Fusing distributed aerodynamic data using Bayesian Gappy Proper Orthogonal Decomposition

Anna Bertram* and Philipp Bekemeyer†
German Aerospace Center (DLR), Institute of Aerodynamics and Flow Technology, 38108 Braunschweig, Germany

Matthias Held‡
Airbus Operations GmbH, Aircraft Aerodynamics, 28199 Bremen, Germany

During the development of an aircraft, a multitude of aerodynamic data is required for different flight conditions throughout the flight envelope. Nowadays, a large portion of this data is routinely acquired by Computational Fluid Dynamics simulations. However, due to modeling and convergence issues especially for extreme flight conditions, numerical data cannot be reliably generated throughout the entire flight envelope yet. Hence, numerical data is complemented by data from wind tunnel experiments and flight testing. However, the data from these different sources will always show some discrepancies to deal with. Data fusion methods aim at combining the individual strengths and weaknesses of data from different sources in order to provide a consistent data set for the entire parameter domain. In this work we propose an extension to the well established gappy proper orthogonal decomposition technique by interpreting the occurring least-squares problem as a regression task. A Bayesian perspective is imposed to account for uncertainties during the data fusion process. This involves a kernelized regression formulation which also leverages the problem of linearity imposed by the dimensionality reduction method. We demonstrate the performance and robustness of the approach investigating an industrial-relevant, large-scale aircraft test case fusing high quality experimental and numerical data.

I. Introduction

During the development of an aircraft, a wealth of aerodynamic data is required for different flight conditions throughout the flight envelope. Quantities like the scalar-valued lift coefficient or the pressure distribution at the surface are, for instance, significant for structural and geometrical design, performance and loads evaluations and the design of the flight control system. In the last decades, this data has been increasingly acquired through Computational Fluid Dynamics (CFD) simulations [1]. However, due to the complexity of the problem, a direct numerical simulation for industrial aircraft configurations is not even feasible for high-performance computers in the foreseeable future. Simplifying assumptions are made to ease the problem, but CFD simulations remain computationally demanding. In addition, due to convergence issues especially for extreme flight conditions, numerical data cannot be reliably provided throughout the entire flight envelope. CFD data is therefore complemented by data from wind tunnel experiments and flight testing. Because of errors mainly introduced by the physical modeling and the discretization of the problem on the one hand and experimental limitations on the other hand, the data from these different sources will, however, always show some differences to deal with. The heterogeneity in the data becomes even more complex when, instead of scalar-valued quantities, field data has to be considered—as often required by industry: For example, for the structural design of an aircraft, one is interested in reliably predicting the location of aerodynamic shocks as indicated by the surface pressure distribution. While CFD provides this quantity in each cell on the surface of the discretized aircraft, it is typically only available at comparatively few sensor locations in wind tunnel and flight tests which results in a high discrepancy in the dimensions of computational and experimental data.

Data fusion techniques aim at combining the individual strengths of different data sources to provide consistent and reliable data sets. For scalar-valued quantities of interest, a popular data fusion strategy is the use of variable-fidelity surrogate models. A base assumption of these models is that the different data sources have different levels of accuracy. The simplest way to fuse such variable-fidelity data is by employing so-called bridge functions: A surrogate model for

*Research Scientist
†Team Leader
‡Expert Aerodynamic Modelling for Loads

the low-fidelity data is used to approximate the high-fidelity data via an additive, multiplicative or hybrid correction, [2]. Another example of data fusion with bridge function is given in [3] where data from two computer codes of different fidelity were combined via a Kriging-based bridge function. Kriging, also known as Gaussian process regression, is based on the assumption that the given data points are realizations of correlated random variables. It can handle highly nonlinear responses and features fast evaluation times which makes it a popular method in various different fields such as design and analysis of computer experiments [4, 5], machine learning [6] and surrogate modeling [7]. Direct extension of Gaussian process regression to the variable-fidelity framework are known as Cokriging [7, p. 177], [8–11], and Hierarchical Kriging, [12]. Both approaches have been used in the context of aerodynamic applications for variable-fidelity modeling based on different sources of computational data, [7, 10, 12–15], and for experimental and numerical data, [16]. A non-hierarchical approach for the fusion of scalar-valued quantities based on Gaussian process regression was recently presented in [17]: A Gaussian process model is built for every data source individually. Afterwards, these models are combined via a weighted sum based on modeling uncertainty and expert knowledge on the model fidelity.

For the surrogate modeling of vector-valued quantities based on data of different levels of fidelity, the variable-fidelity methodology was extended in [18, 19]. The authors considered aerodynamic data from computer simulations of different accuracy on the same computational grid in order to construct a surrogate model for the expensive, high-fidelity computer code. A common orthonormal basis was computed employing the dimensionality reduction technique Proper Orthogonal Decomposition (POD), [20, 21]. The scalar-valued basis coefficients of this POD basis are then interpolated via a variable-fidelity surrogate model like Cokriging or Hierarchical Kriging. Another data fusion approach for vector-valued quantities was recently proposed in [17], where in a Bayesian setting, experimental and numerical data to the same aerodynamic flow conditions and the same spatial grid were combined via a weighted sum. A drawback of these method is however that low- and high-fidelity data have to be given at the same spatial grid which prevents their application for the fusion of sparse experimental sensor data and high-dimensional numerical data.

The latter can be achieved with the so-called Gappy POD method. Gappy POD combines POD with a least-squares problem to reconstruct not completely known data. The idea is that within the subspace spanned by the POD modes a solution can be found which minimizes the differences to reference data at a few discrete locations in a least-square sense. The method was first developed for the reconstruction of human face images from incomplete data sets [22]. Later, the method was extended to fluid dynamic applications in [23], where it was shown that it can be used to reconstruct missing data in CFD snapshots based on a set of complete CFD snapshots for steady aerodynamic flow around an airfoil. Modifications of the original approach have been proposed in [24] in order to enhance its robustness and effectiveness for the reconstruction of spatio-temporal incomplete flow field data. A comparison of Gappy POD and the statistical interpolation method Kriging for the reconstruction of unsteady aerodynamic flow data was done in [25]. In [26], the method used for the reconstruction of unsteady flow data and, based on this procedure, a strategy for optimal sensor placement was introduced. To avoid overfitting, especially when dealing with real experimental data, a regularized version of Gappy POD was introduced in [27]. The authors compared different regularization methods and applied their regularized Gappy POD approach to fuse CFD and experimental surface pressure data from a steady aerodynamic flow around the flap of a transport aircraft. The resulting surface pressure distribution matches the experimental values as close as possible while providing dense information on the whole surface as a CFD analysis would. The approach was further extended such that the fused aerodynamic surface data sums up to the overall integral coefficients as measured by the wind tunnel balance in [28]. Technically this has been achieved by extending the regression to account for an equality constraint. One of the limitations of the aforementioned approach stems from the assumption of linearity within the POD model and the consecutive least-squares problem. Furthermore, the acceptance of such data-driven approaches is often limited by the fact that it is hard to estimate the uncertainty in the predicted data.

In this work, we address these issues by proposing a Bayesian extension for Gappy POD. It is based on the equivalence of the Gappy POD problem of least-squares with a linear regression task for a specific choice of sample data. By taking the regression perspective, we solve the Gappy POD problem by employing Gaussian process regression. The final data fusion result is then given in terms of a probability distribution, which provides valuable information on the predicted modeling accuracy. The limitations of the ordinary approach are overcome by considering nonlinear covariance functions. The approach is demonstrated on an industrial-relevant aerodynamic test case where it can be shown that the nonlinear approach leads to a significant improvement of the data fusion result in terms of agreement with the experimental data.

This paper is structured as follows. First, a theoretical introduction to ordinary and regularized Gappy POD is provided in Section II.A and II.B, followed by a description of the new Bayesian extension in Section II.C. In Section III.A, the use case and simulation environment is presented. Results of the data fusion process are discussed

and compared to regularized Gappy POD and an approach solely relying on CFD in Section III.B. We close this paper by giving a conclusion and an outlook in Section IV.

II. Data Fusion Methodology

In this section the mathematical basics of the Gappy POD data fusion methodology is described. We will briefly review the ordinary and regularized Gappy POD approach in subsection II.A and II.B based on the work of [27–29]. Afterwards, the new Bayesian regression extension is introduced in subsection II.C.

A. Ordinary Gappy POD

Assume that there is a functional dependency between the set of input parameters $x \in \mathcal{D} \subset \mathbb{R}^d$ and the resulting steady-state surface pressure coefficient distribution on the computational grid, $y\colon \mathcal{D} \to \mathbb{R}^N$. This functional dependency is not explicitly given but can be evaluated via CFD simulations for different input parameters $\mathcal{P} = \{\xi^1, \ldots, \xi^n\} \subset \mathcal{D}$ to obtain sampled data. The resulting outputs are stored in the *snapshot matrix* $Y \in \mathbb{R}^{N\times n}$,

$$Y \coloneqq [y^1, \ldots, y^n] = [y(\xi_1), \ldots, y(\xi_n)]\,. \tag{1}$$

For simplicity of the notation and without loss of generality assume that the snapshots are centered with respect to their mean, that is

$$\sum_{i=1}^{n} y^i = 0. \tag{2}$$

Performing a singular value decomposition (SVD) [30, Sec. 2.4] of the snapshot matrix yields

$$Y = U\Sigma V^{\mathrm{T}}, \tag{3}$$

where $U = [u^1, \ldots, u^N] \in \mathbb{R}^{N\times N}$ and $V = [v^1, \ldots, v^n] \in \mathbb{R}^{n\times n}$ are orthonormal matrices, i.e. $U^{\mathrm{T}}U = UU^{\mathrm{T}} = I_N$ and $V^{\mathrm{T}}V = VV^{\mathrm{T}} = I_n$, and $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{N\times n}$ contains the singular values $\sigma_1 \geq \ldots, \geq \sigma_n \geq 0$ in descending order. Suppose the rank of the snapshot matrix Y is $r = \operatorname{rank}(Y)$, then only the first $r \leq n$ singular values are non-zero. The corresponding r left singular vectors, which are the first r columns of the matrix U, constitute an orthonormal basis $\{u^1, \ldots, u^r\}$ of the space spanned by the snapshots $y^1, \ldots, y^n$, the so-called *POD basis*, [30, Corollary 2.4.6].

A key idea of Gappy POD is to interpret a given vector $t \in \mathbb{R}^s$ of experimental data, where $s < N$ is the number of experimental sensors, as a vector $y \in \mathbb{R}^N$ from which only the components $y_{j_1}, \ldots, y_{j_s}$ with $j_1, \ldots, j_s \in \{1, \ldots, N\}$ are known, that is

$$t = \begin{bmatrix} t_1 \\ \vdots \\ t_s \end{bmatrix} = \begin{bmatrix} y_{j_1} \\ \vdots \\ y_{j_s} \end{bmatrix} = P^{\mathrm{T}} y \tag{4}$$

for a mask matrix $P \coloneqq [e_{j_1}, \ldots, e_{j_s}] \in \mathbb{R}^{N\times s}$. The components of the vector t are identified with the components of the vector y via a nearest neighbor search of the s sensor coordinates in the highly resolved computational grid.

Assuming that the vector y can be approximated in the POD subspace, POD basis coefficients $\hat{a} = (\hat{a}_1, \ldots, \hat{a}_r)^{\mathrm{T}} \in \mathbb{R}^r$ can be found such that

$$y \approx \hat{y} = \sum_{j=1}^{r} \hat{a}_j u^j = U_r \hat{a}, \tag{5}$$

where $U_r = [u^1, \ldots, u^r] \in \mathbb{R}^{N\times r}$ is the matrix of POD basis vectors, i.e. the first r columns of U. The basis coefficient vector $\hat{a} \in \mathbb{R}^r$ which yields the smallest L_2 error regarding the observed entries of the vector y is defined by the least squares problem

$$\hat{a} = \arg\min_{a} \left\| P^{\mathrm{T}} U_r a - t \right\|_2^2. \tag{6}$$

Usually, $X = P^{\mathrm{T}}U_r \in \mathbb{R}^{s\times r}$ has full column rank and therefore Eq. (6) has a unique solution given by

$$\hat{a} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}t. \tag{7}$$

Substituting this basis coefficient vector $\hat{a}$ into Eq. (5) yields the ordinary Gappy POD approximation of the vector y.

B. Regularized Gappy POD

To avoid overfitting, especially when dealing with data from different sources, it is often necessary to complement the least squares problem with regularization terms on the basis coefficients $\hat{a}$, [27, 28]. Shrinkage methods give preference to smaller basis coefficients $\hat{a}$ by imposing a penalty on their value, cf. [31, Sec. 3]. The influence of less important POD modes, i.e. basis vectors which correspond to small singular values, can thus be restricted by scaling the basis vectors with their corresponding singular values

$$\tilde{U} = [\tilde{u}^1, \dots, \tilde{u}^r] \coloneqq [\sigma_1 u^1, \dots, \sigma_r u^r] \tag{8}$$

A popular choice for the regularization of least squares problems is ridge regression, sometimes also called Tikhonov regularization, [31, Sec. 3.4.1], [32]. It imposes a L_2 penalty on the basis coefficient vector a. The corresponding Gappy POD problem reads,

$$\hat{a}_{\mathrm{rr}} = \arg\min_a \left\| P^{\mathrm{T}} \tilde{U} a - t \right\|_2^2 + \lambda \|a\|_2^2, \tag{9}$$

where $\lambda \in \mathbb{R}_+$ is a parameter which controls the strength of the regularization: the larger the value of λ, the higher the amount of shrinkage towards 0. If $X = P^{\mathrm{T}} \tilde{U}_r$ has full column rank, the Gappy POD problem has a unique solution,

$$\hat{a}_{\mathrm{rr}} = (X^{\mathrm{T}} X + \lambda I)^{-1} X^{\mathrm{T}} t. \tag{10}$$

As in the ordinary case, the Gappy POD approximation of the vector y is obtained by evaluating the corresponding linear combination of POD basis modes,

$$\hat{y} = \sum_{j=1}^{r} \hat{a}_j u^j = \tilde{U}_r \hat{a}_{\mathrm{rr}}, \tag{11}$$

cf. Eq. (5).

C. Bayesian Gappy POD Extension

The Gappy POD problem Eq. (6) can be interpreted as a linear regression problem with sample data

$$\{(x_i, t_i) \mid i = 1, \dots, s\}, \tag{12}$$

where $x_i \coloneqq (X)_i = (P^{\mathrm{T}} U_r)_i = (U_r)_{j_i} \in \mathbb{R}^r$ denotes the i-th row of the matrix $X = P^{\mathrm{T}} U_r$ or, in other words, the j_i-th row of the matrix of POD modes U_r and $t_i = (t)_i$ is the corresponding sensor response. In this setting, we aim at evaluating the regression model f for all rows of the matrix of POD modes $(U_r)_i, i = 1, \dots, N$, to obtain the Gappy POD solution,

$$\hat{y} = f(U_r) = \left(f\left((U_r)_i \right) \right)_{i=1,\dots,N}. \tag{13}$$

In case of ordinary and regularized Gappy POD, the standard linear regression model,

$$f(x) = x^{\mathrm{T}} w, \tag{14}$$

with weight vector $w \in \mathbb{R}^r$ is considered. In this work, we take a Bayesian perspective on the linear regression problem Eq. (12) which allows us to derive probability distributions for the data fusion result. While in the ordinary Gappy POD formulation, only point estimates are provided as data fusion result, this new perspective allows to provide valuable information about the estimated accuracy of the prediction.

In order to give a better insight into our method, we will first review some basics of linear regression from a statistical point of view based on [6, Sec. 2]. This leads us to the concept of Gaussian Process Regression which will then be used to solve the regression problem Eq. (12) in our Bayesian Gappy POD extension.

Suppose we are given a training data set $\{(x_i, t_i) \mid i = 1, \dots, s\}$ with input variables $x_i \in \mathbb{R}^r$ and corresponding outputs $t_i = t(x_i), i = 1, \dots, s$. Our goal is to model the relationship between inputs and outputs. The linear regression model with Gaussian noise reads,

$$f(x) = \phi(x)^{\mathrm{T}} w, \qquad t(x) = f(x) + \varepsilon, \tag{15}$$

with a fixed set of basis functions $\phi(x) = (\phi_1(x), \dots \phi_m(x))^{\mathrm{T}}$, corresponding weight vector $w \in \mathbb{R}^m$, and additive independent, identically distributed Gaussian noise ε with zero mean and stationary variance $\sigma^2 > 0$. The likelihood

function, i.e. the probability density of the observations given the weights, is given by

$$\begin{aligned} p(t \mid X, w) &= \prod_{i=1}^{s} p(t_i \mid x_i, w) \\ &= \frac{1}{\sqrt{(2\pi\sigma^2)^s}} \exp\left(-\frac{1}{2\sigma^2}\|t - \Phi^{\mathrm{T}} w\|_2^2\right), \end{aligned} \tag{16}$$

where $t = (t(x_1), \ldots, t(x_s))^{\mathrm{T}} = (t_1, \ldots, t_s)^{\mathrm{T}}$ is the vector of observed outputs and

$$\Phi = [\phi(x_1), \ldots, \phi(x_s)] = \begin{bmatrix} \phi_1(x_1) & \ldots & \phi_1(x_s) \\ \vdots & \ddots & \vdots \\ \phi_m(x_1) & \ldots & \phi_m(x_s) \end{bmatrix} \in \mathbb{R}^{m \times s} \tag{17}$$

is the matrix of function values of the basis functions $\phi(x)$ evaluated at the sample points $x_1, \ldots, x_s$.

Instead of determining a specific weight vector w, we express our beliefs about the probability distribution of the weight vector w before seeing the data in terms of a *prior distribution* $p(w)$. Based on this prior distribution and the likelihood function, we can now derive the *posterior distribution* of w, that is the probability distribution of w after seeing the data by making use of Bayes' Theorem, cf. [6, Sec. A.1],

$$\text{posterior} = \frac{\text{likelihood} \cdot \text{prior}}{\text{marginal likelihood}}, \tag{18}$$

$$p(w \mid X, t) = \frac{p(t \mid X, w) \cdot p(w)}{p(t \mid X)}, \tag{19}$$

where the *marginal likelihood*,

$$p(t \mid X) = \int p(t \mid X, w) p(w) \, \mathrm{d}w, \tag{20}$$

is a normalizing constant independent of the weight vector w.

Finally, predictions for the output $f_* \coloneqq f(x_*)$ for a new input variable $x_* \in \mathbb{R}^r$ can be made by evaluating the *predictive distribution*,

$$p(f_* \mid x_*, X, t) = \int p(f_* \mid x_*, w) p(w \mid X, t) \, \mathrm{d}w. \tag{21}$$

One can show that for different choices of prior distributions $p(w)$, the maximum of the predictive distribution equals the prediction of well-known regression techniques such as ordinary linear regression, ridge regression and LASSO, cf. [33, Tab. 7.1, p. 226]. However, the Bayesian point of view enables to go a step further: As we consider the weight vector as a statistical quantity, the output f can be viewed as a statistical quantity as well. Thus, instead of providing a regression model which is a point estimate of the probability distribution for the input-output mapping $f(x)$, we can provide the entire distribution of probable regression models.

Except for some special choices of prior distributions $p(w)$, there is no explicit expression for the predictive distribution in Eq. (21). One of those special choices, which allows for an analytical expression of the predictive distribution, is the assumption of a Gaussian prior,

$$p(w) \sim \mathcal{N}(0, \Sigma_w). \tag{22}$$

Applying some calculus one can show that in this case the posterior distribution of the weight vector w and, consequently, the predictive distribution of f_* are also Gaussian with

$$\mathrm{E}[f_*] = \phi_*^{\mathrm{T}} \Sigma_w \Phi \left(\Phi^{\mathrm{T}} \Sigma_w \Phi + \sigma^2 I\right)^{-1} t, \tag{23}$$

$$\mathrm{Var}[f_*] = \phi_*^{\mathrm{T}} \Sigma_w \phi_* - \phi_*^{\mathrm{T}} \Sigma_w \Phi \left(\Phi^{\mathrm{T}} \Sigma_w \Phi + \sigma^2 I\right)^{-1} \Phi^{\mathrm{T}} \Sigma_w \phi_* \tag{24}$$

where $\phi_* \coloneqq \phi(x_*)$, [6, p. 17]. Defining the function

$$k(x, x') = \phi(x)^{\mathrm{T}} \Sigma_w \phi(x'), \tag{25}$$

we can rewrite these expressions in terms of the function k,

$$\mathrm{E}[f_*] = k(x_*)\left(K + \sigma^2 I\right)^{-1} t, \tag{26}$$

$$\mathrm{Var}[f_*] = k(x_*, x_*) + \sigma^2 - k(x_*)^{\mathrm{T}}\left(K + \sigma^2 I\right)^{-1} k(x_*), \tag{27}$$

where $k(x_*) \coloneqq \left(k(x_*, x_i)\right)_{i=1,\ldots,s} \in \mathbb{R}^s$ and $K \coloneqq \left(k(x_i, x_j)\right)_{i,j=1,\ldots,s} \in \mathbb{R}^{s\times s}$. It is straight forward to show that $f(x)$ defines a zero-mean Gaussian process whose covariance $\mathrm{Cov}[f(x), f(x')]$ for any two inputs $x, x' \in \mathbb{R}^r$ is given by $k(x, x')$,

$$f(x) \sim \mathcal{GP}(0, k(x, x')). \tag{28}$$

Because of this property, the function $k(x, x')$ is also called *covariance function*, [6, p. 12].

In the above setting, a set of basis functions $\phi(x)$ has been chosen and assumptions on the prior distribution of the weight vector w have been made, which implicitly define the covariance of the Gaussian process $f(x)$. The key idea of Gaussian Process Regression, also known as Kriging, is to bypass this concept by defining a positive definite covariance function for $f(x)$ explicitly, [34, p. 160]. In turn, this implicitly defines a (potentially infinite) set of basis functions $\phi(x)$ due to Mercer's theorem, [33, p. 483]. Of course, the choice of the covariance function has a large impact on the predictions. It is typically chosen such that it reflects the property that the outputs $f(x)$ and $f(x')$ of input variables x and x' that are *close* or *similar* to each other are more strongly correlated than the outputs of more distinct input variables, [34, Sec. 6.4.2]. A widely used class of covariance functions is given by

$$k(x, x') = \theta_0 \cdot \exp\left(-\theta_1 \|x - x'\|^2\right) + \theta_2 x^{\mathrm{T}} x' + \theta_3, \tag{29}$$

cf. Eq. (6.63) in [34, Sec. 6.4, p. 307], with hyperparameters $\theta_0, \ldots, \theta_3$. Instead of defining the hyperparameters and the noise variance σ^2 in advance, they are usually determined from the data by maximizing the log marginal likelihood function,

$$\ln p(t \mid X) = -\frac{1}{2} t^{\mathrm{T}} (K + \sigma^2 I)^{-1} t - \frac{1}{2} \ln \det(K + \sigma^2 I) - \frac{s}{2} \ln 2\pi. \tag{30}$$

For a detailed discussion of covariance functions for GPRs and different methods for hyperparameter estimation, the reader is referred to Sec. 4 and Sec. 5.4 of the textbook [6].

The derivation of Gaussian Process Regression from linear regression with a special choice of prior information motivates our Bayesian regression extension for Gappy POD: We propose to solve the Gappy POD problem Eq. (12) by employing GPRs. A kernel function $k(x, x')$ is chosen from the class Eq. (29) such that the log marginal likelihood function Eq.(30) is maximized. Afterwards, a predictive distribution for all rows of the matrix of POD modes, $(U_r)_i, i = 1, \ldots, N$, is obtained by evaluating Eq. (26) and Eq. (27).*

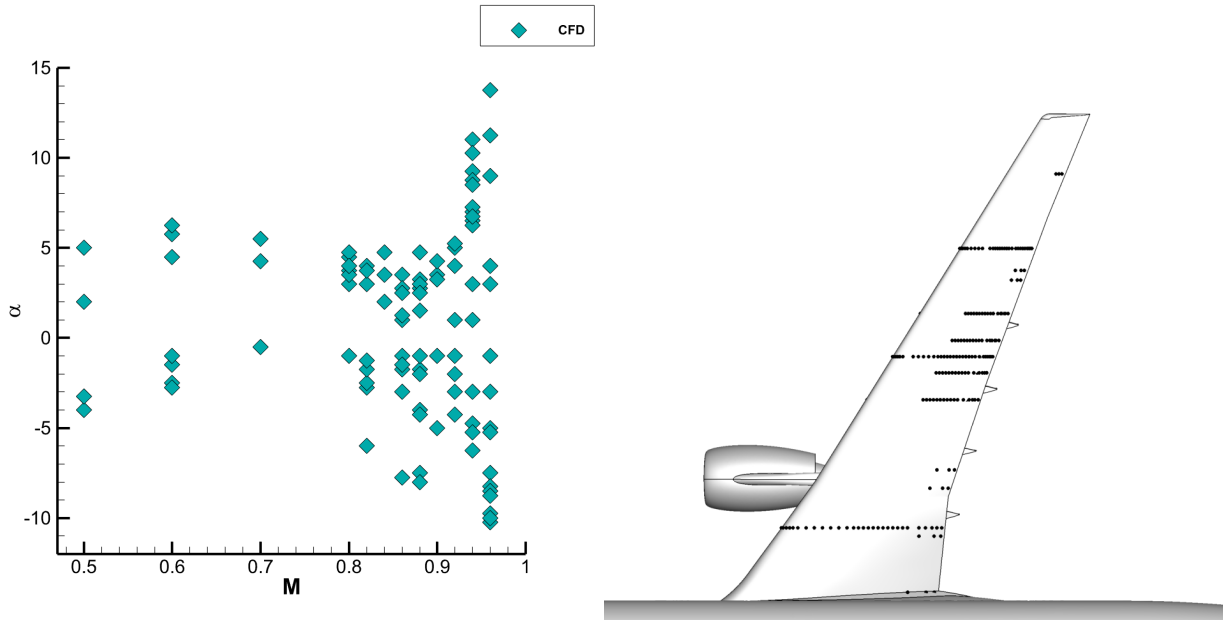
Note that the final data fusion result will in general not lie in the POD subspace, as is the case for ordinary or regularized Gappy POD. Consequently, no vector of POD basis coefficients can be found.

III. Application to a Transport Aircraft Test Case

A. Case study description

The test case configuration considered here is an industrial-relevant aircraft configuration known as XRF1. The XRF1 is an Airbus provided industrial standard multi-disciplinary research test case representing a typical configuration for a long range wide body aircraft. It is used by Airbus to engage with external partners on development and demonstration of relevant capabilities and technologies. For the computational data, a highly resolved CAD model of the XRF1 is considered as half configuration. The corresponding surface grid consists of $N = 388,918$ grid points. High-fidelity RANS-CFD simulations were carried out with the DLR flow solver TAU, [35] using the SST turbulence model. Structural deformation are accounted for by coupling the CFD analysis to a computational structural mechanics (CSM) investigation which relies on a finite element model of the wind tunnel aircraft structure. In this way, a total number of $n = 100$ pressure coefficient (c_p) distributions on the surface of the aircraft were computed for different Mach numbers $\mathrm{M} \in [0.5, 0.96]$ and angles of attack $\alpha \in [-11°, 14°]$ and are considered as CFD snapshots during the following

*In this setting each entry of the Gappy POD solution $(\hat{y})_i, i = 1, \ldots, N$ is a random variable.



(a) CFD data base

(b) Pressure tap locations

Fig. 1 Data-points which are sampled during the numerical analysis as well as the location of pressure taps on the upper wing surface

investigations. The sampling points are displayed in Figure 1(a). The Reynolds number was fixed to Re = $25 \cdot 10^6$. Wind tunnel tests for the XRF1 configuration were carried out in the European Transonic Windtunnel (ETW). The surface pressure coefficient was obtained at 314 pressure taps distributed over 26 section cuts along the wing with locations shown in Figure 1(b) for a number of 196 different combinations of Mach number and angle of attack.

The sampled parameter combinations cover a wide range of aerodynamic phenomena including fully attached flow, severe trailing edge separation, strong shock waves on the upper and lower wing surface as well as shock induced separation. Based on the specific configuration at hand and observed aerodynamic characteristics, the Mach number vs angle of attack diagram has been divided in different region as shown in Figure 2. In particular these are the design range, the linear region and the nonlinear region. Note that, the XRF1 is an aircraft concept provided from Airbus for the means of demonstration of relevant capabilities and technologies and displayed regions should be regarded as a generic classification of flow phenomena based on common aerodynamic knowledge. Moving from the design region towards the edge of the envelope it is expected that the prediction accuracy of numerical simulation tools decreases [36]. Hence, discrepancies between numerical and experimental results are likely to grow. This should directly result in enlarged credible intervals for fused pressure distributions at more challenging conditions.

B. Results

Different points within the envelope are chosen to provide an insight into the data fusion capabilities and are discussed next. Especially within the nonlinear region, towards the edge of the envelope, information provided by CFD snapshots is sparse and likely also inaccurate due to various simulation challenges including turbulence modeling. For all presented cases, data fusion was employed using regularized Gappy POD with ridge regression, cf. II.B, and the new Bayesian Gappy POD approach with Gaussian Process Regression as introduced in Sec. II.C. In the latter case, the kernel function was chosen from the class Eq. (29), where the hyperparameters $\theta_0, \ldots, \theta_3$ and the noise variance σ^2 were determined by maximizing the marginal likelihood function Eq. (30). Note that the noise variance σ^2 is assumed to be constant for all pressure taps and is determined purely from the data without taking any prior knowledge on particular sources of uncertainty into account. Some uncertainties like the measurement uncertainties due to sensor inaccuracy or instrumentation may be (in parts) quantified by manufacturer's specifications or expert knowledge. Furthermore, we assume the noise variance σ^2 to be stationary meaning that the measurements from all pressure taps are corrupted by the same type of noise. Although it is in general possible to incorporate more knowledge on uncertainties into the GPR

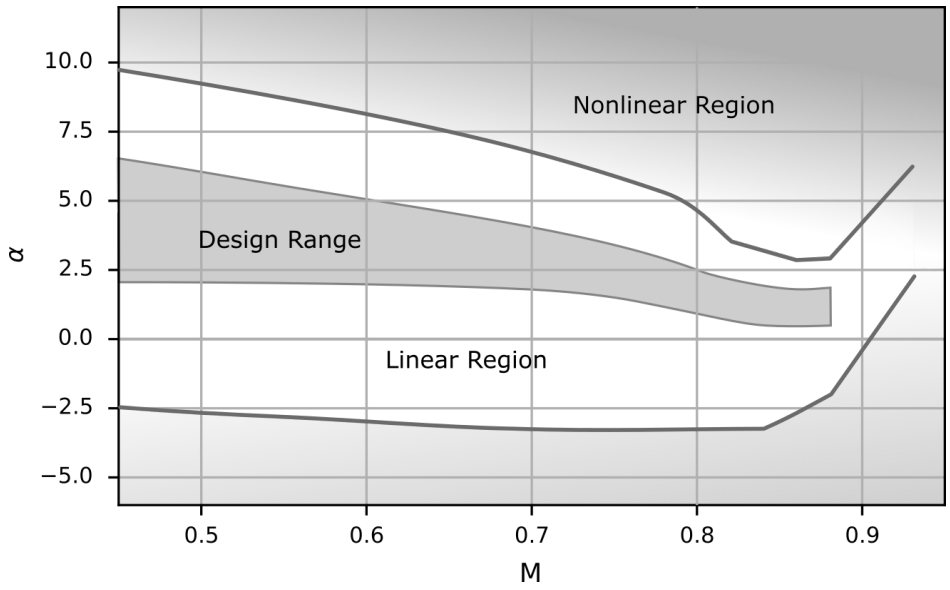


Fig. 2 XRF1 envelope zones, kindly provided by Airbus.

modeling, we neglected these points herein and propose a detailed investigation as a potential topic for future work.

In addition to the data fusion results, a surrogate model was constructed solely based on the same CFD data set by interpolating the POD basis coefficients—a widespread data-driven surrogate modeling technique introduced by [37]. For the interpolation of the basis coefficients, thin plate spline (TPS) interpolation, [38], was considered. For comparison, the CFD-based surrogate model was evaluated at the flow conditions of interest.

For a quantitative assessment of the agreement with the wind tunnel sensor measurements, the results $\hat{y}$ of every investigated method were evaluated by means of the root mean squared error,

$$\mathrm{RMSE}(\hat{y}) \coloneqq \sqrt{\frac{1}{s}\|P^{\mathrm{T}}\hat{y} - t\|_2^2}. \tag{31}$$

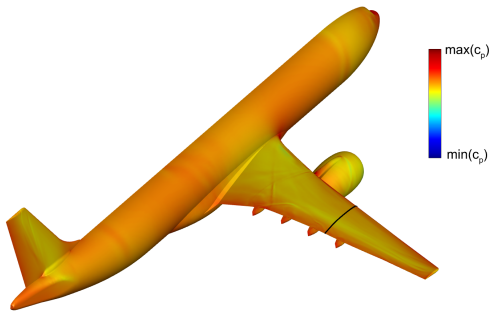
Note that the Bayesian Gappy POD approach introduced in Sec. II.C, yields a predictive distribution instead of a point estimate that is given by the other methods considered in this study. In every surface grid point, the data fusion result is a Gaussian distributed random variable with a certain mean and variance. In the following, we account for this distribution by always showing the predicted mean together with its standard deviation, i.e. the square root of the variance, or giving 95 % credible intervals.[†] As the other investigated methodologies (CFD, POD+TPS and gappy POD based on Ridge regression) are purely deterministic, they only provide mean values for comparison.

The first analyzed flow condition is $\mathrm{M} = 0.50, \alpha = -1.97°$ which is within the linear region. The mean and the standard deviation of the surface pressure distribution obtained from the Gappy POD approach using GPR is shown in 3(a) and 3(b), respectively, while a comparison between different methods in a section cut at $\eta = 0.55$ (indicated by a solid black line in the surface plots) is displayed in 3(c).

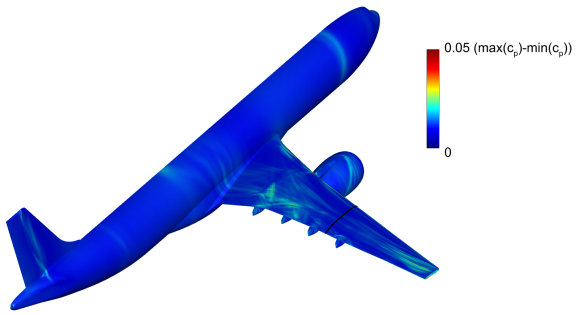
A suction peak is observed on the leading edge of the wing slightly shifted towards the lower side of the wing due to the negative angle of attack. The rest of the wing exhibits are smooth pressure distribution as expected at subsonic flow conditions. When comparing the prediction accuracy for the selected section cut, no differences are observed between both gappy POD methods (Ridge regression and GPR) whereas the purely CFD-based POD + TPS methodology slightly over-predicts the pressure levels on the lower wing surface. The 3D view of the standard deviation in Figure 3(b) shows that it is the smallest in areas of very uniform flow (e.g. fuselage) or in areas where many wind tunnel sensors are located. In areas of non-uniform flow and only a few available measurements (e.g. behind the engine or close to the wing tip) an increase of the standard deviation is observed. This corresponds to a less smooth pressure distribution as displayed in Figure 3(a).

Additional results in the transonic regime are shown in Figure 4.

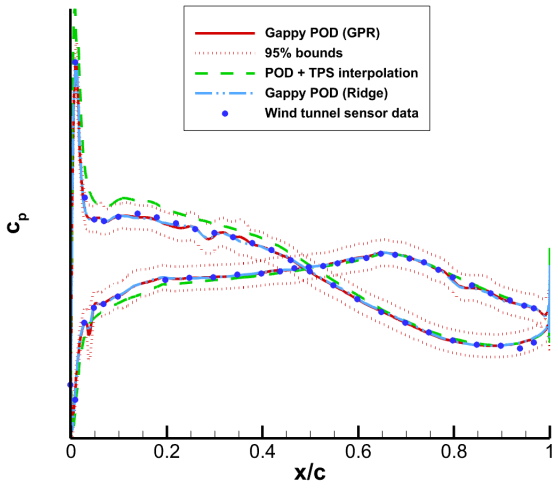
[†]95 % credible intervals correspond to the mean plus/minus 1.96 times standard deviation.



(a) Mean of the predictive distribution of the pressure coefficient on the surface of the aircraft.



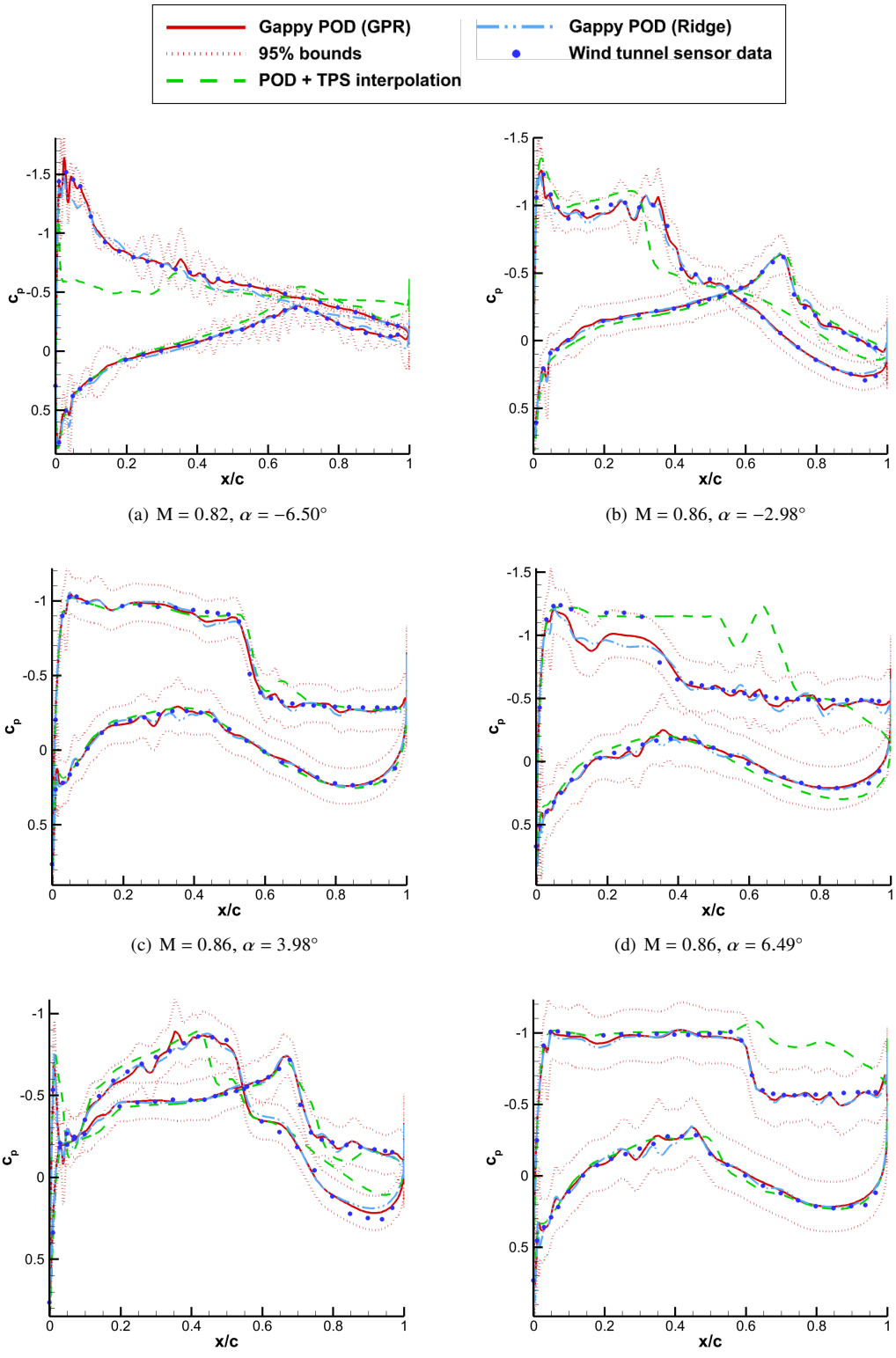
(b) Standard deviation of the predictive distribution of the pressure coefficient on the surface of the aircraft.



(c) Pressure coefficient distribution at selected cut.

Fig. 3 Pressure coefficient approximation for M $= 0.50, \alpha = -1.97°$ via Gappy POD combined with ridge regression and GPR in comparison to POD plus TPS interpolation at a wing section cut where wind tunnel sensor data is available. The solid black line indicates the cut location.

Figure 4(a) shows the data for the case M $= 0.82, \alpha = -6.50°$, which is well within the negative nonlinear region. As in the first case a rather smooth pressure distribution is observed, with the Gappy POD with GPR accurately reproducing the wind tunnel data. While the Gappy POD with Ridge regression approach is able to reproduce the overall trend it fails to exactly match the c_p level on the lower surface. The result of the simple POD + TPS interpolation is far off the wind tunnel test result. This can be attributed to the underlying CFD data where the closest data point features a leading edge separation on the lower surface. In comparison to the section cut plot of the first analyzed parameter combination, Figure 3(c), a completely different behavior of the 95 % credible bounds are observed. While in the first investigated case the predictive variance is dominated by the noise variance, it is strongly driven by the process variance of the GPR model in the second case. Here, the hyperparameter optimization of the covariance function yields a comparatively small noise variance. As a result, the credible intervals are very narrow at the pressure tap locations and get significantly larger apart from the data. These two cases demonstrate that the optimization of the hyperparameters is a challenging task and crucial for resulting variances. The loss function Eq. (30) is often multimodal and thus has multiple local maxima, cf. [6, Sec. 5.4]. For example, by looking only at the data, it is often not clear if the fluctuation in the data originates either from the true relationship of input and output or from noise. More data or information about the uncertainty associated with the data may help to improve the hyperparameter tuning. A detailed discussion is however beyond the scope of this work. Because of this different behavior, additional surface plots of the mean and standard



(a) M = 0.82, $\alpha = -6.50°$

(b) M = 0.86, $\alpha = -2.98°$

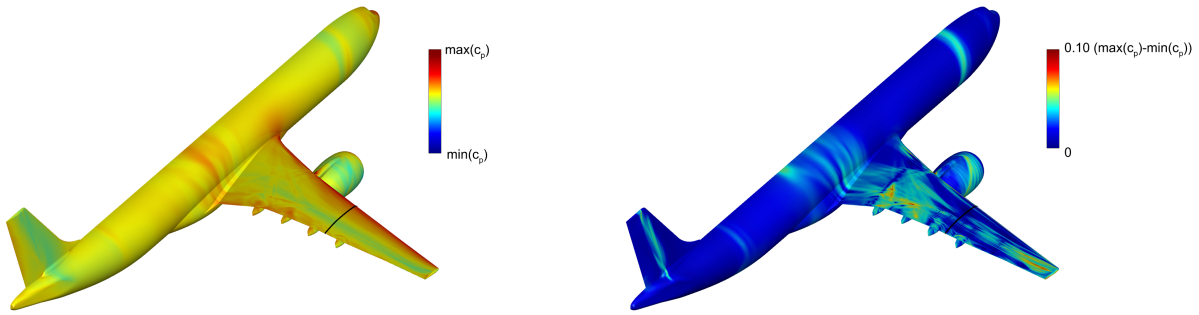
(c) M = 0.86, $\alpha = 3.98°$

(d) M = 0.86, $\alpha = 6.49°$

(e) M = 0.92, $\alpha = 0.02°$

(f) M = 0.92, $\alpha = 7.01°$

Fig. 4 Pressure coefficient approximation via Gappy POD combined with ridge regression and GPR in comparison to POD plus TPS interpolation at the selected wing section cut.



(a) Mean of the predictive distribution of the pressure coefficient on the surface of the aircraft.

(b) Standard deviation of the predictive distribution of the pressure coefficient on the surface of the aircraft.

Fig. 5 Mean and standard deviation of pressure coefficient distribution for M = 0.82, $\alpha = -6.50°$ as estimated with Bayesian Gappy POD.

deviation are given in Figure 5.

As indicated by the section cut plot in Figure 4(a), the surface plot of the standard deviation in Figure 5(b) shows the characteristic increase of the standard deviation in between the pressure taps leading to greenish streaks especially on the wing. On section cuts where many pressure taps are located, these streaks are interrupted by dark blue lines indicating that the standard deviation in this regions is estimated to be small. In addition, a larger increase of the standard deviation in areas of non-uniform flow and stronger recompression (e.g. on the vertical tail plane and forward fuselage) is observed. This can be interpreted that the uncertainty in the model prediction in these regions is estimated to be high. This is partly in line with our expectations on the corresponding uncertainty in the data fusion result as from an aerodynamic point of view it is indeed difficult to estimate accruing flow phenomena in these regions.

Another case in the negative nonlinear region is presented in 4(b). Note that, due to the negative angle of attack, a shock wave is present on the lower side of the wing. Both Gappy POD approaches match the c_p levels of the wind tunnel sensor data including a good representation of both (upper and lower surface) shocks in terms of strength and location. However, not all sensor data are accurately matched, leading to a relative increase of the 95 % bounds. Furthermore some unphysical oscillations especially for the pressure distribution on the lower surface can be observed. The CFD-based POD + TPS surrogate model at the same time shows a good match of the upper surface pressure distribution, while significant differences to the wind tunnel sensor data can be observed for the lower wing surface. Especially the shock position is too close to the leading edge. 4(c) shows a case at a positive angle of attack of $\alpha = 3.98°$ and $Ma = 0.86$. Both Gappy POD approaches accurately reproduce the characteristics of the pressure distribution as measured in the wind tunnel test, including a supersonic plateau and a strong shock on the upper surface. The c_p plateau downstream of the shock, indicating a shock induced separation, is matched as well. The POD + TPS interpolation however fails to accurately reproduce the shock. Further increasing the angle of attack to $\alpha = 6.49°$ (4(d)) and hence extrapolating beyond the range of available CFD data, both applied Gappy POD approaches fail to match the c_p level in front of shock. In contrast to the POD + TPS extrapolation, the shock location however is matched. The 95 % credible interval is increased, showing the increased uncertainty of this data fusion result. The final two data points demonstrate the capability of the presented data fusion method towards the edge of the flight envelope, close to the dive Mach number. In 4(e) the data for M = 0.92, $\alpha = 0.02°$ is presented. A number of different aerodynamic features (leading edge suction peak, shocks on upper and lower surface) are accurately matched. However, the resulting pressure distribution is not smooth everywhere and the uncertainty bounds are increased. 4(f) shows similar patterns as 4(c) but with further increased 95 % bounds.

For a quantitative comparison of the results, the root mean squared error Eq. (31) for all investigated test cases and methods is given in Table 1. The discrepancy between the purely CFD-based POD + TPS interpolation result and the Gappy POD methods are underlined by a large difference in the root mean squared error for all investigated cases. Compared to the established Gappy POD with ridge regression, the results obtained by Gappy POD with GPR show a smaller root mean squared error in all cases. The difference is highest for the second test case of M = 0.82 and $\alpha = -6.50°$. In this case, as discussed above, the noise variance was estimated to be comparatively small and thus the

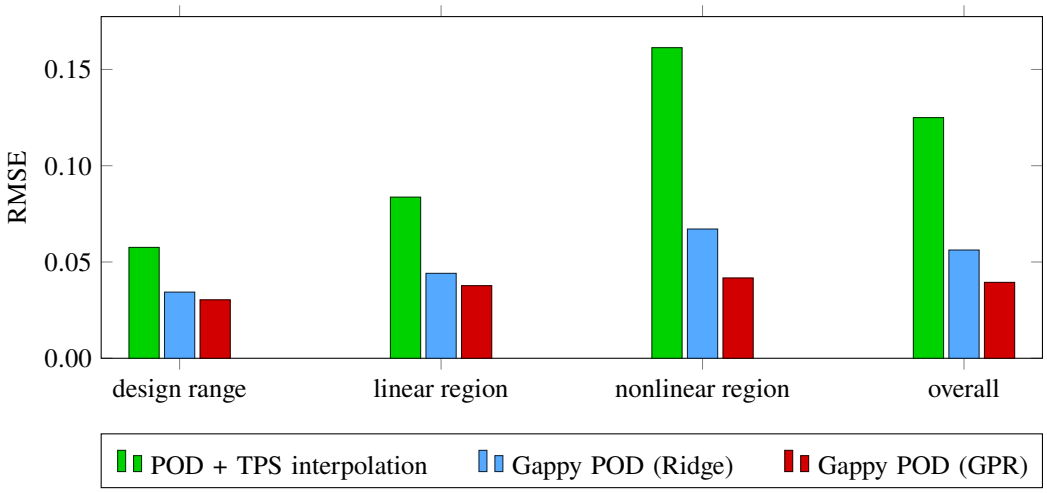


Fig. 6 Average root mean squared error with respect to the 314 wind tunnel pressure probes for Gappy POD with ridge regression, Gappy POD with GPR and POD with TPS interpolation for 196 wind tunnel data sets in the different domains of the envelope.

model tries to give a good fit to the data. This is not always a desired property as it may lead to overfitting.

Table 1 Root mean squared error of the pressure coefficient approximation for POD + TPS interpolation, Gappy POD with GPR and Gappy POD with ridge regression.

Mach	α [°]	RMSE($\hat{y}$) $\cdot 10^2$		
		POD + TPS	Gappy POD (Ridge)	Gappy POD (GPR)
0.50	−1.97	10.2	2.7	2.3
0.82	−6.50	25.0	6.4	0.1
0.86	−2.98	10.1	5.3	2.3
0.86	3.98	8.7	5.2	4.3
0.86	6.49	26.4	8.2	5.8
0.92	0.02	9.1	4.9	3.2
0.92	7.01	20.0	7.2	6.7

In addition to the cases presented above, the root mean squared error was evaluated for all available 196 data sets from the wind tunnel experiment. In Figure 6 the average root mean squared error for the predictions of the investigated methods is displayed divided by the different regions of the envelope.

As can be seen from the bar plot, the Gappy POD approaches outperform the simple POD + TPS interpolation in all regions of the envelope. While the differences between the two Gappy POD approaches in the design range and linear region appear to be very small, Gappy POD with GPR yields a significant improvement in terms of the RMSE in the nonlinear region. Averaged over all 196 wind tunnel test cases, the root mean squared error of Gappy POD with GPR gives a root mean squared error which is about 30 % smaller than the root mean squared error of the established Gappy POD with ridge regression.

IV. Conclusions and Outlook

In this work, an extension of the Gappy POD approach for fusing experimental and numerical data is introduced. The approach alters the Gappy POD regression problem by employing Bayesian regression techniques. The Bayesian perspective allows to define a common statistical framework in which previous Gappy POD methods like ordinary and regularized Gappy POD are special cases. Furthermore, the statistical point of view enables to provide predictive distributions which give information about the uncertainty associated with the model prediction and therefore allow for

a more qualitative evaluation of data fusion results. The limitation of linearity in the Gappy POD approach is leveraged by solving the Gappy POD regression problem using Gaussian process regression with nonlinear covariance functions. The new extension was demonstrated and compared to other methods by means of an industrial-relevant test case for which high-quality numerical and experimental wind tunnel data is available. It was shown that, for the test case at hand, the new extension yields data fusion results which are in better agreement with the wind tunnel measurements than the results obtained with regularized Gappy POD, reducing the root mean squared error by 30 %. The promising outcomes of this initial study motivate further investigations of the Bayesian Gappy POD extension in future research.

The choice of a covariance function for the Gaussian process regression and the tuning of its hyperparameter have a large impact on the data fusion result. In this initial study, we chose a wide-spread covariance function for GPRs with no specific connection to aerodynamic applications as our purpose was the general demonstration of the proposed extension. A more careful choice of a covariance function could however help to further improve the results. Moreover, further work needs to be done regarding the integration of known uncertainties: Some uncertainties like measurement uncertainties due to sensor inaccuracy or systematic errors due to the instrumentation in the wind tunnel on the one hand and errors in the computational data due to modeling or convergence on the other hand may be quantified up to some extend by e.g. manufacturer's specifications, data analysis or expert knowledge. Although the proposed extension would in general allow to incorporate knowledge on uncertainties, an investigation of the effect of taking such information into account needs to be done in future research.

Acknowledgment

The authors would like to thank Airbus for providing the XRF1 test case as a mechanism for demonstration of the approaches presented in this paper.

References

[1] Blazek, J., *Computational fluid dynamics: Principles and applications*, third edition ed., Butterworth Heinemann, Amsterdam, 2015.

[2] Han, Z.-H., Görtz, S., and Zimmermann, R., "Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function," *Aerospace Science and Technology*, Vol. 25, No. 1, 2013, pp. 177–189. https://doi.org/10.1016/j.ast.2012.01.006.

[3] Keane, A. J., "Wing Optimization Using Design of Experiment, Response Surface, and Data Fusion Methods," *Journal of Aircraft*, Vol. 40, No. 4, 2003, pp. 741–750. https://doi.org/10.2514/2.3153.

[4] Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., "Design and Analysis of Computer Experiments," *Statistical Science*, Vol. 4, No. 4, 1989. https://doi.org/10.1214/ss/1177012413.

[5] Santner, T. J., Williams, B., and Notz, W., *Design and analysis of computer experiments*, Springer series in statistics, Springer, New York, 2003. URL http://www.loc.gov/catdir/enhancements/fy0817/2003045444-d.html.

[6] Rasmussen, C. E., and Williams, C. K. I., *Gaussian processes for machine learning*, 3rd ed., Adaptive computation and machine learning, MIT Press, Cambridge, Mass., 2008.

[7] Forrester, A. I. J., Sóbester, A., and Keane, A. J., *Engineering design via surrogate modelling: A practical guide*, Wiley, Hoboken, NJ, USA, 2008.

[8] Kennedy, M. C., and O'Hagan, A., "Predicting the output from a complex computer code when fast approximations are available," *Biometrika*, Vol. 87, No. 1, 2000, pp. 1–13. https://doi.org/10.1093/biomet/87.1.1.

[9] Han, Z.-H., Zimmermann, R., and Görtz, S., "A New Cokriging Method for Variable-Fidelity Surrogate Modeling of Aerodynamic Data," *48th AIAA Aerospace Sciences Meetings*, American Institute of Aeronautics and Astronautics, Reston, Virginia, 2010. https://doi.org/10.2514/6.2010-1225.

[10] Han, Z.-H., Zimmermann, and Görtz, S., "Alternative Cokriging Method for Variable-Fidelity Surrogate Modeling," *AIAA Journal*, Vol. 50, No. 5, 2012, pp. 1205–1210. https://doi.org/10.2514/1.J051243.

[11] Bertram, A., and Zimmermann, R., "Theoretical investigations of the new Cokriging method for variable-fidelity surrogate modeling," *Advances in Computational Mathematics*, Vol. 44, No. 6, 2018, pp. 1693–1716. https://doi.org/10.1007/s10444-017-9585-1.

[12] Han, Z.-H., and Görtz, S., "Hierarchical Kriging Model for Variable-Fidelity Surrogate Modeling," *AIAA Journal*, Vol. 50, No. 9, 2012, pp. 1885–1896. https://doi.org/10.2514/1.J051354.

[13] Forrester, A. I. J., Bressloff, N. W., and Keane, A. J., "Optimization using surrogate models and partially converged computational fluid dynamics simulations," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 462, No. 2071, 2006, pp. 2177–2204. https://doi.org/10.1098/rspa.2006.1679.

[14] Forrester, A. I. J., Sóbester, A., and Keane, A. J., "Multi-fidelity optimization via surrogate modelling," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 463, No. 2088, 2007, pp. 3251–3269. https://doi.org/10.1098/rspa.2007.1900.

[15] Forrester, A. I. J., and Keane, A. J., "Recent advances in surrogate-based optimization," *Progress in Aerospace Sciences*, Vol. 45, No. 1-3, 2009, pp. 50–79. https://doi.org/10.1016/j.paerosci.2008.11.001.

[16] Kuya, Y., Takeda, K., Zhang, X., and Forrester, A. I. J., "Multifidelity Surrogate Modeling of Experimental and Computational Aerodynamic Data Sets," *AIAA Journal*, Vol. 49, No. 2, 2011, pp. 289–298. https://doi.org/10.2514/1.J050384.

[17] Renganathan, S. A., Harada, K., and Mavris, D. N., "Aerodynamic Data Fusion Toward the Digital Twin Paradigm," *AIAA Journal*, Vol. 58, No. 9, 2020, pp. 3902–3918. https://doi.org/10.2514/1.J059203.

[18] Bertram, A., "Data-driven variable-fidelity reduced order modeling for efficient vehicle shape optimization," PhD Thesis, TU Braunschweig, Braunschweig, 2018. https://doi.org/10.24355/dbbs.084-201811231243-0.

[19] Bertram, A., Othmer, C., and Zimmermann, R., "Towards Real-time Vehicle Aerodynamic Design via Multi-fidelity Data-driven Reduced Order Modeling," *2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, American Institute of Aeronautics and Astronautics, Reston, Virginia, 2018. https://doi.org/10.2514/6.2018-0916.

[20] Pinnau, R., "Model Reduction via Proper Orthogonal Decomposition," *Model Order Reduction: Theory, Research Aspects and Applications*, Mathematics in Industry, Vol. 13, edited by H.-G. Bock, F. de Hoog, A. Friedman, A. Gupta, H. Neunzert, W. R. Pulleyblank, T. Rusten, F. Santosa, A.-K. Tornberg, L. L. Bonilla, R. Mattheij, O. Scherzer, W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 95–109. https://doi.org/10.1007/978-3-540-78841-6_5.

[21] Berkooz, G., Holmes, P., and Lumley, J. L., "The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows," *Annual Review of Fluid Mechanics*, Vol. 25, No. 1, 1993, pp. 539–575. https://doi.org/10.1146/annurev.fl.25.010193.002543.

[22] Everson, R., and Sirovich, L., "Karhunen-Loève procedure for gappy data," *Journal of the Optical Society of America A*, Vol. 12, No. 8, 1995, p. 1657. https://doi.org/10.1364/JOSAA.12.001657.

[23] Bui-Thanh, T., Damodaran, M., and Willcox, K. E., "Aerodynamic Data Reconstruction and Inverse Design Using Proper Orthogonal Decomposition," *AIAA Journal*, Vol. 42, No. 8, 2004, pp. 1505–1516. https://doi.org/10.2514/1.2159.

[24] Venturi, D., and Karniadakis, G. E., "Gappy data and reconstruction procedures for flow past a cylinder," *Journal of Fluid Mechanics*, Vol. 519, 2004, pp. 315–336. https://doi.org/10.1017/S0022112004001338.

[25] Gunes, H., Sirisup, S., and Karniadakis, G. E., "Gappy data: To Krig or not to Krig?" *Journal of Computational Physics*, Vol. 212, No. 1, 2006, pp. 358–382. https://doi.org/10.1016/j.jcp.2005.06.023.

[26] Willcox, K., "Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition," *Computers & Fluids*, Vol. 35, No. 2, 2006, pp. 208–226. https://doi.org/10.1016/j.compfluid.2004.11.006.

[27] Franz, T., and Held, M., "Data Fusion of CFD Solutions and Experimental Aerodynamic Data," *ODAS 2017*, edited by DLR and ONERA, 2017. URL https://elib.dlr.de/114707/.

[28] Mifsud, M., Vendl, A., Hansen, L.-U., and Görtz, S., "Fusing wind-tunnel measurements and CFD data using constrained gappy proper orthogonal decomposition," *Aerospace Science and Technology*, Vol. 86, 2019, pp. 312–326. https://doi.org/10.1016/j.ast.2018.12.036.

[29] Bui-Thanh, T., Damodaran, M., and Willcox, K., "Proper Orthogonal Decomposition Extensions for Parametric Applications in Compressible Aerodynamics," *21st AIAA Applied Aerodynamics Conference*, American Institute of Aeronautics and Astronautics, Reston, Virigina, 2003, p. 519. https://doi.org/10.2514/6.2003-4213.

[30] Golub, G. H., and van Loan, C. F., *Matrix computations*, 4$^{\text{th}}$ ed., Johns Hopkins studies in mathematical sciences, Johns Hopkins Univ. Press, Baltimore, MD, USA, 2013.

[31] Hastie, T., Tibshirani, R., and Friedman, J. H., *The elements of statistical learning: Data mining, inference, and prediction*, second edition, corrected at 12th printing 2017 ed., Springer series in statistics, Springer, New York, NY, 2017.

[32] Tikhonov, A. N., and Arsenin, V. J., *Solutions of ill-posed problems*, Scripta series in mathematics, Winston, Washington, D.C., 1977.

[33] Murphy, K. P., *Machine learning: A probabilistic perspective*, Adaptive computation and machine learning series, The MIT Press, Cambridge, Massachusetts and London, England, 2012.

[34] Bishop, C. M., *Pattern recognition and machine learning*, corrected at 8th printing 2009 ed., Information science and statistics, Springer, New York, NY, 2009.

[35] Schwamborn, D., Gerhold, T., and Heinrich, R., "The DLR TAU-Code: Recent Applications in Research and Industry," *ECCOMAS CFD 2006: Proceedings of the European Conference on Computational Fluid Dynamics*, edited by P. Wesseling, Delft University of Technology, [s. l.], 2006.

[36] Abbas-Bayoumi, A., and Becker, K., "An industrial view on numerical simulation for aircraft aerodynamic design," *Journal of Mathematics in Industry*, Vol. 1, No. 10, 2011. https://doi.org/10.1186/2190-5983-1-10.

[37] Ly, H. V., and Tran, H. T., "Modeling and control of physical processes using proper orthogonal decomposition," *Mathematical and Computer Modelling*, Vol. 33, No. 1-3, 2001, pp. 223–236. https://doi.org/10.1016/S0895-7177(00)00240-5.

[38] Buhmann, M. D., *Radial Basis Functions*, Cambridge University Press, Cambridge, 2003. https://doi.org/10.1017/CBO9780511543241.