Voskuijl. Data fusion and aerodynamic surrogate modeling for handling qualities analysis. HAL preprint, 2020.

[7] J. Kou and W. Zhang. Multi-fidelity modeling framework for nonlinear unsteady aerodynamics of airfoils. *Applied Mathematical Modelling*, 76, 2019.

[8] Y. Kuya, K. Takeda, X. Zhang, and A. Forrester. Multifidelity surrogate modeling of experimental and computational aerodynamic data sets. *AIAA Journal*, 49:289–298, 2011.

[9] J. Mukhopadhaya, B. Whitehead, J. Quindlen, and J. Alonso. Multi-fidelity modeling of probabilistic aerodynamic databases for use in aerospace engineering. *arXiv: Fluid Dynamics*, 2019.

[10] L. He, W. Qian, T. Zhao, and Q. Wang. Multi-fidelity aerodynamic data fusion with a deep neural network modeling method. *7th European Conference for Aeronautics and Space Sciences (EUCASS)*, 22, 2020.

[11] A. Feldstein, D. Lazzara, N. Princen, and K. Willcox. Multifidelity data fusion: Application to blended-wing-body multidisciplinary analysis under uncertainty. *AIAA Journal*, 58(2):889–906, 2020.

[12] N. Nigam, S. Mohseni, J. Valverde, S. Voronin, J. Mukhopadhaya, and J. Alonso. Toolset for creation of multi-fidelity probabilistic aerodynamic databases. pages 1–20, 2021.

[13] A. Bertram and R. Zimmermann. Theoretical investigations of the new coKriging method for variable-fidelity surrogate modeling: Well-posedness and maximum likelihood training. *Advances in Computational Mathematics*, 44, 2018.

[14] M. Meliani, N. Bartoli, T. Lefebvre, M. Bouhlel, J. Martins, and J. Morlier. Multi-fidelity efficient global optimization: Methodology and application to airfoil shape design. In *AIAA Aviation Forum*, pages 1–18. 2019.

[15] M. Bouhlel, N. Bartoli, J. Morlier, and A. Otsmane. Improving Kriging surrogates of high-dimensional design models by partial least squares dimension reduction. *Structural and Multidisciplinary Optimization*, 53:935–, 2016.

[16] B. Peherstorfer, K. Willcox, and M. Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60:550–591, 2018.

[17] D. Khatamsaz and D. Allaire. A comparison of reification and coKriging for sequential multi-information source fusion. pages 1–14, 2021.

[18] L. Le Gratiet. *Multi-Fidelity Gaussian Process Regression for Computer Experiments*. Theses, Université Paris-Diderot - Paris VII, 2013.

[19] AGARD Advisory Report No. 304. Quality assessment for wind tunnel testing. pages 1–81. NATO, 1994.

[20] J. Vassberg, M. Dehaan, M. Rivers, and R. Wahls. Development of a common research model for applied CFD validation studies. pages 1–22, 2008.

[21] J. Carrara and A. Masson. Three years of operation of the ONERA pressurized subsonic wind tunnel. In *12th Congress of the International Council of the Aeronautical Sciences*, pages 778–792, Munich, Germany, 1980. ICAS.

[22] A. Cartieri, D. Hue, A. Chanzy, and O. Atinault. Experimental investigations on common research model at ONERA-S1MA–Drag prediction workshop numerical results. *Journal of Aircraft*, 55, 2017.

[23] M. Rivers and A. Dittberner. Experimental investigation of the NASA common research model. In *28th AIAA Applied Aerodynamics Conference*, pages 1–29, 2010.

[24] M. Rivers, J. Quest, and R. Rudnik. Comparison of the NASA common research model European transonic wind tunnel test data to NASA national transonic facility test data. *CEAS Aeronautical Journal*, 9:307–317, 2018.

[25] X. Vaucheret. Recent calculation progress on wall interferences in industrial wind tunnels. *La Recherche Aérospatiale*, pages 45–47, 1988.

[26] J. Hantrais-Gervois and J. Piat. A methodology to derive wind tunnel wall corrections from RANS simulations. In *Advanced Wind Tunnel Boundary Simulation*, pages 1–24. ST0/NATO, 2018.

[27] A. Cartieri and D. Hue. Using RANS computations to calculate support interference effects on the common research model. In *Advanced Wind Tunnel Boundary Simulation*, pages 1–18. ST0/NATO, 2018.

[28] L. Cambier, S. Heib, and S. Plot. The ONERA elsA CFD software: Input from research and feedback from industry. *Mechanics & Industry*, 14:159–174, 2013.

[29] D. Hue. 5th drag prediction workshop: Computational fluid dynamics studies carried out at ONERA. *Journal of Aircraft*, 51(4):1295–1310, 2014.

[30] M. Mckay, R. Beckman, and W. Conover. A comparison of three methods for selecting vales of input variables in the analysis of output from a computer code. *Technometrics*, 21:239–245, 1979.

[31] A. Da Ronch, M. Panzeri, M. Bari, R. d'Ippolito, and M. Franciolini. Adaptive design of experiments for efficient and accurate estimation of aerodynamic loads. In *6th Symposium on Collaboration in Aircraft Design*, 2016.

[32] M. Bouhlel, J. Hwang, N. Bartoli, R. Lafage, J. Morlier, and J. Martins. A Python surrogate modeling framework with derivatives. *Advances in Engineering Software*, page 102662, 2019. URL `https://github.com/SMTorg/smt`.

[33] R. Jin, W. Chen, and A. Sudjianto. An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, 134: 268–287, 2005.

[34] J. Walter, W. Lawrence, D. Elder, and M. Treece. Development of an uncertainty model for the national transonic facility. In *27th AIAA Aerodynamic Measurement Technology and Ground Testing Conference*, pages 1–18, 2012.

[35] W. Oberkampf and F. Blottner. Issues in computational fluid dynamics code verification and validation. *AIAA Journal*, 36(5):687–695, 1998.

[36] A. Mishra, J. Mukhopadhaya, G. Iaccarino, and J. Alonso. Uncertainty estimation module for turbulence model predictions in SU2. *AIAA Journal*, 57(3):1066–1077, 2019.

[37] P. Spalart and V. Venkatakrishnan. On the role and challenges of CFD in the aerospace industry. *Aeronautical Journal*, 120(1223):209–232, 2016.

[38] C. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, Cambridge, MA, 2005.

[39] L. Le Gratiet. Bayesian analysis of hierarchical multifidelity codes. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):244–269, 2013.

[40] M. Kennedy and A. O'Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13, 2000.

[41] L. Le Gratiet and J. Garnier. Recursive co-Kriging model for design of computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5):365–386, 2014.

[42] B. Ankenman, B. Nelson, and J. Staum. Stochastic Kriging for simulation metamodeling. *Operations Research*, 58:371–382, 2010.

[43] M. Binois, R. B. Gramacy, and M. Ludkovski. Practical heteroscedastic Gaussian process modeling for large simulation experiments. *Journal of Computational and Graphical Statistics*, 27(4):808–821, 2018.

[44] A. Forrester, A. Sobester, and A. Keane. Multi-fidelity optimization via surrogate modelling. *Proceedings of the Royal Society A*, 463:3251–3269, 2007.

[45] J. Bailly and D. Bailly. Multifidelity aerodynamic optimization of a helicopter rotor blade. *AIAA Journal*, 57:1–12, 2019.

[46] R. Jin, W. Chen, and A. Sudjianto. On sequential sampling for global metamodeling in engineering design. 28th Design Automation Conference, pages 539–548, 2002.

[47] V. Picheny, D. Ginsbourger, O. Roustant, R. Haftka, and N. Kim. Adaptive designs of experiments for accurate approximation of a target region. *Journal of Mechanical Design*, 132(7):071008, 2010.

APPENDIX

7 ADDITIONAL CFD SIMULATION

In this appendix, we show two additional CFD results: an example where elsA has properly converged after 1000 iterations, and another example where limit-cycle oscillations are encountered (see Fig.11).

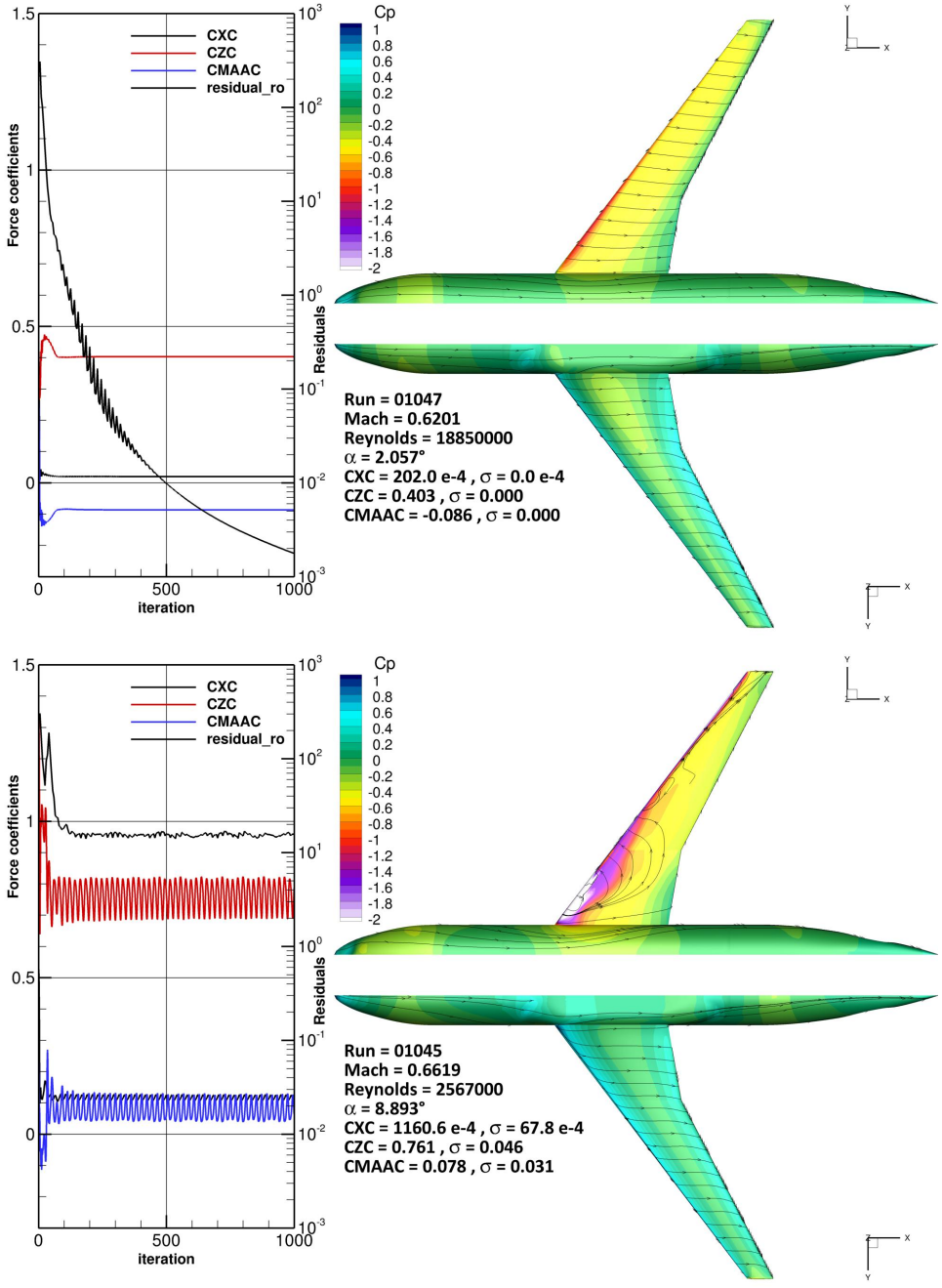


Figure 11: Examples of CFD simulations from the database constructed in Section 2. The panel description is the same as the one in Fig.6.

8 EXPERIMENTAL UNCERTAINTY

In this appendix, we give some complementary information required in Section 3. More precisely, we provide further details about the estimation of error associated to the angle of attack α, the aerodynamic forces, and the dynamic pressure. According to the structure of the database constructed in Section 2, a zero sideslip angle is considered.

Experimental error in aerodynamic forces in model axes. The measurement of the aerodynamic forces is carried out in the body axes by an internal strain-gauge balance. Each force component is considered to be affected by an uncertainty equal to $c = 0.5\times10^{-3}$ of the balance capacity on this component, independently of the other components:

$$\tau_A = cA_{\text{capa}}, \qquad \tau_N = cN_{\text{capa}}, \qquad \tau_{M_A} = cM_{A,\text{capa}}.$$

We neglect uncertainties associated to the positioning of the balance axes with respect to the model axes, and to the calculation of the weight of the model, that is subtracted from raw balance readings to deduce aerodynamic forces.

Error measurements associated to the angle of attack. Since the slideslip angle is zero, α is given by

$$\alpha = \theta + \gamma_a + \Delta\alpha, \tag{12}$$

where θ [deg] is the pitch angle of the model, γ_a [deg] is the mean upwash angle of the flow in the wind tunnel, and $\Delta\alpha$ [deg] is a corrective term associated to the wall and support effects. Assuming that θ, γ_a and $\Delta\alpha$ are independent Gaussian random variables, the variance parameter for α is given by

$$\tau_\alpha^2 = \tau_\theta^2 + \tau_{\gamma_a}^2 + \tau_{\Delta\alpha}^2, \tag{13}$$

where τ_θ^2, $\tau_{\gamma_a}^2$ and $\tau_{\Delta\alpha}^2$ are the corresponding noise variance parameters for θ, γ_a and $\Delta\alpha$, respectively. τ_θ^2 is deduced from uncertainty analysis of the pitch angle sensor. $\tau_{\gamma_a}^2$ results from an uncertainty analysis of the procedure used to determine upwash angle, namely the model inversion method, which implies again the pitch angle sensor, and the force balance. $\tau_{\Delta\alpha}^2$ is computed by considering that most of the wall correction in angle of attack stems from lift-induced effect. Therefore, $\tau_{\Delta\alpha}$ is made proportional to the lift coefficient.

Experimental error in dynamic pressure. The dynamic pressure is derived from the value of the total pressure p_i and the static pressure p. These values are themselves deduced from measurements made on reference pressure taps of the wind tunnel, using tunnel calibration laws. Finally, wall and support corrections results in a blockage correction Δq. Finally the dynamic pressure q reads:

$$q = \frac{\gamma}{\gamma-1}\, p \left(\left(\frac{p_i}{p} \right)^{\frac{\gamma-1}{\gamma}} - 1 \right) + \Delta q$$

Hence, defining $\lambda = 1 + \frac{\gamma-1}{2}M^2$, the uncertainty on q can be written as:

$$\tau_q^2 = \left(\frac{\gamma}{2}M^2 + \lambda(\lambda^{-\frac{\gamma}{\gamma-1}} - 1) \right)^2 \tau_{p_i}^2 + \left(1 - \frac{M^2}{2} \right)^2 \tau_p^2 + \tau_{\Delta q}^2.$$

Note that the two first terms result from the propagation of the errors in pressure measurements. The last term is added to account for uncertainties in the tunnel calibration and the wall corrections. Variances $\tau_{p_i}^2$ and τ_p^2 are evaluated thanks to an uncertainty analysis of the pressure sensors involved in the measurements of the reference total and static pressures, and confirmed by considering redundancy in the measurements. $\tau_{\Delta q}$ was chosen to be proportional to q.

Experimental error in force coefficients in wind axes. The expression of the variance of the error on the drag coefficient C_x was already provided in Section 3. Reminding that A and N are the axial and normal aerodynamic forces in model axes (respectively), the lift coefficient is expressed as $C_z = \dfrac{-A\sin\alpha + N\cos\alpha}{qS_{\text{ref}}}$. Consequently, the variance of the error is given by:

$$\tau_{C_z}^2 = \left(\frac{\sin\alpha}{qS_{\text{ref}}}\right)^2 \tau_A^2 + \left(\frac{\cos\alpha}{qS_{\text{ref}}}\right)^2 \tau_N^2 + C_z^2 \left(\frac{\tau_q}{q}\right)^2 + C_x^2 \tau_\alpha^2.$$

The pitching moment coefficient equals to $C_M = \dfrac{M_A}{qS_{\text{ref}}L_{\text{ref}}}$, where M_A is the aerodynamic pitching force. Then the variance of the error is equal to:

$$\tau_{C_M}^2 = \left(\frac{1}{qS_{\text{ref}}L_{\text{ref}}}\right)^2 \tau_{M_A}^2 + C_M^2 \left(\frac{\tau_q}{q}\right)^2.$$