Low Order Model for Transonic Afterbody Aerodynamic Characteristics

G. Zuccolo[1], R. Christie[2], D. MacManus[3], I. Goulos[4]
Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom

P. Martin[5]
Defence Science and Technology Laboratory, Portsdown West, Fareham, PO17 6AD, United Kingdom

A key aspect in the preliminary design of new combat aircraft is the prediction of the afterbody and exhaust system aerodynamic drag. To meet the various operating conditions requirements for a multi-role vehicle the afterbody typically includes a variable geometry. Within the preliminary design context, this makes the aerodynamic performance prediction a difficult challenge. This research investigates reduced order models for prediction of the aerodynamic performance of axisymmetric transonic afterbody and nozzle systems for a range of aerodynamic conditions and geometric degrees of freedom. The aerodynamic performance metric of interest is afterbody drag coefficient (C_D). Two reduced order models are investigated: artificial neural network and Gaussian process. The geometric variables include boattail closing angle, nozzle throat to exit area ratio and afterbody mean angle and the aerodynamic parameters are free-stream Mach number and nozzle pressure ratio. The results show that these types of reduced order models can be used for preliminary design aerodynamic performance prediction. The Gaussian process C_D prediction is less accurate compared to the artificial neural network with the latter giving a prediction uncertainty of approximately ±0.01 in C_D with a 2σ confidence level. The Gaussian process prediction uncertainty is approximately ±0.013 C_D.

I. Nomenclature

Roman symbols:

A_9/A_8 = exit to throat area ratio

A_1 = afterbody shoulder cross-sectional area $A_1 = \pi r_i^2$

C_D = afterbody drag coefficient based on A_1

M_∞ = free-stream Mach number

NPR = Nozzle total to static Pressure Ratio $NPR = \frac{P_{07}}{p_\infty}$

$DNPR$ = Design Nozzle total to static Pressure Ratio

R = gas constant for dry air: $R = 287 \; J/_{Kg\,K}$

T_7 = nozzle inlet total temperature

---

[1] PhD Candidate, Centre for Propulsion Engineering, g.zuccolo@cranfield.ac.uk.
[2] Research Fellow in Propulsion Aerodynamics, Centre for Propulsion Engineering, r.christie@cranfield.ac.uk.
[3] Head of Gas Turbine Technology Group, Centre for Propulsion Engineering, D.G.Macmanus@cranfield.ac.uk.
[4] Lecturer in Propulsion Integration, Centre for Propulsion Engineering, i.goulos@cranfield.ac.uk.
[5] Principal Scientist, Defence Science and Technology Laboratory.

Greek symbols:

γ = specific heat ratio

β = boattail closing angle

β_1 = Prandtl-Glauert factor $\rightarrow \beta_1 = \sqrt{1 - M_\infty^2}$ if $M_\infty < 1$, $\beta_1 = \sqrt{M_\infty^2 - 1}$ if $M_\infty > 1$

ϑ = afterboby mean angle

II.   Introduction

Some new generation fighter aircraft are expected to be multi-role vehicles designed to perform a large variety of missions. The required flexibility makes the design process more challenging and different requirements must be taken into account over a wide range of flight Mach numbers and engine power settings. These requirements may include high manoeuverability, low infrared signature, thrust vectoring, variable area nozzle and low radar signature as well as extended range [1]. To comply with these performance demands the integration of the propulsion system within the aircraft is a key consideration. Intakes and afterbody-nozzle systems can include moving parts to efficiently accommodate different power settings and flight speeds. In addition, due to weight and balance requirements, fighter-type aircraft tend to incorporate short, steep afterbodies which may be prone to external flow separation [2].

Although an aircraft has many sources of drag, not all the components contribute in the same manner. For instance, in some vehicles the afterbody can account for up to 30% of their zero-lift drag [3]. For this reason it is important to understand the afterbody drag sensitivity to geometric features, engine power settings and aircraft flight conditions from the early design stages. Experimental studies have addressed the effect of M_∞, NPR and afterbody mean angle (ϑ) on drag for simple geometry parametrizations such as circular arc and conical afterbodies [2], [4]. The data was used to develop empirical correlations for the prediction of afterbody aerodynamic performance (C_D). These correlations allowed the prediction of C_D for the subsonic (M_∞=[0.4-0.96]) and supersonic (M_∞=[1.15-1.3]) flight regime separately with an accuracy of $\pm 0.01 C_D$ and $\pm 0.05 C_D$ for circular arc and conical afterbodies respectively [5]. These correlations were made available in the form of carpet plots for specific M_∞ and graphical interpolation is required to assess C_D at intermediate M_∞ [5]. Correlations of this form are difficult to use to address the problem of multivariate aircraft optimization therefore other types of afterbody performance correlations were explored [6]. However these were in the form of simple equations and allow the assessment of a reduced number of degrees of freedom on afterbody and nozzle system aerodynamic performance.

The use of Response Surface Models (RSMs) could represent a significant advantage both in terms of accuracy and suitability for optimization algorithms by replacing the current simple correlations [6]. The aim of this paper is to assess quantitatively the performance of RSMs for the prediction of afterbody drag across the M_∞ regime and for different throttle settings (NPR), boattail closing angle (β), nozzle exit to throat area ratio (A_9/A_8) and ϑ. In the past RSMs have been successfully used in the civil aerospace applications for the prediction of aerodynamic loads on aircraft and for the performance prediction of low NPR subsonic exhaust systems [7], [8]. However, to the authors' knowledge they have not been used for drag prediction of transonic afterbody and exhaust systems across a wide range of geometric and aerodynamic variables.

III.Background

A.   Afterbody and Exhaust Parametrization

The afterbody geometry parametrization is based on the CST method. This analytical definition of the geometry is infinitely differentiable and offers the advantage of generating smooth curves with good aerodynamic properties using an arbitrary number of intuitive Degrees of Freedom (DoF) [9]. In this study the afterbody and exhaust system geometry is defined by six independent geometric DoF, afterbody length (L), shoulder radius (ri), nozzle base thickness (δ), boattail angle (β) exhaust duct exit to throat area ratio (A_9/A_8) and straight shoulder length (L_1) (Fig. 1a). The divergent section of the exhaust duct is designed using the Method of Characteristics (MoC) [10]–[12]. This is to allow the ideal expansion of the exhaust gas and to minimize the external afterbody-exhaust flow interaction terms at the design NPR. The geometry of the divergent section of the nozzle duct is univocally defined by A_9/A_8 while the convergent section is parametrized using a CST curve. This afterbody and exhaust geometry parametrization method has been introduced and discussed in detail in previous work and it has shown that it offers improved flexibility compared to traditional circular arc or conical afterbodies [13] (Fig. 1b).

(a)   (b)
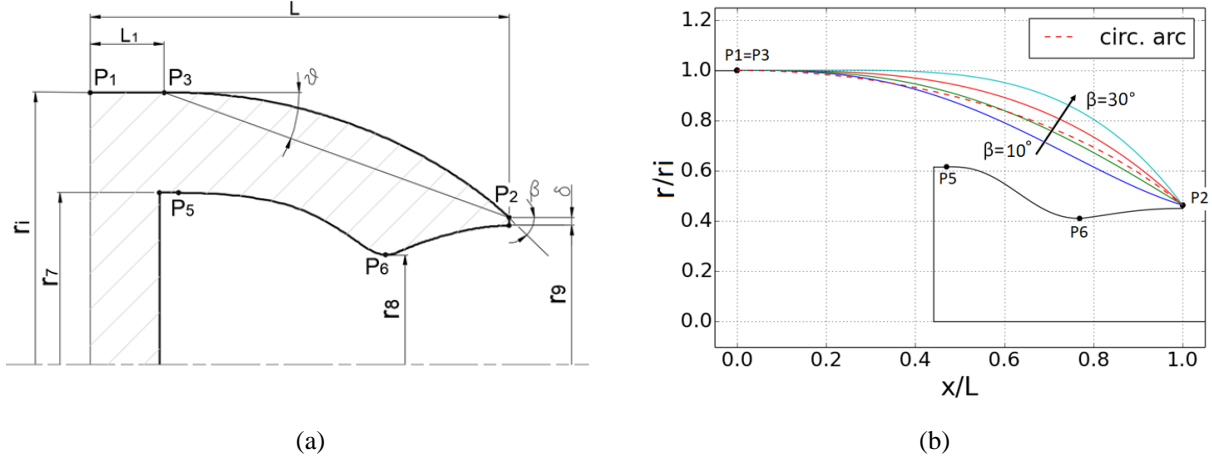
Fig. 1: (a) CST afterbody parametrization and (b) Effect of β

B. Aerodynamic Simulations

The data used for the Response Surface Models has been generated using Computational Fluid Dynamics (CFD) simulations. The axisymmetric computational domain consists of a 2D circular far-field of radius $80r_i$, a sting that extends $22r_i$ upstream of the afterbody from the point $P_1=(0,r_i)$. The circular far-field was modelled with a pressure far-field boundary condition where the static pressure and temperature were prescribed. The free-stream Mach number (M_∞) was altered by changing the total pressure at the far-field with static pressure and static temperature kept constant. The sting was modelled as an inviscid wall for the first portion adjacent to the pressure far field ($L_S=6.6r_i$) and as a viscous wall boundary condition for the remaining part. All the remaining walls were modelled as no-slip adiabatic wall boundary conditions. The inlet of the nozzle duct was modelled with a pressure inlet boundary condition where the NPR was changed by prescribing the inlet total pressure. The computational mesh was generated using a hybrid approach encompassing rectangular cells in the near-wall region and an unstructured mesh made of triangular cells in the remaining part of the domain. The near-wall region was discretized with a set of 60 inflation layers spaced with a growth ratio of 1.15. In this region the y^+ was approximately 1. A mesh independence study based on the GCI [14] led to the choice of a mesh size of approximately 440 thousands elements [13]. An implicit, density based, axisymmetric RANS solver was used. The conservation equations were discretized with a second order scheme and the gradients were computed with a Green-Gauss node-based scheme. This computational approach has been validated against experimental results [13].

The dataset used to build the RSM (Dataset-A) consists of 4608 points. This generated performance correlations aimed to assess the effect of afterbody external geometry parameters (β, ϑ), flight condition (M_∞) and nozzle operating condition (NPR, A_9/A_8) on external afterbody aerodynamic performance. The dataset comprised 36 geometries and 128 aerodynamic conditions and was based on full factorial spacing of the data points (Table 1 [13]).

DoF	Values		DoF	Range
A_9/A_8	1.2, 1.4, 1.6		NPR	3.5-8.25
β	10°, 20°, 30°		M_∞	0.6-1.4
L_1	0.0L, 0.2L, 0.4L, 0.6L			
(ϑ)	(~5°-20°) *Dependent variable*			

(a)   (b)

Table 1: Geometric DoF (a) and Aerodynamic DoF (b) ranges of the aerodynamic simulations

Additional independent datasets were used to assess the performance of the RSMs. The aim of these additional datasets was to enable the assessment of RSMs accuracy and performance robustness for C_D prediction. The RSM prediction accuracy test was based on an independent dataset (Dataset-B) made of 500 configurations selected from a Latin Hypercube Sampling (LHS) of the bounds of Dataset-A (Table 1). The LHS gives combinations of the 5 DoF in a way that provides the optimum distribution of the bounds of the DSE. For this case it consists of 500 different

geometries each one run at a unique combination of NPR and M_∞. The RSMs performance and robustness was assessed with 20 independent datasets (Dataset-C) each one based on an LHS containing 25 afterbody and exhaust systems configurations with each one assessed at a different combination of aerodynamic conditions. This enabled the assessment of the impact of random aspect of the test data selection.

C. Performance Metrics

The main metric used to assess the aerodynamic performance of the afterbody is C_D. This is based on the stream force and control volume approach [15] and is defined following the schematics and engine station numbers of (Fig. 2). The drag coefficient accounts for the effect of the viscous and pressure forces acting on the afterbody external surface including the nozzle base (Eq. 1).

$$C_D = \frac{\phi_{aft}}{\frac{1}{2} A_1 V_\infty^2 \rho_\infty} \tag{1}$$

Where ϕ_{aft} is the afterbody external drag force, V_∞ is the free-stream velocity, A_1 is the afterbody maximum cross sectional area and ρ_∞ is the free-stream density.
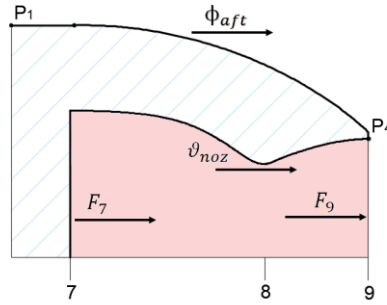

Fig. 2 Control volumes and forces definition for a generic afterbody and exhaust system

D. Performance Correlation State of the Art

To enable the afterbody and exhaust aerodynamic performance prediction some low order model were developed. These were built on wind-tunnel based experimental data on transonic axisymmetric afterbody and nozzle systems. The vast majority of the data was generated using circular arc afterbodies with simple convergent nozzles. These performance correlations enabled the assessment the dependency of C_D on a number of geometric and aerodynamic parameters. The afterbody geometry parametrization used for most of the experimental work was a simple circular arc, therefore the geometric DoF were afterbody length (L) and mean angle (ϑ). The aerodynamic DoF explored were NPR and M_∞. Although the M_∞ range was relatively large, spanning from subsonic ($M_\infty = 0.4$) to supersonic ($M_\infty = 1.3$) these correlations enabled the prediction of C_D for the subsonic ($M_\infty=[0.4\text{-}0.96]$) and supersonic ($M_\infty=[1.15\text{-}1.3]$) flight regime separately. with an accuracy of $\pm0.01C_D$ and $\pm0.05C_D$ for circular arc and conical afterbodies respectively [5] No information was provided for the transonic M_∞ range ($M_\infty=0.96$ to $M_\infty=1.15$). Other correlations provided information for the transonic regime but were based on jet-off experimental work and simple circular arc or conical geometry parametrizations. These correlations were presented in the form of carpet plots for specific M_∞ requiring graphical interpolation to determine C_D at intermediate M_∞ [5]. Correlations of this form are difficult to use to address the problem of multivariate aircraft optimization therefore other types of afterbody performance correlations were explored ([6]). These only enabled the prediction of C_D as a function of the geometric features of the afterbody (L, ϑ) and no accuracy assessment was provided.

IV. Methodology

A. Gaussian Process

The Gaussian process method used within this research is based on the work by Lophaven [16]. The typical use of this method is to construct an approximated model of a physical phenomenon based on a discrete collection of pairs of inputs and the related output [17]. The approximated model, known as Response Surface Model (RSM), can be used as a surrogate for the computer model and it offers a continuous interpolation of the input data. A Gaussian

process RSM approximates a generic function y(x) using a model made of the combination of a regression function and a correlation function, μ and z respectively (Eq. 2**(1)**).

$$\tilde{y}(x) = \mu + z(x) \qquad\qquad 2$$

Where $\tilde{y}(x)$ is the RSM approximation of y(x). The correlation function adopted within this formulation is the linear combination of p known polynomials of order $0<=d<=2$ which give a constant, linear and quadratic regression function respectively (Eq.3).

$$\mu = \sum_{k=1}^{p} \alpha_k f_k(x) \qquad\qquad 3$$

Where $\alpha=(\alpha_1, \dots \alpha_k, \dots \alpha_p)$ is the p-dimensional vector containing the regression coefficients. For a dataset made of N samples a constant regression model will have p=1 and $f_k=1$ [17]. The correlation function z is expressed in the form of a stochastic process assumed to have zero mean and standard deviation σ [17] which is the combination of stationary, one-dimensional correlations that are a function of the relative position of an untried point x relative to all the points in the computer experiment (Eq. 4):

$$z = \sum_{i=1}^{N} \gamma_s \prod_{h=1}^{dim} f\big(R_h(\vartheta_h, d_{ij}), r(\vartheta_h, d_{ix})\big) \qquad\qquad 4$$

Where N is the size of the dataset, dim is the dimensionality of the problem, γ_s are weighing coefficients, R_h is one of the *dim* NxN matrices of stochastic process correlation, $d_{ij} = (x_i - x_j)$ is the distance between the i-th and j-th input point of the input dataset with i,j=1…N and the parameter ϑ_h accounts for the correlation between input points. $r(\vartheta_h, d_{ix})$ is the column vector that takes into account the distances of the untried, generic point x from all the other experimental points with $d_{ix} = (x_i - x)$. Large values of ϑ_h lead to faster correlation decrease. For a given untried point x and a given distance d_{ix}, as ϑ_h increases the influence of the experimental point x_i on the RSM output in x decreases. The method used in this work offers the choice of several correlation functions (Table 2). In all the cases the correlation decreases as d_{ij} and ϑ_h increase and it reaches its maximum when $d_{ij} = 0$, for i=j.

NAME	$R_h(\vartheta_h, d_{ij})$
Absolute exponential	$\dfrac{1}{e^{\vartheta_h\|d_{ij}\|}}$
Squared exponential (Gaussian)	$\dfrac{1}{e^{\vartheta_h(d_{ij})^2}}$
Cubic	$1 - 3\varepsilon^2 + 2\varepsilon^3 \qquad \varepsilon = min(1, \vartheta_h\|d_{ij}\|)$
Linear	$max(0, 1 - \vartheta_h\|d_{ij}\|)$

Table 2: List of correlation functions used for the Gaussian process RSM

The parameters ϑ_h, γ_s and α_k are determined by the algorithm through maximum likelihood estimation and ϑ_h is assumed to be constant in all the dimensions, giving an isotropic problem. In the case of a Gaussian correlation function, for a given index h, the correlation matrix would be the following (Eq. 5):

$$R_h(\vartheta_h, d_{ij}) = \begin{bmatrix} R_h(\vartheta_h, d_{1,1}) & \cdots & R_h(\vartheta_h, d_{1,N}) \\ \vdots & \ddots & \vdots \\ R_h(\vartheta_h, d_{N,1}) & \cdots & R_h(\vartheta_h, d_{N,N}) \end{bmatrix} = \begin{bmatrix} 1 & \cdots & \dfrac{1}{e^{\vartheta_h(d_{1,N})^2}} \\ \vdots & \ddots & \vdots \\ \dfrac{1}{e^{\vartheta_h(d_{N,1})^2}} & \cdots & 1 \end{bmatrix} \qquad 5$$

The fact that the diagonal only contains ones means that the RSM resulting from the combination of all these R_h will be forced to pass through the input data points. In the case where there is a known level of uncertainty or some known amount of noise on the output data due to the precision of the computational model used for the input data generation, it is possible to allow the RSM to float around the experimental points. This is done by adding a non-zero value, called nugget (η), to the diagonal of R_h (Fig. 3).
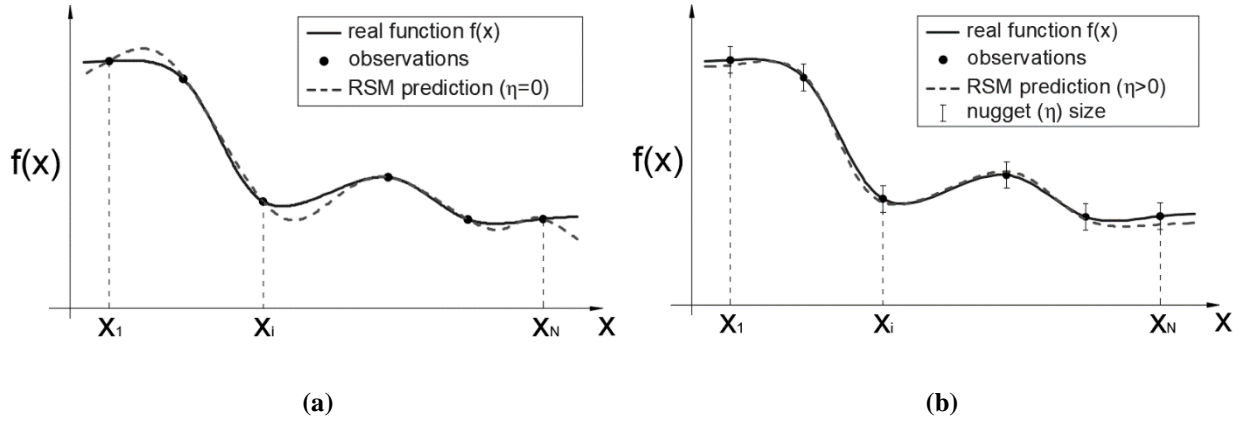

Fig. 3: (a) model with η=0 and (b) effect of non-zero η (based on [18])

The generic $r(\vartheta_h, d_{ix})$ for an untried point x reads as follows (Eq.**6**):

$$r(\vartheta_h, d_{ix}) = \left[\frac{1}{e^{\vartheta_h(d_{1,x})^2}}, \dots, \frac{1}{e^{\vartheta_h(d_{i,x})^2}}, \dots, \frac{1}{e^{\vartheta_h(d_{N,x})^2}}\right]^T \quad i = 1, \dots, N \qquad 6$$

The final RSM is then the combination of regression and correlation function. The regression function is a retro-fitting of the experimental results while the correlation function models the deviation of the complex physics from the simpler regression function [18] (Fig. 4).
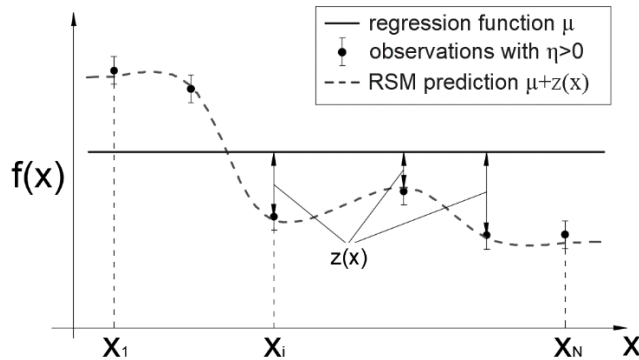
Fig. 4: Correlation and regression function example (based on [18])

In this research the Gaussian process is used to model the aerodynamic performance of axisymmetric transonic afterbody and nozzle system as a function of the 5 DoF explored within Dataset-A ($C_D = f(NPR, M_\infty, L_1, \beta, A_9/A_8)$).

B. Gaussian Process RSM Performance Assessment

The performance of an RSM is usually assessed with the k-fold cross-validation method. The input dataset is randomly split into k different, mutually exclusive, subsets of equal size known as the folds. k different RSMs are then built using an input dataset made of k-1 folds. The excluded subset is used to test the RSM and assess the difference between RSM prediction and actual computer input data giving some measure of the quality of the RSM. The Leave One Out (LOO) is the extreme case of k-fold cross-validation, where k=N and the size of each fold is just one point. One of the metrics that is conventionally used to assess the quality of the Gaussian process RSM is the root mean squared (RMS) of the LOO error [19] (Eq. **7**).

$$RMS_{LOO} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\varepsilon_i}{y_{i\_exp}}\right)^2} \qquad 7$$

Where $\varepsilon_i = \tilde{y}_{i\_LOO} - y_{i\_exp}$ is the LOO error for the i-th fold. $\tilde{y}_{i\_LOO}$ is the prediction of the RSM built with N-1 input points when interrogated in the missing fold, y_{i_exp} is the value obtained from the computer experiment for the i-th point. Along with RMS_{LOO} in this study the average value (Eq.**8**) and the standard deviation (Eq.**9**) of the LOO error are also of interest.

$$\bar{\varepsilon}_{LOO} = \frac{\sum_{i=1}^{N}\varepsilon_i}{N} \qquad 8$$

$$\sigma_{LOO} = \sqrt{\frac{\sum_{i=1}^{N}(\varepsilon_i - \bar{\varepsilon})^2}{N}} \qquad 9$$

The disadvantage of this method is that the algorithm that provides the N RSMs used to assess the LOO needs to run N times, which increases the computational time required by a factor of N. The RMS error of the LOO can be seen as a representative error of the prediction of the RSM in the area covered by the initial input dataset [19]. To assess which combination of regression function, correlation function and nugget size gave the best RSM all the possible combinations of these parameters have been explored by carrying out a full factorial analysis. In this assessment the regression functions explored were of order 0, 1 and 3 while the correlation functions used were linear, absolute exponential and squared exponential. The range of η explored is $[10^{-6} - 10^{-3}]$. This gives a total number of RSM assessed equals to 72. For each one of them the RMS error of the LOO model is computed along with the mean value and the standard deviation of the LOO error. Since RMS_{LOO} can sometimes be misleading [19] an additional Gaussian process RSM performance test was conducted. This was based on the root mean squared error (RMS_{IT}, (Eq.10)) and the standard deviation (σ_{IT} (Eq.**11**)),of the RSM C_D prediction for the configurations explored within Dataset-B.

$$RMS_{IT} = \sqrt{\frac{1}{N_{IT}}\sum_{j=1}^{N_{IT}}\left(\frac{\varepsilon_j}{y_{j\_exp}}\right)^2} \qquad 10$$

$$\sigma_{IT} = \sqrt{\frac{\sum_{j=1}^{N_{IT}}(\varepsilon_j - \bar{\varepsilon})^2}{N_{IT}}} \qquad 11$$

Where in this case N_{IT}=500, $\varepsilon_j = \tilde{y}_{j\_RSM} - y_{j\_exp}$ is the RSM prediction error with $\tilde{y}_{j\_RSM}$ being the RSM predicted C_D for the j-th case and y_{j_exp} being the actual value computed for Dataset-B. Other performance metrics computed to have a better insight on the RSM accuracy are the mean absolute error (MAE_{IT}, Eq. 12) and the RSM maximum error ($E_{max}=\max(\varepsilon_j)$). These are based on Dataset-B.

$$MAE_{IT} = \left[\frac{1}{N_{IT}} \sum_{j=1}^{N_{IT}} |\varepsilon_j| \right]$$

12

C. Artificial Neural Network

An Artificial Neural Network (ANN) is a biologically inspired computational model [20]. In biological neural networks, a neuron is a specialist cell, which processes information [21]. A neuron is formed of a cell body known as a soma and out-reaching branches known as axons and dendrites [21]. In an ANN artificial neurons are nodes with connections between them with coefficient weights bound to the connections[20]. Neuron k in Fig. 5 can be described mathematically as:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j$$

13

and

$$y_k = \vartheta(v_k)$$

14

Where x_1, x_2, …, x_m are the input signals, w_1, w_2, …, w_m, are the synaptic weights, v_k is the sum of the input signals weighted by their respective synaptic strengths and the bias b_k, $\varphi(\cdot)$ is the activation function and y_k is the output signal of the neuron [22]. The output is defined by the activation function. When non-linear activation functions are employed complex problems can be modelled with only a small number of nodes. The most basic activation function is the identity function in which the net output is equal to the output of the neuron:

$$f(v_k) = v_k$$

15

and the output from neuron k (y_k) is simply calculated as a linear combination of the inputs [23]. If linear activation functions are used then the last hidden layer will always be a linear function of the first hidden layer and therefore adds no complexity to the system and can be collapsed to a single hidden layer. To allow the network to create complex mappings as more hidden layers are added non-linear activation functions must be used. Two of the most common non-linear activation functions are the logistic and hyperbolic tangent functions (Eq. **16** and Eq. **17**).

$$f(v_k) = \frac{1}{1 + exp(-v_k)}$$

16

$$f(v_k) = \tanh(v_k)$$

17

Both logistic and hyperbolic functions can have slow convergence due to the vanishing gradient problem where for very high or low inputs (v_k) there is almost no change in the prediction and the error vanishes as it gets propagated back [24]. The Rectified Linear unit (ReLu) activation function (Eq. 18) overcomes the vanishing gradient problem and is commonly used in multilayer perceptron and convolutional neural networks. However one potential problem with the ReLu activation function is thaet the gradient is 0 when the neuron is not active. Therefore the network cannot perform back-propagation and prevents learning [25].

$$f(v_k) = max(0, v_k) \qquad\qquad 18$$

The basic architecture of a neural network (Fig. 5b) consists of three types of neuron layers: input, hidden and output layers [26]. Hidden layers and their constituent hidden nodes allow the network to cope with non-linearly separable problems and are so called because they do not interact with the external environment [27]. The selection of the number of hidden layers and neurons is a crucial decision. A large number of hidden neurons will allow the network to correctly predict the data on which it has been trained but can compromise the generalization ability of the network [26]. However, with too few hidden neurons it may not be possible to train the network to have sufficiently low error [26].

In a feed-forward network, signal flow is strictly from input to output nodes. The network can contain many hidden layers but no feedback connections are present [26]. One class of feed-forward ANNs, the Multi-layer Perceptron (MLP) has been used in this paper [18]. The MLPregressor model has outputs, which are continuous values, and optimizes the squared-loss [18]. A quasi-Newton method, L-BFGS and two stochastic gradient-based optimizers, SGD and Adam [28] are implemented as solvers for the weight optimization.

The space of possible network functions grows exponentially large with the depth of the network and may lead to an overfit of the training data and the network to poorly predict outcome values for non-training data. This type of error is known as generalization error[29]. Overfitting of data can be reduced by the addition of a penalty or regularization term (α) for controlling the magnitude of the model parameters to the error function [30]. The MLPregressor model employs the well-known L2norm or ridge regression method [18].

(a)   (b)
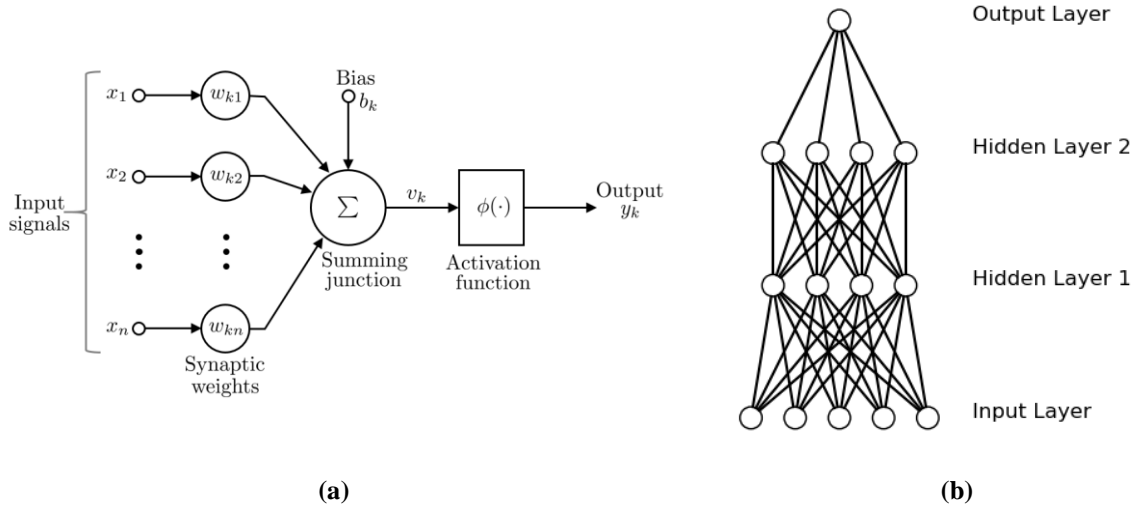
Fig. 5: (a) Model of a neuron adapted from [22] and (b) neural network architecture for a network with five inputs, two hidden layers with four neurons each and a single neuron in the output layer

D. Neural Network Performance Assessment

The independent dataset, Dataset-B, was used to compare the performance of the ANN with the Gaussian –process surrogate model. Therefore all of Dataset-A was used to train the ANNs. The model prediction error was quantified by the maximum error, $(E_{max} = max(\varepsilon_j))$ the Root Mean Squared Error, $(RMS_{IT}$, Eq. 10), the Mean Absolute Error, $(MAE_{IT}$, Eq. 12),the standard deviation of the error, $(\sigma_{IT}$, Eq. 11), and the percentage of predictions with an absolute error greater than 0.01 C_D.

E. RSMs Robustness Testing

Although the test cases contained in the LHS-based independent dataset are optimally spaced to provide the best coverage of the hypervolume defined by the bounds of the independent DoF, the combinations of geometric DoF $(L_1$, β, $A_9/A_8)$ and aerodynamic boundary conditions (NPR, M_∞) is not uniquelly defined. Different LHSs containing the same number of samples and based on the same ranges of DoF will give different geometries operated at different aerodynamic boundary conditions. For this reason, additional testing was conducted to assess the robustness of the performance of the RSMs and the impact of test data spacing. For this purpose smaller LHS-based datasets (Dataset-

C) were used to quantify the average value of the RMS error and standard deviation distribution (E_{RMS}, σ_C) of the RSM predictions. These are defined as:

$$E_{RMS} = \frac{1}{n}\sum_{k=1}^{n} RMS_k \qquad\qquad 19$$

$$\sigma_C = \sqrt{\frac{\sum_{k=1}^{n}(\sigma_k - \overline{\sigma})^2}{n}} \qquad\qquad 20$$

Where n is the number of independent datasets contained in Dataset-C (n=20), RMS_K is the RMS error of the RSM tested with the k-th dataset in Dataset-C, $\overline{\sigma}$ is the average value of the standard deviation of the error distribution ($\overline{\sigma} = \frac{1}{n}\sum_{k=1}^{n}\sigma_k$) and σ_k is the standard deviation of the error distribution for the k-th RSM test performed. E_{RMS} and σ_C quantify the sensitivity of the RSM performance to the testing data used and therefore the robustness of the RSMs.

V. Results

A. Gaussian Process Performance

1. *LOO-based performance evaluation*

The initial assessment of the Gaussian process RSM performance was based on the LOO error. This identified the best combination of regression function, correlation function and nugget size (η). The best performing RSM was the result of the combination of linear correlation function and quadratic regression function with a η=5x10⁻⁴ (Fig. 6). For this model the performance metrics are: $RMS_{LOO} = 0.00126$, $\sigma_{LOO} = 0.00127$, $\overline{\varepsilon}_{LOO} = 1.1e - 06$ and the maximum error is E_{max}=0.0196. This performance assessment method tests the relative importance of each one of the input points in Dataset-A on the overall quality of the RSM. The maximum error E_{max} of the best RSM corresponds to the prediction of the LOO model built with Dataset-A with the exception of the input point given by the configuration with A_9/A_8=1.2 (DNPR=3.75), L_1=0.6L and β=10° when operated at NPR=5.23 and M_∞=1.059.
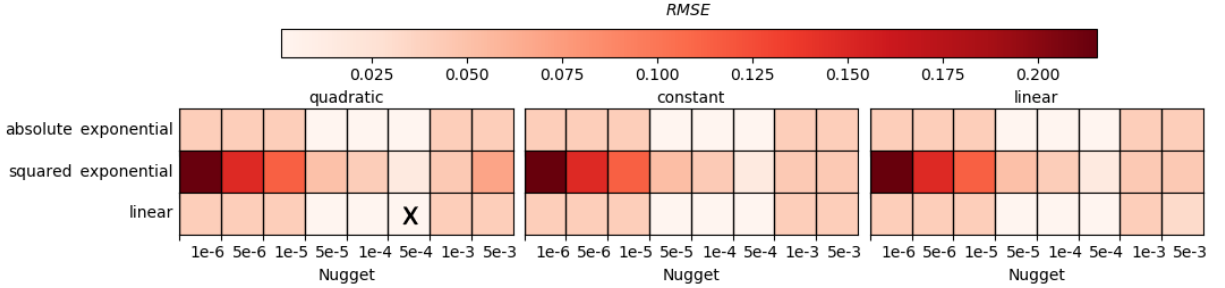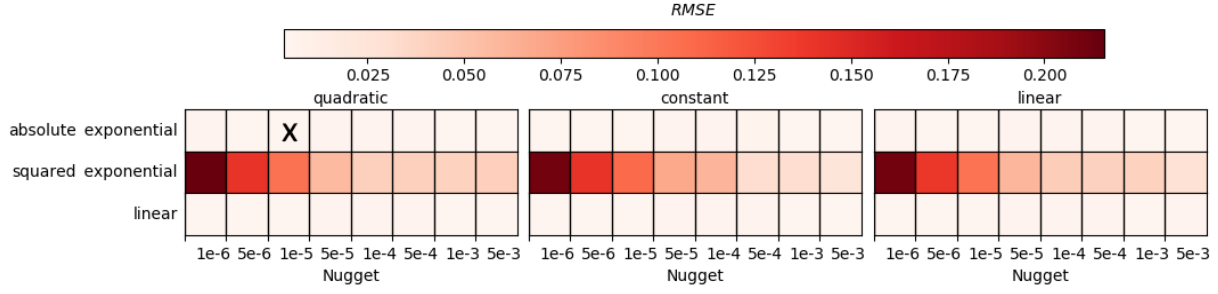

Fig. 6: Gaussian process hyperparameter study for the LOO-based performance assessment

2. *Independent test*

Dataset-B was used to independently quantify the RSMs performance. Rather than quantifying the relative effect of each single data point on the RSM quality this performance assessment method assesses the overall behavior of the model across the hypervoulme defined by the DoF bounds. This assessment identified the best performing RSM as the one resulting from the combination of absolute exponential correlation function, quadratic regression function and a η=5x10⁻⁵ (Fig. 7). For this model the correlation between Dataset-B and model prediction is very good and 1.4% of the predictions were outside of the interval ±0.01 C_D (Fig. 8a). The RSM prediction performance metrics are: $RMS_{IT} = 0.0034$, $\sigma_{IT} = 0.0031$, $\overline{\varepsilon}_{IT} = 0.0014$ and the maximum error is E_{max}=0.01593 (Fig. 8b). The Gaussian process performance proved robust and repeatable when tested for the effect of randomness in the testing data (Fig. 9). For this case the average standard deviation was $\overline{\sigma}$=0.0035 while the standard deviation of the distribution of the standard deviations was σ_C=0.00064. It can be concluded that the Gaussian process accuracy on C_D for axisymmetric afterbody and exhaust systems is $\overline{\varepsilon}_{IT}\pm2\sigma_{IT} = 0.0014\pm0.0062$ with a $2\sigma_{IT}$ confidence level. The value of σ_{IT} can vary within the range $\pm2\sigma_C = 0.00128$ giving a worse case error on the prediction of C_D of 0.0135 with a 2σ confidence level.

Fig. 7: Gaussian process hyperparameter study for independent performance assessment. X marks the best combination of regression function, correlation function and nugget size.

E_{max}	0.0159
RMS_{IT}	0.0034
MAE_{IT}	0.0023
% outside ±0.01 C_D	1.4%
σ_{IT}	0.0031

(a)   (b)
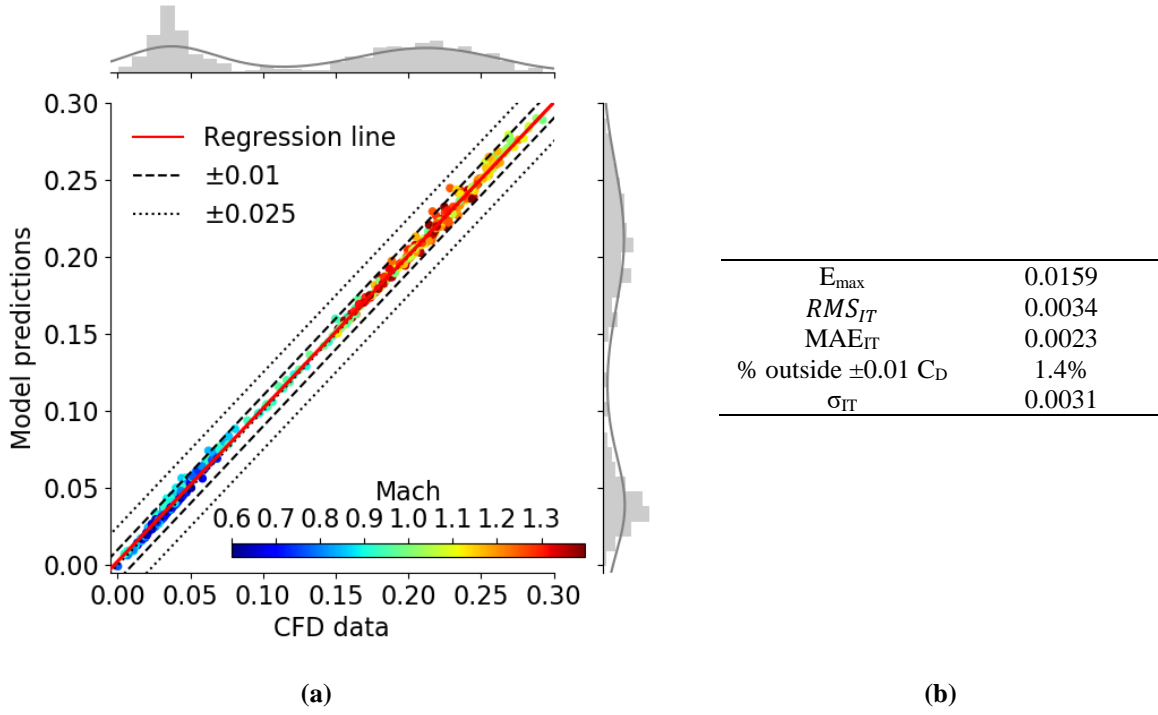
Fig. 8: (a) Correlation of Gaussian process prediction and Dataset-B and (b) performance metrics of the best Gaussian process
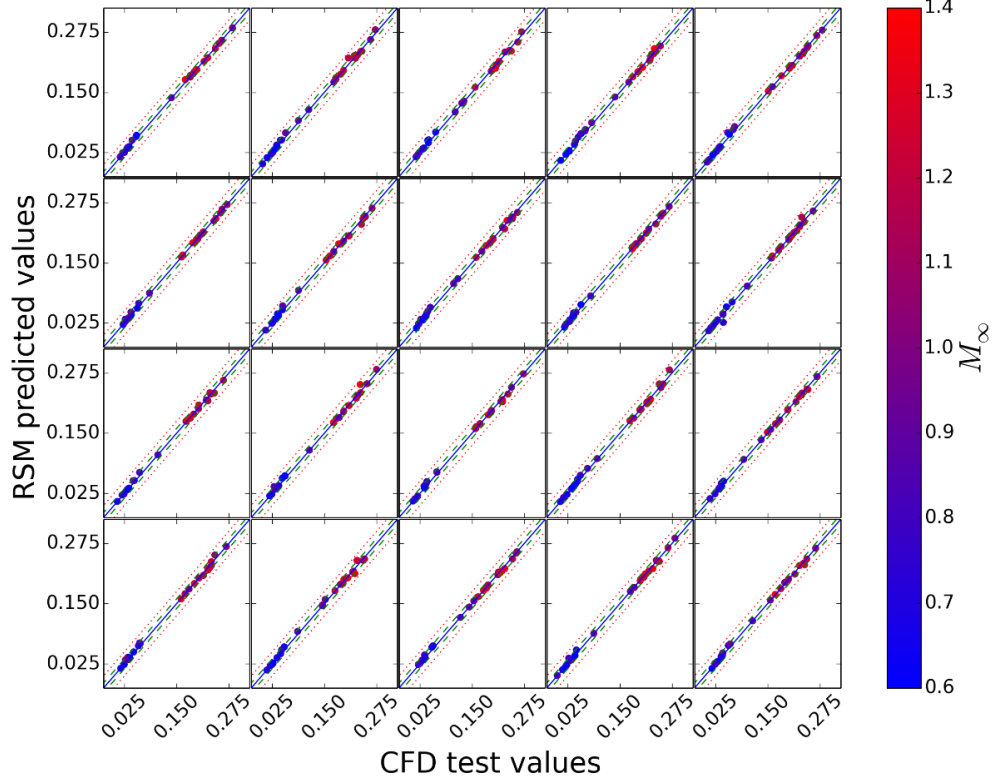
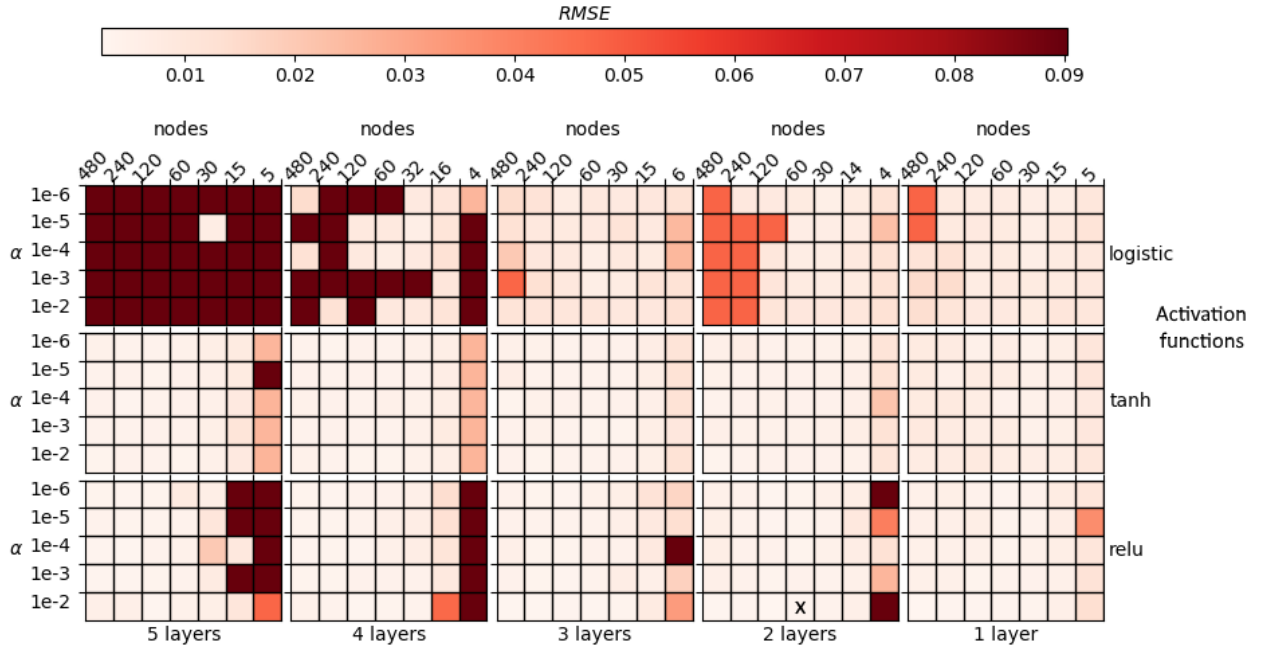**Fig. 9: Correlation of test values and RSM prediction for the 20 LHS dataset in Dataset-C**

B. Artificial Neural Network Performance

To select the hyperparameters for the ANN surrogate model a full factorial investigation was carried out varying the L2 penalty term, α, the activation function, the number of hidden layers and the overall number of nodes. The model fitness was assessed by the RMS_{IT} of predictions made for Dataset-B (Fig. 10). Three activation functions were investigated, including the logistic, hyperbolic tangent and rectified linear unit functions. For each activation function various network structures were modelled. The overall number of hidden neurons and the number of hidden layers were varied from five to 480 and one to five respectively. In addition a range of L_2 norm regularization parameters ($1\times10^{-6} \leq \alpha \leq 1\times10^{-2}$) were studied. The minimum RMS_{IT} error (0.0025 C_D) was found for a ANN with a ReLu activation function, a L_2 norm regularization parameters equal to 0.01 and 2 hidden layers each consisting of 60 hidden neurons. The ANN prediction performance metrics are: $RMS_{IT} = 0.0025$, $\sigma_{IT} = 0.0024$ and the maximum error is $E_{max}=0.0110$ (Fig. 11a and b). The ANN performance also proved robust and repeatable when tested for the effect of randomness in the testing data (Fig. 12). For this case the average standard deviation was $\bar{\sigma}=0.0027$ while the standard deviation of the distribution of the standard deviations was $\sigma_C=0.00051$. It can be concluded that the ANN prediction accuracy on C_D for axisymmetric afterbody and exhaust systems is $\bar{\varepsilon}_{IT}\pm2\sigma_{IT} = 0.0\pm0.0048$ with a $2\sigma_{IT}$ confidence level. The value of σ_{IT} can vary within the range $\pm2\sigma_C = 0.00102$ giving a worse case error on the prediction of C_D equals to 0.098 with a $\pm2\sigma$ confidence level.

C. RSM comparison

Both RSM showed good accuracy on the prediction of C_D for axisymmetric aerodynamic afterbody and exhaust systems. The overall performance of the ANN is better than the Gaussian process both in terms of systematic error ($\bar{\varepsilon}_{IT}$) and standard deviation of the error distribution (σ_{IT}). For these reasons ANN is more suitable for accurate afterbody and exhaust systems aerodynamic performance and it is a better candidate for the use within a Multivariate Vehicle Optimization (MVO) tool. On the other hand, the Gaussian process type RSM can still be a good fit for a MVO tool due to its relatively small σ_{IT}. Although this RSM has a non-zero mean error ($\bar{\varepsilon}_{IT}$) and the prediction of C_D will be affected by a systematic error, this might not be relevant for an optimization process. The aim of MVO tools is not the prediction of C_D as an absolute value but to identify the areas in the design space where the "good" configurations are. These region in the design space are evaluated relatively to a starting case, which will also be affected by a systematic error (Fig. 13). Another important parameter that needs to be considered is the computing

time required for these models. The ANN algorithm takes few minutes to produce an RSM while the Gaussian process takes hours.
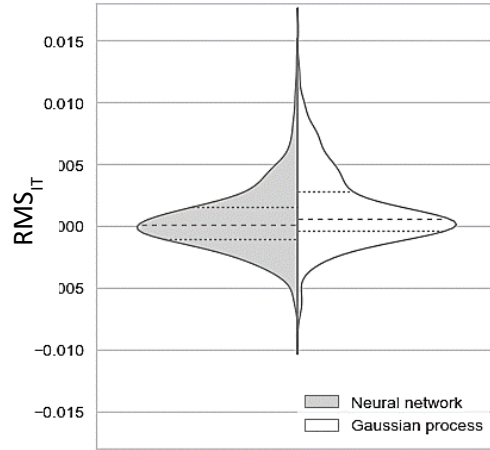

Fig. 10: ANN hyperparameter study for independent performance assessment. X marks the best ANN

E_{max}	0.0114
RMS_{IT}	0.0025
MAE_{IT}	0.0018
% outside ± 0.01 C_D	0.6%
σ_{IT}	0.0024

(a)   (b)
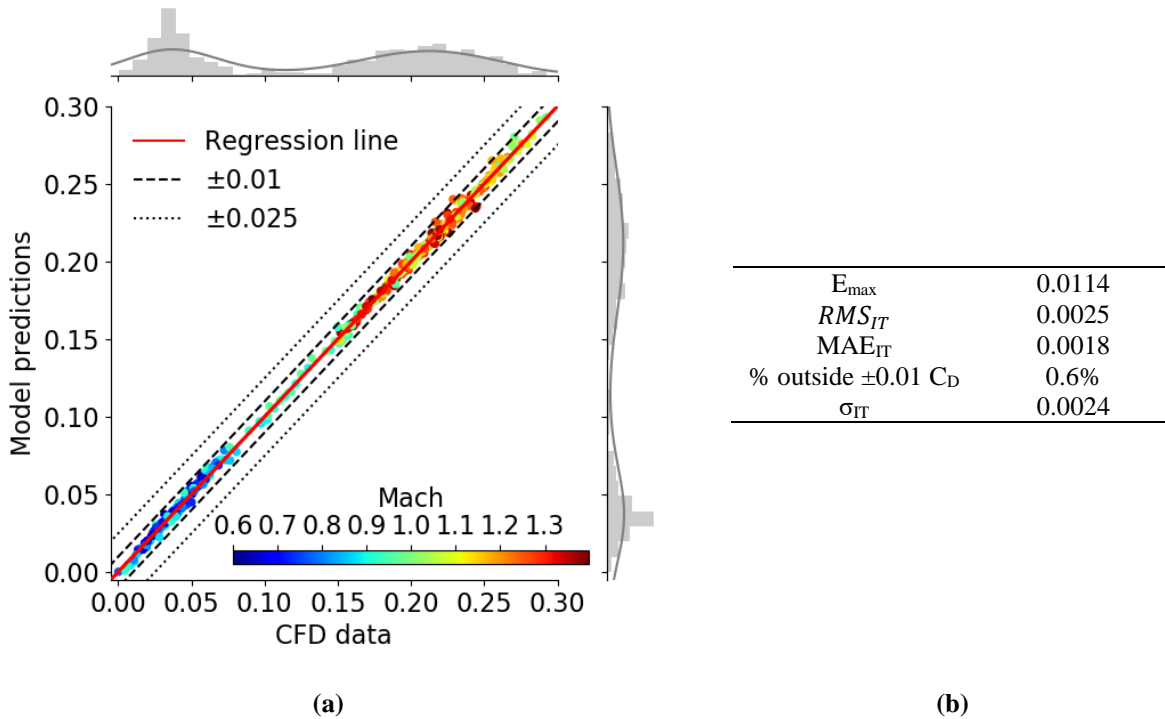
Fig. 11: (a) Correlation of ANN prediction and Dataset-B and (b) performance metrics of the best ANN
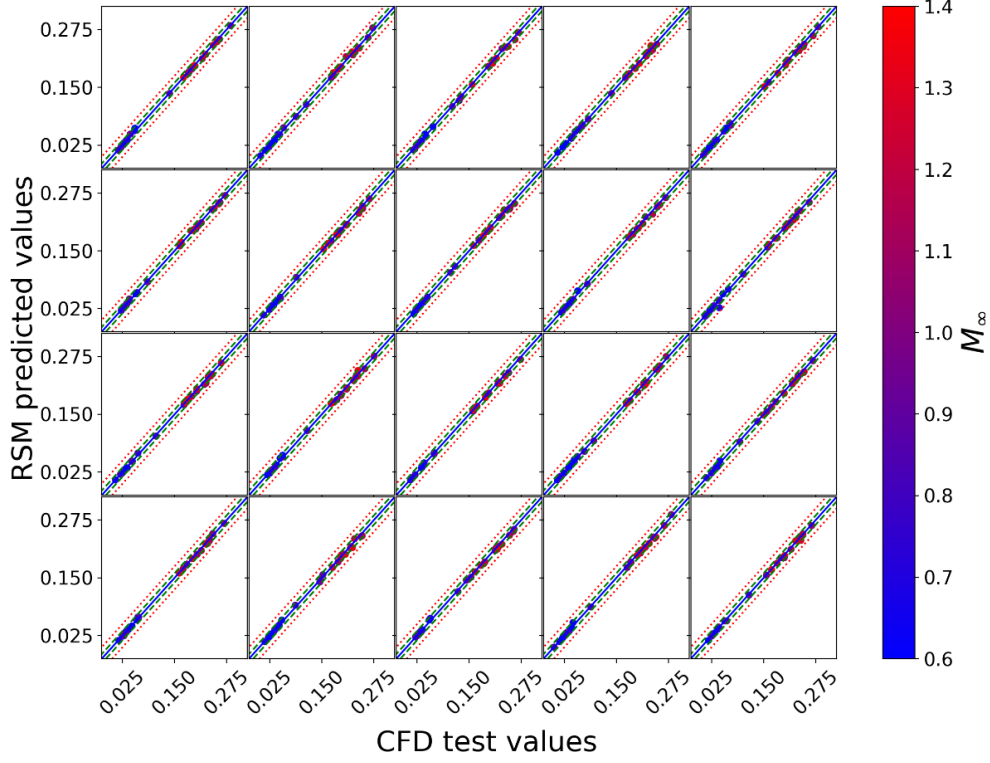
Fig. 12: Correlation of test values and ANN prediction for the 20 LHS dataset in Dataset-C

Fig. 13: Probability density function of prediction error distributions of the Gaussian process and Neural Network models assessed with independent dataset-B

VI. Conclusions

The aim of this research was to quantify the performance of Gaussian process and Artificial Neural Network (ANN) RSMs on afterbody and nozzle system C_D prediction. The objective was to provide a low order model able to predict C_D as a function of several geometric parameters (L_1, β, A_9/A_8) and aerodynamic (NPR, M_∞) conditions. The models were tested using two i[21]ndependent datasets. Dataset-B was used for the RSM prediction performance assessment while Dataset-C used to assess the RSMs robustness and performance repeatability. Overall the Gaussian process and the ANN had prediction uncertainties of 0.0014±0.0062 and ±0.0048 with a 2sigma confidence, respectively. The maximum error was $E_{max}=0.01593$ for the Gaussian process and $E_{max}=0.0110$ for ANN. The ANN

showed no bias in the performance prediction assessed based on Dataset-B and only 0.2% of the predictions were outside the range ±0.01 C_D.

The most important conclusion is that both these low order methods provide an improvement over the conventional models. The prediction accuracy of ANN and Gaussian process is comparable to previous performance correlations. However, the proposed correlations enable the assessment of afterbody drag over a continuous range of flight Mach number including the transonic regime as well as additional geometric control on the afterbody definition to enable a greater set of designs to be considered.

References

[1]   D. Cler, "A parametric investigation of nozzle planform and internal /external geometry at transonic speeds," in *31st Joint Propulsion Conference and Exhibit*, 1995.

[2]   D. E. Reubush and J. F. Runckel, "Effect of Fineness Ratio on Boattail Drag of Circular-Arc Afterbodies Having Closure Ratio of 0.50 With Jet Exhaust at Mach Numbers up to 1.30," NASA TN D-7192, Hampton, Virginia, 1973.

[3]   P. R. Payne, R. M. Hartley, and R. M. Taylor, "Afterbody Drag: Volume 1 - Drag of Conical and Circular Arc Afterbodies without Jet Flow," Bethesda, Maryland, 1980.

[4]   D. E. Reubush, "Effect of Fineness and Closure Ratios on Boattail Drag of Circular-Arc Afterbody Models With Jet Exhaust at Mach Numbers up to 1.30," NASA TN D-7163, Hampton, Virginia, 1973.

[5]   Royal Aeronautical Society, "Subsonic pressure drag of boat-tails with negligible annular base area in the presence of a central propulsive jet," ESDU 01012, 2001.

[6]   D. A. Lovell, "The Application of Multivariate Optimization to Combat Aircraft Design," Farnborough, 1988.

[7]   G. Giangaspero and D. MacManus, "Surrogate models for the prediction of the aerodynamic performance of exhaust systems," vol. under revi.

[8]   A. Heidebrecht and D. G. MacManus, "Surrogate model of complex non-linear data for preliminary nacelle design," *Aerosp. Sci. Technol.*, vol. 84, pp. 399–411, 2019.

[9]   R. Christie, A. Heidebrecht, and D. MacManus, "An Automated Approach to Nacelle Parameterization Using Intuitive Class Shape Transformation Curves," *J. Eng. Gas Turbines Power*, vol. 139, no. 6, 2017.

[10]   T. Zebbiche, "Supersonic axisymmetric minimum length nozzle conception at high temperature with application for air," *KSAS Int. J.*, vol. 9, no. 1, pp. 1–30, 2008.

[11]   B. M. Agrow and G. Emanuel, "Comparison of minimum length nozzles - Agrow, B. M..pdf," *J. Fluids Eng.*, vol. 110, pp. 283–288, 1988.

[12]   J. D. Anderson, *Modern Compressible Flow With Historical Perspective*, Second Ed. McGraw-Hill, 1990.

[13]   G. Zuccolo, D. MacManus, I. Goulos, and P. Martin, "Geometry Parametrisation and Aerodynamic Characteristics of Axisymmetric Afterbodies."

[14]   P. J. Roache, "Perspective: A Method for Uniform Reporting of Grid Refinement Studies," *J. Fluids Eng.*, vol. 116, pp. 405–413, 1994.

[15]   AGARD 237, "Guide to In-Flight Thrust Measurement of Turbojets and Fan Engines," 1979.

[16]   S. N. Lophaven, H. B. Nielsen, and J. Søndergaard, "DACE: a Matlab kriging toolbox Vol.2. IMM. Informatics and Mathematical Modelling," 2002.

[17]   A. I. J. Forrester, A. Sbester, and A. J. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide - Intro*. 2008.

[18]   Pedregosa et al, "Scikit-learn: Machine Learning in Python," 2011. [Online]. Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy_targets.html. [Accessed: 14-Nov-2019].

[19]   R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," vol. 5, p. 7, 1995.

[20]   S. Shanmuganathan and S. Samarasinghe, *Artificial neural network modelling*. 2014.

[21]   A. K. Jain, J. Mao, and K. Mohiuddin, "Arti cial Neural Networks : A Tutorial," *IEE Neural Comput.*, pp. 1–49, 1996.

[22]   S. Haykin, *Neural Networks and Learning Machines*, Third. Hamilton, Ontario: Pearson, 2008.

[23]   G. Villarrubia, J. F. De Paz, P. Chamoso, and F. De Prieta, "Neurocomputing Artificial neural networks used in optimization problems," vol. 272, pp. 10–16, 2018.

[24]   S. Hochreiter, "The Vanishing Gradient Problem During Learning Recurrent Neural Nets And Problem Solutions," *Int. J. Uncertainty, Fuzziness Knowledge-Based Syst.*, pp. 1–10, 1998.

[25]   A. L. Maas, A. Y. Hannum, and A. Y. Ng, "Rectifier Nonlinearities Improve Neural Network Acoustic

Models," in *Proceedings of the 30th International Conference on Ma- chine Learning*, 2013, vol. 28.

[26]   A. Abraham, *Handbook of Measuring Systems*. John Wiley and Sons Ltd, 2005.

[27]   I. A. Basheer and M. Hajmeer, "Artificial neural networks : fundamentals , computing , design , and application," *J. Microbiol. Methods*, vol. 43, pp. 3–31, 2000.

[28]   D. P. Kingma and J. L. Ba, "ADAM: A Method for Stochastic Optimization," *ICLR*, pp. 1–15, 2015.

[29]   Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin, *Learning From Data*. New York: AMLbook.com, 2012.

[30]   E. Phaisangittisagul, "An Analysis of the Regularization between L 2 and Dropout in Single Hidden Layer Neural Network," *2016 7th Int. Conf. Intell. Syst. Model. Simul.*, pp. 174–179, 2016.

2020-01-05

Low order models for transonic afterbody aerodynamic characteristics

Zuccolo, Giovanni

AIAA