MULTI-FIDELITY DESIGN OPTIMISATION OF A TRANSONIC COMPRESSOR ROTOR

C. J. Brooks - A. I. J. Forrester - A. J. Keane

University of Southampton, Southampton, UK, chris.brooks@soton.ac.uk

S. Shahpar

Aerothermal Methods Group, Rolls-Royce plc, Derby, UK

ABSTRACT

This paper demonstrates the application of Kriging based approximations to the aerodynamic shape optimisation of a transonic compressor rotor where multiple fidelities of analysis are available. An extension to the co-Kriging method is implemented, and attention is paid to the efficient construction of the approximate models. The initial data sample within each level of the model is selected using space-filling designs, and the covariance structure of the co-Kriging models is obtained using a maximum likelihood estimation based on gradient optimisation with adjoint functions. An efficient leave-one-out cross-validation method is used to estimate model accuracy and is used to guide initial model refinement.

NOMENCLATURE

η	isentropic efficiency	$\dot{m}$	massflow rate
P_r	pressure ratio	ρ	scaling hyperparameter
$\hat{\mu}$	regression model coefficients	$\hat{\sigma}_Z^2$	residual process variance
$\mathcal{L}$	concentrated log likelihood	$\hat{\sigma}_{\hat{y}}^2$	prediction variance

INTRODUCTION

The performance of turbomachinery components is often assessed at the design stage using high-fidelity Reynolds-Averaged Navier-Stokes (RANS) computational fluid dynamics solvers. Such solvers are computationally expensive in nature, and their use in design optimisation becomes infeasible for large design problems given a limited computational budget. This paper addresses the aerodynamic shape optimisation of a transonic compressor rotor on a limited computational budget.

Kriging response surface models of the rotor performance measures are used as an inexpensive approximation in place of direct solver calls. Kriging has found extensive use in the aerospace industry, providing a highly flexible method capable of representing complicated responses (Forrester et al., 2008). Given a strict computational budget, any global optimisation framework based on Kriging models must be able to cope with potentially sparse data sampling and high problem dimensions.

Computational Fluid Dynamics (CFD) simulation codes can often be run at different levels of complexity both in terms of the mathematical model being used, and in terms of the resolution of the computational grid. For example, a high-fidelity, slow solution can be obtained on a fine grid and a low-fidelity solution with a much faster computation time can be obtained on a coarse grid. Here, a multi-fidelity co-Kriging method based on the auto-regressive model of Kennedy and O'Hagan (2000) is described, where response surface models of the high-fidelity function are constructed from a combination of computationally expensive high-fidelity data and cheap low-fidelity data. Global response surface accuracy is improved by correlating densely sampled low-fidelity data with sparsely sampled high-fidelity data. Strategies for updating the Kriging models with new designs are outlined,

and are designed to maximise both rotor performance, global response surface accuracy and usage of the available parallel computing resource. The aerodynamic shape optimisation of a transonic compressor rotor is used to demonstrate the methodology.

CO-KRIGING WITH TWO LEVELS OF ANALYSIS

Co-Kriging has been used for some time in the area of geostatistics as a multivariate method for the joint estimation of multiple output variables, which incorporates both spatial and intervariable correlation (Chiles and Delfiner, 1999). In the current problem, the estimation of the output of the expensive function is required, given data at two levels of fidelity. The following assumption is made about the two fidelity levels

$$\mathrm{Cov}\{Y_e(\mathbf{x}^{(i)}), Y_c(\mathbf{x})|Y_c(\mathbf{x}^{(i)})\} = 0\ , \forall \mathbf{x} \neq \mathbf{x}^{(i)}\ , \tag{1}$$

where Y_c is the cheap model and Y_e is the expensive model. This is a Markov property, which implies that, at a sample point, no more can be learnt about the expensive function from the cheap function if the data of the expensive function is known at that point. To construct a co-Kriging model, the auto-regressive formulation of Kennedy and O'Hagan (2000) is used, which approximates the expensive function as the sum of the scaled cheap function realization, and a Gaussian process Y_d representing the difference between the expensive function and the scaled cheap function

$$Y_e(\mathbf{x}) = \rho Y_c(\mathbf{x}) + Y_d(\mathbf{x})\ . \tag{2}$$

The classical Kriging approach assumes that the observed data attribute $\left\{y(\mathbf{x}), \mathbf{x} \in R^d\right\}$ is the realization of a Gaussian process F, composed of a mean model $m(\mathbf{x})$ and a mean zero random process $Z(\mathbf{x})$, and subject to random noise $\epsilon(\mathbf{x}) \sim N(0, \sigma_\tau^2)$, described by

$$Y(\mathbf{x}) = F(\mathbf{x}) + \epsilon(\mathbf{x}) = m(\mathbf{x}) + Z(\mathbf{x}) + \epsilon(\mathbf{x})\ . \tag{3}$$

The covariance of Y is

$$\mathrm{Cov}\left\{y\left(\mathbf{x}^{(i)}\right), y\left(\mathbf{x}^{(j)}\right)\right\} = \hat{\sigma}_Z^2 C(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \hat{\sigma}_Z^2 R(\mathbf{h}) + \hat{\sigma}_\tau^2 \delta(\mathbf{h})\ , \tag{4}$$

where $\hat{\sigma}_Z^2$ is the residual process variance, and the lag $\mathbf{h} = \mathbf{x}^{(i)} - \mathbf{x}^{(j)}$. On assuming a constant mean model, the best linear unbiased predictor (BLUP) is given by

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \mathbf{r}^T \mathbf{C}^{-1}(\mathbf{Y} - \mathbf{1}\hat{\mu})\ , \tag{5}$$

where $\mathbf{r}(\mathbf{x}) = [R(\mathbf{x}^{(1)}, \mathbf{x}), \cdots, R(\mathbf{x}^{(n)}, \mathbf{x})]^T$ and $\mathbf{C}$ is the correlation matrix (Chiles and Delfiner, 1999). The prediction variance can be determined as follows

$$\hat{\sigma}_{\hat{y}}^2(\mathbf{x}) = \hat{\sigma}_Z^2 \left[C\left(\mathbf{x}, \mathbf{x}\right) - \mathbf{r}^T \mathbf{C}^{-1} \mathbf{r} + \left(\mathbf{1}^T \mathbf{C}^{-1} \mathbf{r} - 1\right)^T \left(\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}\right)^{-1} \left(\mathbf{1}^T \mathbf{C}^{-1} \mathbf{r} - 1\right)\right]\ . \tag{6}$$

The reinterpolation method of Forrester et al. (2008) can be used to to obtain a noise-free Kriging prediction which includes a component modelling the simulation error. The predictions of the noisy Kriging model at the data sample locations $\hat{\mathbf{Y}}(\mathbf{X})$ are used as the data attribute in a noise-free Kriging model, the predictor of which is given by

$$\hat{y}_r(\mathbf{x}) = \hat{\mu} + \mathbf{r}^T \mathbf{R}^{-1}(\hat{\mathbf{Y}}(\mathbf{X}) - \mathbf{1}\hat{\mu})\ . \tag{7}$$

The prediction variance is given by

$$\hat{\sigma}_{\hat{y}_r}^2(\mathbf{x}) = \hat{\sigma}_Z^2 \left[R\left(\mathbf{0}\right) - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} + \left(\mathbf{1}^T \mathbf{R}^{-1} \mathbf{r} - 1\right)^T \left(\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}\right)^{-1} \left(\mathbf{1}^T \mathbf{R}^{-1} \mathbf{r} - 1\right)\right]\ . \tag{8}$$

The prediction variance now reduces to zero at the sample locations, conforming to the deterministic nature of the sample data. Initially, it is assumed that both the cheap and expensive observations are

subject to input independent random noise, ϵ_c and ϵ_e, respectively. However, the aim is to correlate only the underlying features of the cheap function with those of the expensive function, rather than one which is "corrupted" by noise ϵ_c. The random field describing the expensive function is constructed as the scaled sum of the noise-free cheap function model $Y_{c,r} = m_{c,r} + Z_{c,r}$, obtained via the reinterpolation method, and a difference process Y_d

$$Y_e(\mathbf{x}) = \rho Y_{c,r}(\mathbf{x}) + Y_d(\mathbf{x}) \,. \tag{9}$$

Different correlation structures are permitted at the two levels, thus the data covariance matrix is constructed from component matrices as follows

$$\mathbf{C}_{\tau_d} = \begin{bmatrix} \hat{\sigma}^2_{Z_{c,r}} \mathbf{R}_c(\mathbf{X}_c, \mathbf{X}_c) & \rho \hat{\sigma}^2_{Z_{c,r}} \mathbf{R}_c(\mathbf{X}_c, \mathbf{X}_e) \\ \rho \hat{\sigma}^2_{Z_{c,r}} \mathbf{R}_c(\mathbf{X}_e, \mathbf{X}_c) & \rho^2 \hat{\sigma}^2_{Z_{c,r}} \mathbf{R}_c(\mathbf{X}_e, \mathbf{X}_e) + \hat{\sigma}^2_{Z_d} \left[\mathbf{R}_d(\mathbf{X}_e, \mathbf{X}_e) + \tau_d \mathbf{I}_e \right] \end{bmatrix} \,. \tag{10}$$

The BLUP for the expensive model is given by

$$\hat{y}_e(\mathbf{x}) = \mathbf{q}^T \hat{\mu} + \mathbf{r}^T \mathbf{C}^{-1}_{\tau_d} (\mathbf{Y} - \mathbf{Q}\hat{\mu}) \,, \tag{11}$$

where

$$\mathbf{q} = \begin{bmatrix} \rho \\ 1 \end{bmatrix} , \hat{\mu} = \begin{bmatrix} \hat{\mu}_c \\ \hat{\mu}_d \end{bmatrix} , \mathbf{r}(\mathbf{x}) = \begin{bmatrix} \rho \hat{\sigma}^2_{Z_{c,r}} R_c(\mathbf{X}_c, \mathbf{x}) \\ \rho^2 \hat{\sigma}^2_{Z_{c,r}} R_c(\mathbf{X}_e, \mathbf{x}) + \hat{\sigma}^2_{Z_d} R_d(\mathbf{X}_e, \mathbf{x}) \end{bmatrix} , \mathbf{Q} = \begin{bmatrix} \mathbf{1}_c & \mathbf{0}_c \\ \rho \mathbf{1}_c & \mathbf{1}_e \end{bmatrix} .$$

The residual difference process variance $\hat{\sigma}^2_{Z_d}$ and the GLS estimate of the regression model coefficients $\hat{\mu}_d$, are given by

$$\hat{\sigma}^2_{Z_d} = \frac{1}{n_e} (\mathbf{Y}_d - \mathbf{1}_e \hat{\mu}_d)^T \mathbf{C}_d^{-1} (\mathbf{Y}_d - \mathbf{1}_e \hat{\mu}_d) \,, \quad \hat{\mu}_d = \left(\mathbf{1}_e^T \mathbf{C}_d^{-1} \mathbf{1}_e \right)^{-1} \mathbf{1}_e^T \mathbf{C}_d^{-1} \mathbf{Y}_d \,. \tag{12}$$

The expensive function prediction variance is determined as follows

$$\hat{\sigma}^2_{\hat{y}_e}(\mathbf{x}) = \rho^2 \hat{\sigma}^2_{Z_{c,r}} + \hat{\sigma}^2_{Z_d} (1 + \tau_d) - \mathbf{r}^T \mathbf{C}^{-1}_{\tau_d} \mathbf{r} + \left(\mathbf{Q}^T \mathbf{C}^{-1}_{\tau_d} \mathbf{r} - \mathbf{q} \right)^T \left(\mathbf{Q}^T \mathbf{C}^{-1}_{\tau_d} \mathbf{Q} \right)^{-1} \left(\mathbf{Q}^T \mathbf{C}^{-1}_{\tau_d} \mathbf{r} - \mathbf{q} \right) \,. \tag{13}$$

Constructing Kriging models for the cheap and expensive function data

The cheap function data $\mathbf{Y_c}$ is considered to be independent of the expensive function data $\mathbf{Y_e}$, so the Kriging models for each function can be constructed in a separate manner. The covariance structure of the cheap function random process Z_c is estimated from the cheap sample data using maximum likelihood estimation. The correlation used here is of the form

$$R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}; \theta, \mathbf{p}) = \prod_{l=1}^{d} \exp \left(-\theta_l |x_l^{(i)} - x_l^{(j)}|^{p_l} \right) , \tag{14}$$

where θ is a width parameter controlling the spatial extent of a sample point's influence, and p is a parameter controlling the local smoothness. The $2d$ hyperparameters $\Theta_c \equiv (\theta_c, \mathbf{p}_c)$ are then estimated by maximising the concentrated log likelihood function given by

$$\mathcal{L}_c(\Theta_c) = -\frac{n_c}{2} \ln \left(\hat{\sigma}^2_{Z_c} \right) - \frac{1}{2} \ln |\mathbf{C}_c| \,. \tag{15}$$

The generalised least squares estimates of the residual process variance and the regression model coefficients are given by

$$\hat{\sigma}^2_{Z_c} = \frac{1}{n_c} (\mathbf{Y}_c - \mathbf{1}_c \hat{\mu})^T \mathbf{1}_c^{-1} (\mathbf{Y}_c - \mathbf{1}_c \hat{\mu}_c) \,, \quad \hat{\mu}_c = \left(\mathbf{1}_c^T \mathbf{C}_c^{-1} \mathbf{1}_c \right)^{-1} \mathbf{1}_c^T \mathbf{C}_c^{-1} \mathbf{Y}_c \,. \tag{16}$$

The computational cost of hyperparameter optimisation can become significant for high dimensional problems and large data samples. A gradient-based approach is advocated here, where derivatives

of the likelihood function are obtained in an efficient manner via adjoint methods. Standard analytic methods can be applied to obtain the derivatives, but Toal et al. (2009) have demonstrated that the adjoint formulation provides a more efficient method. The approach taken here is to apply reverse algorithmic differentiation to the original concentrated log likelihood function by handcoding, as opposed to simply applying reverse mode automatic differentiation software to the original computer code. Highly efficient adjoint routines are obtained at a higher level by using well known matrix derivative results in numerical linear algebra (Giles, 2008). The reverse algorithmic differentiation of the likelihood function yields the adjoint correlation matrix given by

$$\overline{\mathbf{C}_c} = \frac{1}{2\hat{\sigma}_Z^2}\mathbf{C}_c^{-1}\left(\mathbf{Y}_c - \mathbf{1}_c\hat{\mu}_c\right)\left(\mathbf{Y}_c - \mathbf{1}_c\hat{\mu}_c\right)^T\mathbf{C}_c^{-1} - \frac{1}{2}\mathbf{C}_c^{-1} \,. \tag{17}$$

The chain rule is then used to obtain the likelihood derivatives with respect to the hyperparameters of the correlation function giving

$$\frac{\partial\mathcal{L}}{\partial\Theta_l} = \sum_{ij}\overline{\mathbf{C}_{c_{ij}}}\frac{\mathbf{C}_{c_{ij}}}{\partial\Theta_l} \,. \tag{18}$$

If a noise model is included, $2d+1$ hyperparameters $\Theta \equiv (\theta, \mathbf{p}, \tau)$ must then be estimated, where $\tau = \hat{\sigma}_\tau^2/\hat{\sigma}_Z^2$ is the noise parameter. The computational cost of evaluating the adjoint likelihood function is independent of the number of hyperparameters and is around twice that of evaluating the standard likelihood function (Toal et al., 2009). The covariance structure of the difference process is also estimated via maximum likelihood. The scaling parameter ρ is obtained together with the $2d+1$ correlation hyperparameters by optimising the concentrated log-likelihood function for the difference model given by

$$\mathcal{L}_d(\Theta_d) = -\frac{n_e}{2}\ln\left(\hat{\sigma}_{Z_d}^2\right) - \frac{1}{2}\ln\left|\mathbf{C}_d\right| \,. \tag{19}$$

The likelihood derivative with respect to the scaling parameter is given by

$$\frac{\partial\mathcal{L}_d}{\partial\rho} = -\frac{1}{\hat{\sigma}_{Z_d}^2}\mathbf{Y}_c\left(\mathbf{X}_e\right)^T\mathbf{C}_d^{-1}\left(\mathbf{Y}_d\left(\mathbf{X}_e\right) - \mathbf{1}_e\hat{\mu}_d\right) \,. \tag{20}$$

Efficient cross-validation

In the absence of an independent data sample subset, assessment of the Kriging model accuracy is often based on a leave-one-out cross-validation. The kriging predictor for the left-out data point is usually found by solving a new Kriging system using the remaining $n-1$ data points. However, Dubrule (1983) showed that the new predictor can be obtained directly from the inverse of original Kriging system matrix. An efficient leave-one-out cross-validation method can be obtained which reduces the typical computational cost from $\mathcal{O}(n^4)$ to $\mathcal{O}(n^3)$. The Kriging system block matrix and it's inverse are given by

$$\mathbf{K} = \begin{bmatrix} \mathbf{C} & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{0} \end{bmatrix} \,, \qquad \mathbf{K}^{-1} = \begin{bmatrix} \mathbf{B} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{W} \end{bmatrix} \,, \tag{21}$$

noting that $Q = \mathbf{1}_c$ for the cheap function Kriging model. The elements of $\mathbf{K}^{-1}$ are obtained using the blockwise matrix inverse lemma, and the cross-validation estimates of the prediction error are given by

$$\mathbf{e}^\star = \mathbf{Y} - \hat{\mathbf{Y}}^\star = \mathbf{D}_\mathbf{B}^{-1}\mathbf{B}\mathbf{Y} \,, \tag{22}$$

where $\mathbf{D}_\mathbf{B}^{-1}$ is a diagonal matrix with entries $d_{ii} = 1/b_{ii}$. This method takes into account the effect of estimating the regression coefficients. Cross-validation estimates of the prediction variances are given

(a) Coarse mesh (b) Fine mesh

Figure 1: **Meridonal views of the cheap and expensive model computational meshes.**

by $\hat{\sigma}^{2\star}_{\hat{y}_i} = \hat{\sigma}^2_Z d_{ii}$. Several metrics for assessing the Kriging model accuracy may then be formed, the most pertinent being the root mean squared cross-validation error, the root mean squared standardised cross-validation error and the Pearson correlation coefficient, given respectively by:

$$\mathrm{RMSE}^\star = \sqrt{\tfrac{1}{n}\mathbf{e}^{\star T}\mathbf{e}^\star}\,, \quad \mathrm{RMSSE}^\star = \sqrt{\tfrac{1}{n}\sum_{i=1}^n \tfrac{e_i^{\star\,2}}{\hat{\sigma}^{2\star}_{\hat{y}_i}}}\,, \quad R^{2\star} = \left(\frac{\mathrm{Cov}(\mathbf{Y},\hat{\mathbf{Y}}^\star)}{\sqrt{\mathrm{Var}(\mathbf{Y})\mathrm{Var}(\hat{\mathbf{Y}}^\star)}}\right)^2. \tag{23}$$

MULTI-FIDELITY OPTIMISATION WITH CO-KRIGING MODELS

In this section we consider the problem of aerodynamic shape optimisation of a transonic compressor rotor. The co-Kriging model is applied to data from three-dimensional viscous flow solutions for NASA compressor rotor 37, a popular public domain testcase (Reid and Moore, 1978). To obtain multi-fidelity data, flow solutions on two computational grids were obtained using a non-linear, unstructured viscous flow solver, HYDRA (Lapworth and Shahpar, 2004). The Reynolds-Averaged steady Navier-Stokes (RANS) equations with the Spalart-Allmaras turbulence model were solved. Convergence to steady-state was accelerated by employing a multigrid algorithm with preconditioning. A four-level multigrid was used for the examples here. A multi-block meshing system PADRAM was used for the geometry parameterisation and meshing (Lapworth and Shahpar, 2004).

Methodology

The cheap data was obtained using a coarse mesh model, shown in Fig. 1a, consisting of 240000 nodes, whilst the expensive data was obtained using a fine mesh of 740000 nodes, shown in Fig. 1b. The tip gap in both models was gridded with six mesh layers. In addition, the expensive model geometry included a 2.5mm radius fillet at the hub as per the original test geometry (Dunham, 1998), and a stationary platform which abuts the hub without a gap. Note that, in the original test geometry there is a 0.75mm gap at either end of the rotating hub separating it from the stationary platform, which is not modelled here. Several authors (such as Shabbir et al. (1997)) have shown that the presence of the hub cavities (either with or without a net leakage flow) can improve the predictions of rotor 37. The rotor characteristics of the datum fine mesh solution are compared with the published experimental results in Fig. 6. The numerical result for pressure ratio agrees well with the experimental data, but the efficiency is underestimated by 2–3%. The same underestimation has been noted in previous studies (Dunham, 1998; Samad and Kim, 2008). The radial profiles are compared in Fig. 5, and highlight the discrepancy in pressure ratio in the hub region attributed to the gap-flow. The efficiency is generally underestimated in the region from mid-span to tip.

The design space was defined by 28 parameters, which included 5 types of design perturbations: sweep, lean, skew, leading and trailing edge re-cambering. Each type of design pertubation was

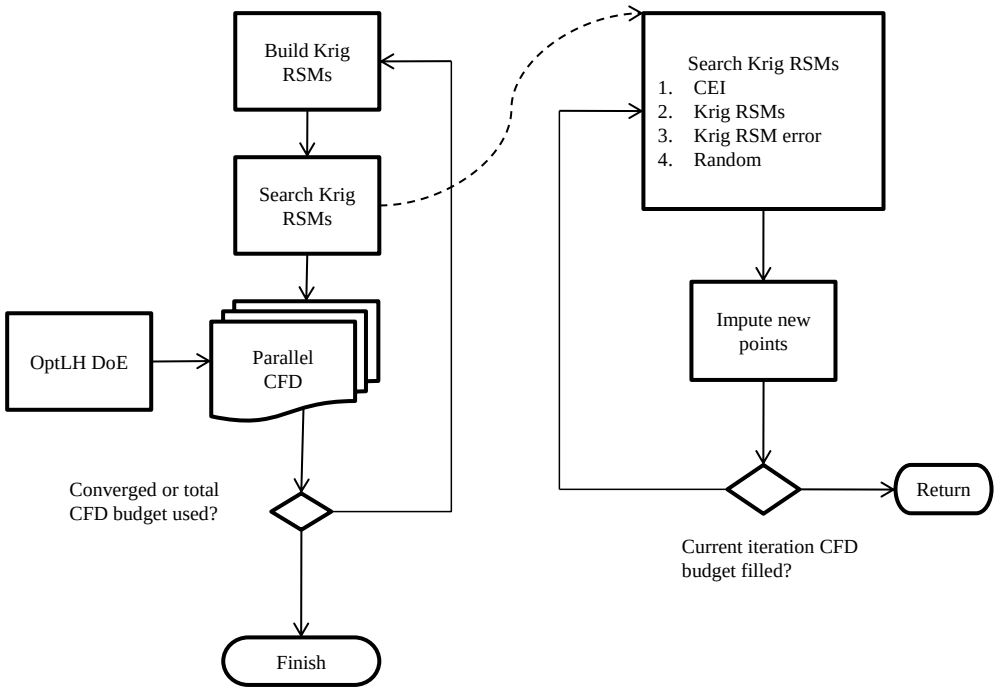


Figure 2: **Optimisation strategy based on Kriging response surface methods.**

applied at 5 locations along the blade span (20%, 40%, 60%, 80% and 100% of blade height). In addition, skew, leading and trailing edge re-cambering were allowed to vary at the blade root. The spanwise variation of the design pertubations was interpolated using NURBS.

The aim of the optimisation was to maximise the stage isentropic efficiency η of the rotor at the working conditions (97.67% choke massflow, 100% engine speed), whilst minimising the variation in the stage pressure ratio $P_r = {}^{P_{outlet}}/_{P_{inlet}}$ to within 1% and massflow rate $\dot{m}$ to within 0.5% of the values of the datum rotor geometry. The optimisation strategy, as outlined in Fig. 2, is derived from the Efficient Global Optimisation method of Jones et al. (1998). Starting with a space-filling Optimised Latin Hypercube DoE, Kriging models of the objective and constraints are tuned and constructed, the Kriging models are used in extensive global searches to obtain new design locations which are then evaluated in parallel. The evaluated designs are then added to the existing database of results, and the Kriging models are updated. Given a parallel computing resource of N_{NODES} nodes and a total computing budget of N_{CFD} CFD evaluations, a total of $\approx N_{\mathrm{CFD}}/N_{\mathrm{NODES}}$ database updates are available.

The parallel resource is filled at each update iteration by N_{NODES} new CFD simulations, the design locations of which are obtained by using a number of criteria listed in order of preference in Fig. 2. An unconstrained search of the Constrained Expected Improvement (CEI) criterion is the preferred method for selecting new design locations (Forrester et al., 2008). In the event that no new design locations are found and the parallel resource is not yet filled, the Kriging model predictions are used in a constrained search of the design space. The next search undertaken is an unconstrained search of the Kriging model errors, where the Kriging error estimate $\hat{\sigma}^2_{\hat{y}}$ is maximised to find the location of maximum uncertainty in the model prediction. Where a number of Kriging models are being used, the model with the lower cross-validation correlation coefficient $R^{2\star}$ is given preference. If the parallel resource is still not yet filled, new design locations are selected at random, with the condition that they must not coincide with existing designs within the database.

When a new design location is found the response values are imputed, penalised using a statistical

bound $\hat{y} \pm \hat{\sigma}^2$ (Forrester et al., 2008), added to the database of results and used to construct the Kriging models without re-tuning. Searches for further design locations will be diverted away from locations previously found within the current update iteration. This avoids the clustering of points around just one promising location, allowing other promising areas of the design space to be explored effectively. Once the new designs have been evaluated by the CFD software, the imputed response values are replaced by the actual values. In the event of any CFD evaluations failing, the failed designs are imputed and maintained in the database of results. Doing this avoids stalling the optimisation process and further searches are diverted away from regions where CFD solutions are failing. For consistency, the Kriging models are only tuned using actual response data, and any failed points are re-imputed before the Kriging models are constructed at each update iteration.

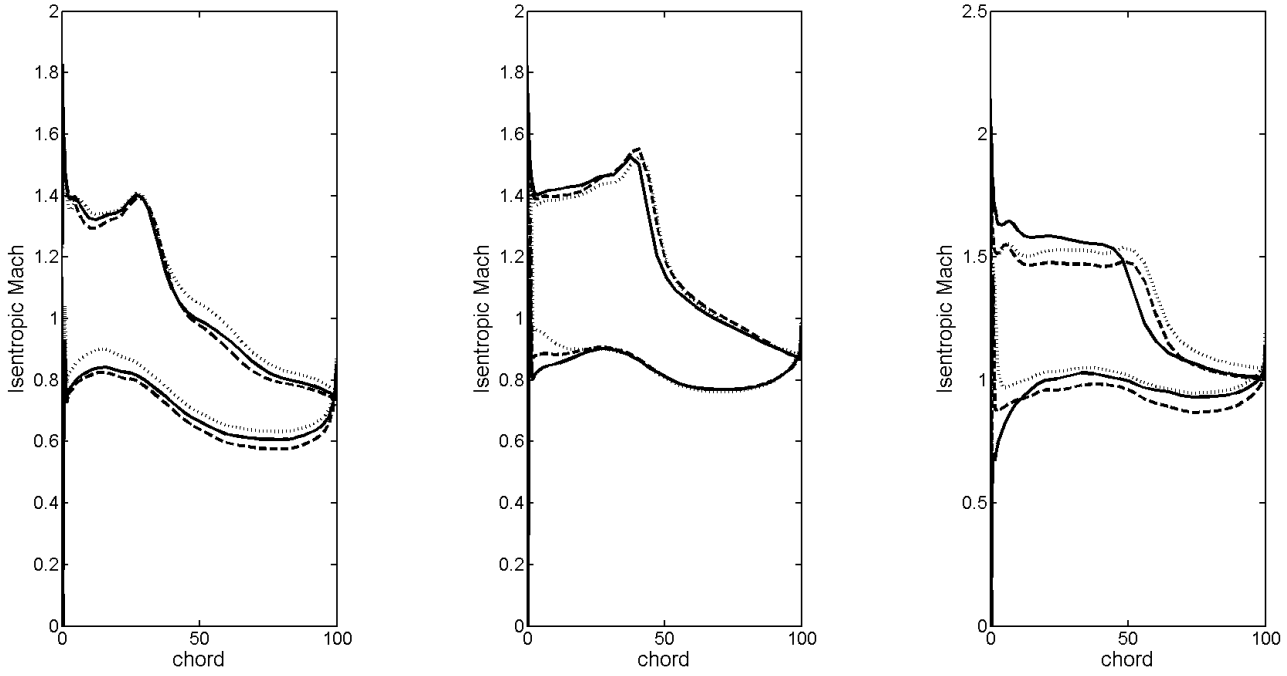
For the current rotor optimisation, a strict computational budget of $N_{\mathrm{CFD}} = 160$ CFD evaluations is prescribed, with a parallel resource of $N_{\mathrm{NODES}} = 10$ compute nodes made available. To provide a comparison, an optimisation based upon standard Kriging models is undertaken in addition to an optimisation based on multi-fidelity co-Kriging models. The parallel computing budget for the multi-fidelity optimisation is split between the cheap and expensive model CFD evaluations. Given that the cheap model has a computational cost of around one third that of the expensive model, 5 nodes are allocated to evaluate 15 cheap model CFD solutions, whilst 5 nodes are allocated to evaluate 5 expensive model CFD solutions. For the initial DoE, an optimal Latin Hypercube DoE of 15 points is defined and used for the cheap model design locations. The 5 design locations for the expensive model are selected as a space-filling subset of the original 15 points (Forrester et al., 2008). At each update iteration, the cheap model is evaluated at the same design locations as those found searching the co-Kriging model of the expensive data. The remaining 10 design locations for updating Kriging models of the cheap data are obtained by searching the cheap Kriging model errors $\hat{\sigma}^2_{\hat{y}_c}$. The total cost per update iteration is then equivalent to 10 expensive model CFD evaluations.

Results

The results of the two optimisations are summarised in Table 1. The design obtained using standard Kriging models provided a 1.79% improvement in efficiency, whilst satisfying the constraints. This is in line with previous optimisation studies (Samad and Kim, 2008). The design obtained using multi-fidelity co-Kriging models provided a 2.34% improvement in efficiency, whilst satisfying the constraints. A large proportion of cheap model CFD runs failed to converge due to the sparse mesh, but by imputing the response values, the failures did not adversely affect the optimisation progress. By comparing the cross-validation error metrics it can be seen that the multi-fidelity co-Kriging models provide a much improved estimation of the responses, in particular the constraint responses, which were estimated quite poorly by the standard Kriging models. The isentropic Mach number surface distributions for each design are compared in Fig. 3. The near-hub and mid-span distributions are similar, but the near-tip distributions show the shock is reduced and is moved towards the trailing edge on the optimised designs. Some blade sections along the span are compared in Fig. 4 along with the 3D geometry. On average over the blade span, both blades are swept backward in the flow direction, but the multi-fidelity design has a more non-uniform radial profile with a stronger backward sweep at the tip. Both exhibit lean at the mid-point away from the direction of rotation, but vice versa towards the endwalls. The multi-fidelity design exhibits a slightly more aggressive radial twist due to the skew of the blade sections changing gradually from the direction of rotation at the hub to the opposite direction at the tip, further delaying the shock in the tip region (and the resultant flow separation). Radial profiles of the pressure ratio and efficiency are shown in Fig. 5. The design obtained using multi-fidelity Kriging models provides improvement in efficiency from 10% to 80% blade height, whilst the design obtained using standard kriging models shows improvements from 30% to 80% blade height. The general backward sweep, forward lean nature of the designs found

RSM	% failed	Response	% datum	RMSE*	RMSSE*	$R^{2\star}$
Standard	3.1	η	1.79	0.0065	1.12	0.67
		P_r	-0.84	0.1636	1.14	0.025
		$\dot{m}$	0.2	0.2865	1.4	8.6e-5
Multi-fidelity	3.9 (39.8)	η	2.34	0.0022 (0.0026)	0.91 (0.96)	0.93 (0.97)
		P_r	-0.98	0.0014 (0.0041)	0.92 (0.95)	0.99 (0.95)
		$\dot{m}$	0.27	6.4e-4 (0.0021)	0.87 (0.93)	0.96 (0.90)

Table 1: **Rotor optimisation results and final Kriging model error metrics. Cheap function Krig values in brackets.**



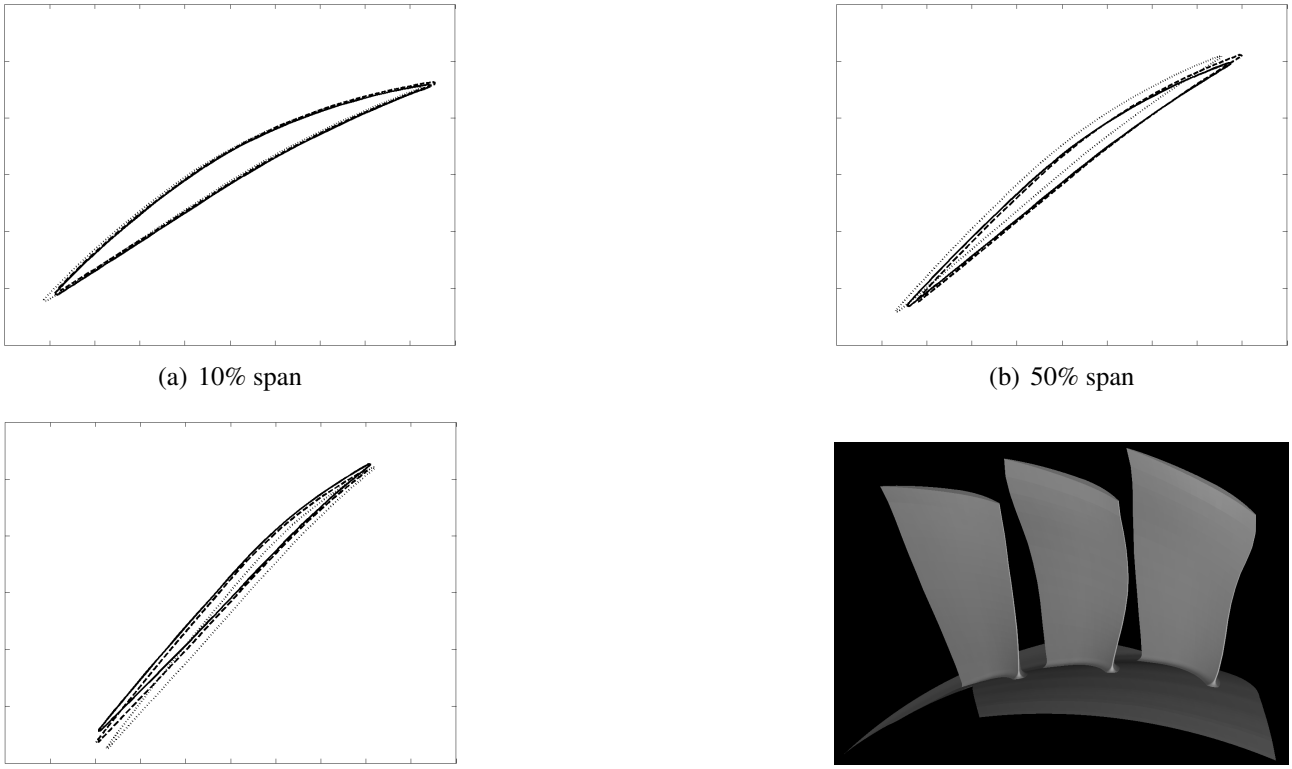
(a) 10% span (b) 50% span (c) 90% span

Figure 3: **Isentropic Mach number surface distributions at different spanwise locations. $-$, datum; $--$, standard Krig design; $\cdots$, multi-fidelity Krig design.**

here conforms with the findings of Benini and Biollo (2007). The full rotor characteristics are shown in Fig. 6. By plotting the rotor characteristic we find that the increase in efficiency of the optimised designs is consistent from choke to stall, however, the designs produce a large change in the choke massflow which is unacceptable from an engineering standpoint. In future, this could be minimised by implemented a constraint on the variation of the choke massflow from the datum, at the expense of an extra CFD simulation at each design point.

CONCLUSIONS

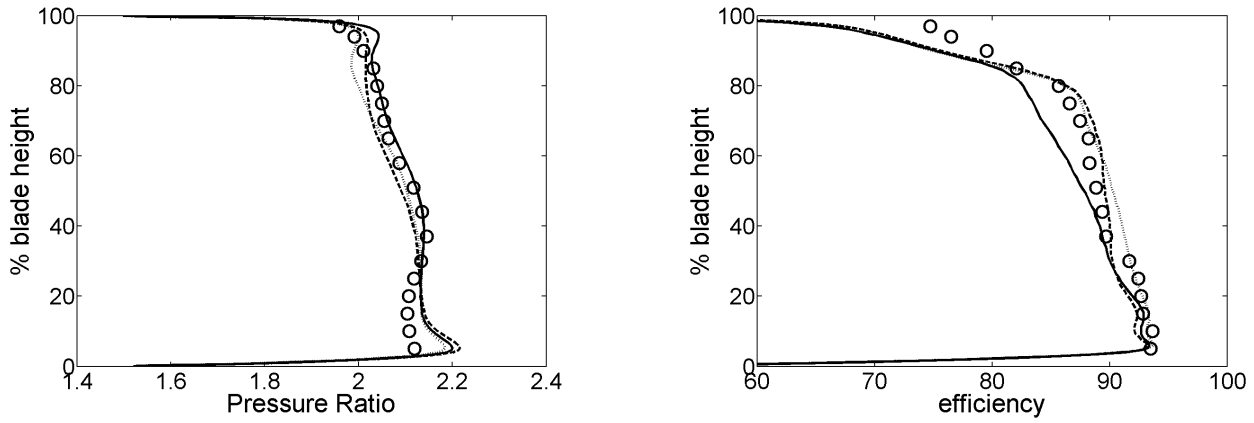
An optimisation strategy based on multi-fidelity co-Kriging models has been outlined and applied to the aerodynamic shape optimisation of a transonic compressor rotor. The use of a co-Kriging model to correlate densely sampled low-fidelity data with sparsely sampled high-fidelity data is found to outperform Kriging models based on high-fidelity data alone given the same computational budget. Efficient tuning of the co-Kriging models was achieved by employing likelihood gradients obtained via adjoint methods. Metrics for assessing the co-Kriging model accuracy were determined from an efficient leave-one-out cross-validation.



(a) 10% span (b) 50% span

(c) 90% span (d) 3D geometry

Figure 4: **Blade geometries and profile at different spanwise locations. −, datum (left blade); − −, standard Krig design (middle blade); · · ·, multi-fidelity Krig design (right blade).**



(a) Pressure ratio (b) Efficiency

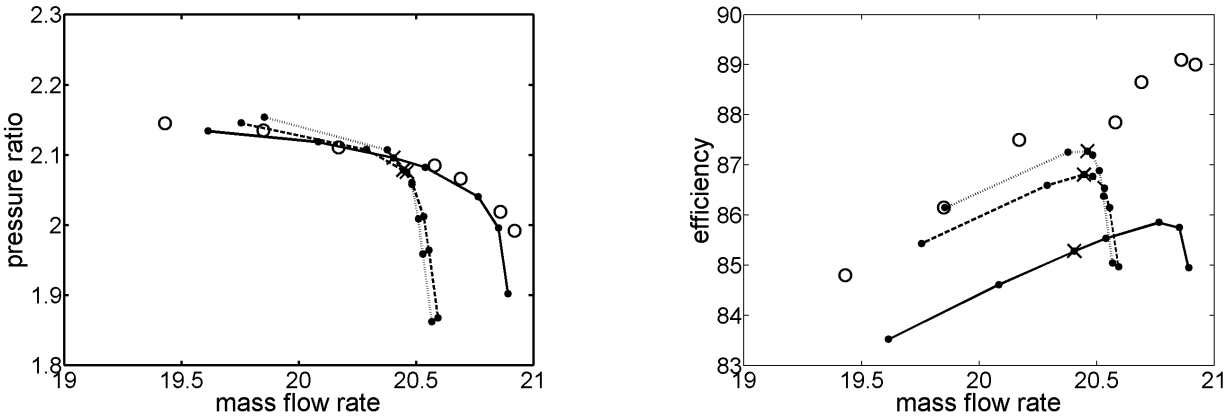
Figure 5: **Radial profiles of pressure ratio and efficiency for datum and optimised rotor blade geometries. ○, experimental data.**

ACKNOWLEDGEMENTS

This work has been undertaken as part of the Centre for Fluid Mechanics Simulation (CFMS) Core Programme, a United Kingdom consortium of industry partners and academic organisations led by Airbus, and joined by the United Kingdom Department of Trade and Industry. The continued support of Rolls-Royce plc. is greatfully acknowledge.

REFERENCES

E. Benini and R. Biollo (2007) *Aerodynamics of swept and leaned transonic compressor-rotors.* Ap-



(a) Pressure ratio (b) Efficiency

Figure 6: **Full characteristics for datum and optimised rotor blade geometries. $\times$, location of single-point optimisation.**

plied Energy, 84(10), p.1012-1027.
J.P. Chiles and P. Delfiner (1999) Geostatistics. Modeling spatial uncertainty. Wiley, New York.
O. Dubrule. (1983) *Cross validation of kriging in a unique neighborhood.* Math. Geo., 15(6), p.687-699.
J. Dunham (1998) *Cfd validation for propulsion system components.* AGARD Advisory Report AR-355.
A.I.J. Forrester, A. Sóbester, and A.J. Keane (2008) Engineering design via surrogate modelling: a practical guide. Wiley, Chichester.
M.B. Giles (2008) *Collected matrix derivative results for forward and reverse mode algorithmic differentiation.* In *Lecture Notes in Computational Science and Engineering*, 64, p.35-44.
D.R. Jones, M. Schonlau, and W.J. Welch (1998) *Efficient global optimisation of expensive black-box functions.* J. of Global Opt., 13, p.455-492.
M. Kennedy and A. O'Hagan (2000) *Predicting the output from a complex computer code when fast approximations are available.* Biometrika, 87, p.1-13.
B.L. Lapworth and S. Shahpar (2004) *Design of gas turbine engines using cfd.* In *ECCOMAS 2004*, Jyvaskyla.
L. Reid and R.D. Moore (1978) *Design and overall performance of four highly-loaded, high-speed inlet stages for an advanced, high-pressure-ratio core compressor.* NASA Tech Report TP-1337.
A. Samad and K.-Y. Kim (2008) *Shape optimization of an axial compressor blade by multi-objective genetic algorithm.* Proc. IMechE, Part A: J. Power and Energy, 222(6), p.599-611.
A. Shabbir, M.L. Celestina, J.J. Adamczyk, and A.J. Strazisar (1997) *Effect of hub leakage flow on two high speed axial flow compressor rotors.* ASME Paper 97-GT-346.
D.J.J. Toal, A.I.J. Forrester, N.W. Bressloff, A.J. Keane, and C. Holden (2009) *An adjoint for likelihood maximization.* Proc. R. Soc. A, 465(2111), p.3267-3287.