Combining a Reduced Polynomial Chaos Expansion Approach with Universal Kriging for Uncertainty Quantification

Justin Weinmeister*, Nelson Xie†, Xinfeng Gao‡
Computational Fluid Dynamics and Propulsion Laboratory
Colorado State University, Fort Collins, CO 80523, USA
Aditi Krishna Prasad§and Sourajeet Roy¶
High Speed System Simulation Laboratory
Colorado State University, Fort Collins, CO 80523, USA

Engineering design optimization studies, based on the large number of computational fluid dynamics simulations necessary for uncertainty quantification, are computationally expensive. Polynomial chaos expansion methods have the potential to save computational costs by reducing the number of input design parameters. Kriging methods are able to accurately predict off-design values and give an estimate of their error. However, each has its limitations. In this paper, we combine a reduced dimensional polynomial chaos approach with a universal Kriging method as a new non-intrusive metamodeling method for fast uncertainty quantification and optimization in a simplified engine nacelle inlet design. Its performance is benchmarked against the reduced dimensional polynomial chaos approach and universal Kriging. Results show the reduced-polynomial-chaos-Kriging method gives more accurate results than the reduced dimensional polynomial chaos approach for non-smooth solutions. However, the new method is highly-dependent on the experimental design. The application of a standalone Kriging method on the reduced model produced excellent stability and indicates refinement of the method is possible.

I. Introduction

Computational fluid dynamics (CFD) is an effective tool in the design process for new aircraft propulsion systems to improve aerodynamic performance and reduce aircraft fuel consumption. However, the number of CFD model simulations required for optimization over a large parameter space can be computationally prohibitive. Recently, we developed a novel reduced dimensional polynomial chaos method (RPC)[1] for model parameter space reduction which reduces the number of CFD simulations required. This approach begins by performing a high-dimensional model representation of an engine nacelle inlet design problem which leads to quantification of the relative impact of each random model parameter acting alone on the variance of the CFD solution variables. The resulting impact factors are then used to guide the parameter reduction by identifying only the parameters of maximum importance to the nacelle inlet design. The reduction of the parameter space translates to a substantially smaller number of CFD simulations required to perform the parametric uncertainty quantification study. Nevertheless, due to the highly nonlinear CFD model for the fluid dynamics problems, such as flows around/in the engine nacelle inlet geometry, the solution does not show smooth dependency on the input parameters. The RPC method is not reliable in these cases. In order to improve the accuracy of this method, a higher order of PC expansion would be needed. However, the

*Undergraduate Research Assistant, Email: jrwein@rams.colostate.edu AIAA Student Member
†Research Associate, Email: nelson.xie@colostate.edu
‡Assistant Professor, Email: xinfeng.gao@colostate.edu, AIAA Member
§Graduate Research Assistant, Email: aditikp@colostate.edu
¶Assistant Professor, Email: sjeetroy@engr.colostate.edu

number of simulations required to fit a higher order expansion scales exponentially with the design space. Given the fact that each CFD simulation for practical geometries under realistic operating conditions is very computationally expensive, this renders that even a RPC approach can be inapplicable for strongly nonlinear, high-dimensional problems. This is known as the curse of dimensionality. Kriging methods, best linear unbiased predictors developed for geostatistics,[2] have recently seen use in CFD applications for their ability to accurately predict off-design values for non-linear problems.[3] The most basic method is ordinary Kriging which assumes a constant mean for each solution variable across the design space. This method is limited when these solution variables have significant trends. The universal Kriging (UK) method accounts for these trends in the data. However, the UK method requires that this trend is known before fitting. Schöbi, Sudret, and Wiart[4] have approached this problem by combining a polynomial chaos expansion approach with Kriging. Their method uses a least angle regression method to find a sparse set of the full polynomial chaos expansion to reduce computational cost.

In the present study, we propose a reduced-polynomial-chaos-Kriging method (RPC-K) which can address the weaknesses of each respective method for a minimal increase in the cost of creating the metamodel. This method uses the RPC function to describe the trend in a UK process. This metamodel simulates the output of the CFD model with a simple analytical function which can be evaluated cheaply enough to run a Monte Carlo analysis. The RPC process allows us to run a sensitivity study to reduce the number of input parameters and gives a good trend function to use in the Kriging process. The Kriging process is then run to refine the metamodel's predictions of the solution variables without having to run additional simulations to increase the number of input parameters or the order of the fit. As such, the method can alleviate the curse of dimensionality. This differs from the work of Schöbi et al.[4] where they remove the least impactful terms instead of the least impactful dimensions.

The RPC-K program is created in the MATLAB® programming language to take advantage of pre-programmed packages as the computational time for all of the elements in the RPC-K method is insignificant compared to CFD codes. The proposed method is first verified on a convection-diffusion-reaction problem, and it is then validated by the engine nacelle inlet design problem to evaluate its performance for both smooth and non-smooth flows.

II. Reduced-Polynomial-Chaos-Kriging Method

This section will describe the reduced-polynomial-chaos-Kriging (RPC-K) method in detail. For completeness and convenience, we will briefly review the RPC approach and the universal Kriging method. Then, the new non-intrusive meta-modeling approach, RPC-K, will be illustrated and discussed.

II.A. Reduced Dimensional Polynomial Chaos Expansion Approach

We have developed a dimension reduction based PC approach.[1] Although such an expansion contains a small subset of the overall PC coefficients, it is comparable in accuracy to the full-dimensional expansion since only the least impactful dimensions are removed.

Assume that the parametric uncertainty in the CFD modeling problem is represented by n mutually uncorrelated random parameters, $\vec{\omega} = [\omega_1, \omega_2, \cdots, \omega_n]^{\mathrm{T}}$ where $\vec{\omega} \in \Omega$ (a multidimensional parameter space). For example, the random parameters considered in this study are relevant to the nacelle inlet geometric parameters, flow conditions (Mach number, angle of attack, free stream conditions, etc.), and operating conditions. Typically, the conservation laws are solved by CFD modeling.

The stochastic CFD solution variables, $\mathbf{U}$, can be approximated using a PC expansion as

$$\mathbf{U}(t,\vec{\omega}) \approx \mathbf{U}^{(RPC)}(t,\vec{\omega}) = \sum_{k=0}^{p} \phi_k(\vec{\omega})\mathbf{U}_k(t)\,, \tag{1}$$

where $\phi_k(\vec{\omega})$ is the k^{th} degree multivariate orthogonal polynomial basis and $\mathbf{U}_k(t)$ is the corresponding coefficients for time t. For the purpose of sensitivity analysis, the PC expansion of any particular solution variable $u(t,\vec{\omega}) \in \mathbf{U}(t,\vec{\omega})$ can be rearranged into a high dimensional model representation (HDMR) formulation. Such a formulation describes the model solution variables as a hierarchical superposition of functions

describing the interactions among the random dimensions as[5]

$$u(t,\vec{\omega}) = u_0(t) + \sum_{i=1}^{n} u_i(t,\omega_i) + \sum_{1\le i,j\le n} u_{ij}(t,\omega_i,\omega_j) + \cdots + u_{12\ldots n}(\omega_1,\ldots,\omega_n,t)\,, \tag{2}$$

where $u_0(t)$ is the mean value of $u(t,\vec{\omega})$, $u_i(t,\omega_i)$ represents the contribution of ω_i to $u(t,\vec{\omega})$ acting alone, $u_{ij}(t,\omega_i,\omega_j)$ represents the pairwise correlated contribution of $\vec{\omega}_i$ and $\vec{\omega}_j$ to $u(t,\vec{\omega})$, and so on. Specifically, $u_0(t)$ is defined by

$$u_0(t) = u(t,\vec{\omega}^{(0)})\,, \tag{3}$$

where the superscript (0) in $\vec{\omega}^{(0)}$ represents that the random variables assume their mean values. Assuming that the degree of expansion required to minimize the error of the approximation of Eq. (2) below a prescribed tolerance[6] is m, the total number of terms in Eq. (1) is $p+1 = \frac{(n+m)!}{n!m!}$. In this work, according to the principle of sparsity of effects, the HDMR expansion of Eq. (2) is truncated to include only the most statistically significant terms (i.e. up the first order terms) as

$$u(t,\vec{\omega}) \approx u_0(t) + \sum_{i=1}^{n} u_i(t,\omega_i)\,. \tag{4}$$

The second term in Eq. (4) can be expressed using the cut-HDMR[5] as

$$u_i(t,\omega_i) = u(t,\vec{\omega})|_{\vec{\omega}^{(0)}\setminus\vec{\omega}_i} - u_0(t)\,, \tag{5}$$

where the notation $\vec{\omega}^{(0)} \setminus \omega_i$ represents the vector where all components of $\vec{\omega}$ except ω_i are set to their mean values. It is noted that the terms $u_i(t,\omega_i)$ represent the impact of the random dimension ω_i on the model output acting alone. The coefficients $u_i(t)$ can always be evaluated using any non-intrusive sampling of Eq. (5).[7] Now, based on the truncated HDMR expansion given in Eq. (4), the variance of the model output can be decomposed as the sum of variances of the first order terms as

$$\sigma_u^2 \approx \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2\,, \tag{6}$$

where σ_u^2 represents the total variance of $u(t,\vec{\omega})$ and σ_i^2 represents the variance of $u_i(t,\omega_i)$. Next, the sensitivity index of $u_i(t,\omega_i)$ is defined as the ratio of the variance due to $u_i(t,\omega_i)$ to the total variance of the output $u(t,\vec{\omega})$ as

$$S_i(t) = \frac{\sigma_i^2}{\sigma_u^2}\,. \tag{7}$$

Based on Eq. (1), the above sensitivity indices can be expressed in a closed form manner as [8,9]

$$S_i(t) = \frac{\sigma_i^2}{\sigma_u^2} = \frac{\sum_{k=0}^{m} (u_i^k)^2}{\sum_{i=1}^{n}\sum_{k=0}^{m} (u_i^k)^2}\,. \tag{8}$$

It is noted that the above indices are positive quantities bounded between the limits of [0,1], and their magnitude directly reflects the relative impact of each random parameter.[8,9] The sum of the sensitivities equals one. In order to account for the time varying nature of these indices, the L_1 norm of the indices of Eq. (8) are considered to identify the least impactful random parameters that can be pruned from the original random space. Thereafter, a new and sparse PC expansion can be efficiently generated from the reduced dimensional random subspace. The RPC method has been tested by Gao et al.[1] and will be used in conjunction with the universal Kriging method in the present study.

II.B. Universal Kriging

Universal Kriging is a stochastic interpolation algorithm and approximates the CFD output $\mathbf{U}(t,\vec{\omega})$ with a Gaussian process realization,

$$\mathbf{U}(t,\vec{\omega}) \approx \mathbf{U}^{(UK)}(t,\vec{\omega}) = \boldsymbol{\beta}^{\mathrm{T}} \cdot \boldsymbol{f}(\vec{\omega}) + \sigma^2 Z(\vec{\omega})\,, \tag{9}$$

where $\boldsymbol{\beta}^T \cdot \boldsymbol{f}(\vec{\omega}) = \sum_{j=0}^{p} \beta_j f_j(\vec{\omega})$ is the trend of the Gaussian process, $Z(\vec{\omega})$ is the zero-mean, unit-variance Gaussian process, and σ^2 is the Kriging variance. The Gaussian process, $Z(\vec{\omega})$, is represented by an autocorrelation function $\mathrm{R} = \mathrm{R}(|\vec{\omega} - \vec{\omega}'|; \boldsymbol{\theta})$. The autocorrelation function R correlates two samples of the input parameter space, $\vec{\omega} - \vec{\omega}'$, based on fitting its parameters, $\boldsymbol{\theta}$, to empirical covariance values. These parameters of the Kriging model are known as hyperparameters as they belong to the autocorrelation function.

Consider that there are N sets of samples for the design parameter space, $\vec{\omega} = \left\{\vec{\omega}^{(1)}, ..., \vec{\omega}^{(N)}\right\}$ with corresponding CFD output values $\mathbf{U} = \left\{\mathbf{U}^{(1)}, ..., \mathbf{U}^{(N)}\right\}$. Following Schöbi et al.[4] and Dubourg,[10] based on the generalized least-squares solution to the maximum likelihood function, the metamodel parameters $\left\{\boldsymbol{\beta}, \sigma^2\right\}$ can be obtained by

$$\boldsymbol{\beta}(\boldsymbol{\theta}) = \left(\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{U}\,, \tag{10}$$

$$\sigma^2_{\mathbf{U}}(\boldsymbol{\theta}) = \frac{1}{N}(\mathbf{U} - \mathbf{F}\boldsymbol{\beta})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{U} - \mathbf{F}\boldsymbol{\beta})\,, \tag{11}$$

where $\mathbf{R}_{ij} = \mathrm{R}(\vec{\omega}^{(i)} - \vec{\omega}^{(j)}; \boldsymbol{\theta})$ is the correlation matrix of the design parameters and $\mathbf{F}_{ij} = f_i(\vec{\omega}^{(i)})$ is a matrix of the polynomial basis calculated for each design point, the Vandermonde matrix. $\sigma^2_{\mathbf{U}}$ is the optimal Kriging variance and $\boldsymbol{\beta}$ are the optimal trend coefficients. The solution predicted at a new sample, $\vec{\omega}$, is described by a mean value $\mu_{\hat{\mathbf{U}}}$ and a variance $\sigma^2_{\hat{\mathbf{U}}}$ given as:

$$\mu_{\hat{\mathbf{U}}}(\vec{\omega}) = \boldsymbol{f}(\vec{\omega})^{\mathrm{T}}\boldsymbol{\beta} + \mathbf{r}(\vec{\omega})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{U} - \mathbf{F}\boldsymbol{\beta})\,, \tag{12}$$

where $\mu_{\hat{\mathbf{U}}}(\vec{\omega})$ is the new prediction, a Gaussian variable.

$$\sigma^2_{\hat{\mathbf{U}}}(\vec{\omega}) = \sigma^2_{\mathbf{U}}\left(1 - [\boldsymbol{f}(\vec{\omega})^{\mathrm{T}}\mathbf{r}(\vec{\omega})^{\mathrm{T}}]\begin{bmatrix}\mathbf{0} & \mathbf{F}^T\\ \mathbf{F} & \mathbf{R}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{f}(\vec{\omega})\\ \mathbf{r}(\vec{\omega})\end{bmatrix}\right), \tag{13}$$

where $\mathrm{r}_i(\vec{\omega}) = \mathrm{R}(\vec{\omega} - \vec{\omega}^{(i)}; \boldsymbol{\theta})$ is the correlation between the new sample $\vec{\omega}$ and the design points $\vec{\omega}^{(i)}$. The autocorrelation and error estimation must be modeled since they play an important role in the Kriging method.

II.B.1. Autocorrelation Functions

The autocorrelation function describes the correlation of two data points based only on the distance between them. It is evaluated between every data point in order to construct the autocorrelation matrix, $\mathbf{R}$, for Kriging. Weak stationarity of the data are assumed for ordinary Kriging processes, where the mean, variance, and covariance vary based only on separation distance and not their position.[2] Matheron developed the intrinsic hypothesis for universal Kriging which allows for nonstationary mean and thus violation of weak stationarity.[2] The intrinsic hypothesis requires either a known trend function or autocorrelation function for the resulting predictor to be unbiased, each of which is unknown for most practical problems. This is remedied by finding the maximum likelihood of the trend function. This minimizes the bias with the limit that the function will be unbiased as the sampling becomes dense. The maximum likelihood function for a Gaussian process[10] can be written as

$$L(\boldsymbol{u}|\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}) = \frac{1}{\left((2\pi\sigma^2)^N[det\mathbf{R}(\boldsymbol{\theta})]\right)^{1/2}} exp\Big[\frac{-1}{2\sigma^2}(\mathbf{Z} - \mathbf{F}\boldsymbol{\beta})^{\mathrm{T}}\mathbf{R}(\boldsymbol{\theta})^{-1}(\mathbf{Z} - \mathbf{F}\boldsymbol{\beta})\Big]\,. \tag{14}$$

This equation can be rewritten and solved for in terms of $\boldsymbol{\theta}$ alone. Simplified, this equation can be reduced to the inverse log likelihood function[10, 11]

$$\psi(\boldsymbol{\theta}) = \hat{\sigma^2}(\boldsymbol{\theta})[det\mathbf{R}(\boldsymbol{\theta})]^{1/N}\,, \tag{15}$$

where N is again the number of design points. When this equation is minimized, the hyperparameters have been optimized for the current data points and the intrinsic hypothesis can be used.

Various autocorrelation functions exist with different properties, though they all contain hyperparameters. Common autocorrelation functions are given by Webster and Oliver,[2] and Dubourg.[10] In the present study, we only use the bounded linear autocorrelation described below.

The linear model can be described by

$$\mathrm{R}(\vec{\omega}-\vec{\omega}';\vec{l})=\prod_{i=1}^{n}\begin{cases}\frac{\omega_i-\omega_i'}{l_i} & \text{for}(\omega_i-\omega_i')\leq l_i,\\ 0 & (\omega_i-\omega_i')>l_i\end{cases} \tag{16}$$

where l_i is the range parameter for input parameter i and $(\omega_i-\omega_i')$ is the distance between the same input parameter from two separate design points. The linear model is a bounded model where l_i is the finite value where the autocorrelation function reaches a maximum.

The above autocorrelation function is separable in that it is a product of univariate functions. This allows each input parameter to have its own hyperparameters and thus achieve a more accurate fit. General guidelines[2] exist for determining the correct autocorrelation function for specific data, though the process generally involves a combination of statistical techniques and visual inspection.

II.B.2. Error Estimation

Universal Kriging predicts the value at a new location as a random process described by the mean given in Eq. (12) and a variance given in Eq. (13). Because the variance prediction is local, Schöbi et al.[4] recommends using a global error term to evaluate the metamodel. In this work, only the local error is dealt with in order to evaluate the method's effectiveness for CFD applications. This variance is estimated based on the estimation point's distance from the design points. For estimated values at a design point, the estimation variance is 0. It then increases along with the separation distance according to the autocorrelation function's fit to the data. If the estimation point is beyond the ranges of the autocorrelation function, $\vec{l}$, the Kriging variance is a constant maximum. Relating these local error measures to global error measures is left for a future study.

II.C. The RPC-K Method

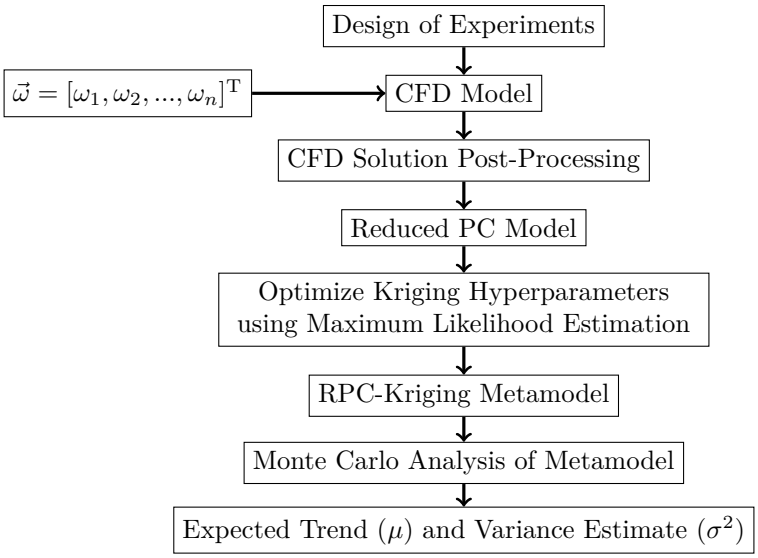
When combining the RPC method with universal Kriging, we follow Schöbi et al.[4] where Eq. (9) is rewritten as

$$\mathbf{U}(t,\vec{\omega})\approx\mathbf{U}^{(RPC-K)}(t,\vec{\omega})=\sum_{k=0}^{p}\phi_k(\vec{\omega})\mathbf{U}_k(t)+\boldsymbol{\sigma}^2Z(\vec{\omega},\vec{\zeta})\,. \tag{17}$$

This is the core of the RPC-K method. We have gained experience[1] to obtain the polynomial basis for the first term on the right hand side of Eq. (17), of which the procedure is detailed in SectionII.A. To emphasize, we replace the least angle regression method of Schöbi et al.[4] with the RPC method. The second term on the right hand side of Eq. (17) and the polynomial coefficients can be computed by the procedure as presented in Section II.B.

The RPC basis can be evaluated for each design point to construct the Vandermonde matrix, $\mathbf{F}$. With this data, the hyperparameters that maximize the likelihood function are then found by minimizing Eq. 15. With these hyperparameters found, the best linear unbiased predictor can be constructed using polynomial coefficients solved for using Eq. 10 and the Kriging variance using Eq. 13. This completes the RPC-K model for the system using the known design points. It can then be used to predict the values of solution variables within the input parameter domain. This process is depicted in the following flowchart.

First, the design of the experiments must be determined in order to give the best training data for the RPC-K model. This data is dependent on the input parameters selected, the design of experiments, and the CFD model used. The modeled data can be primitive or derived solution variables. Once a sensitivity study is performed and the number of simulations required to construct the RPC bases is known, we can create a RPC-K metamodel based on the simulation data. The RPC-K can be evaluated cheaply for many design points. This gives a simulated Monte Carlo analysis of the original design problem with both output trend and standard deviation estimates. It is worth mentioning that because the Kriging method gives estimates of the variance for each prediction, it can be used to quantify the uncertainty across the input parameter space. This information can then be used for adaptive refinement of the model by selecting design points which carry the most uncertainty. However, this adaptive method is left for a future study.



III. Verification

The methods were verified on a one-dimensional transient convection-diffusion-reaction problem. This same problem was used in our previous study[1] to verify the RPC method, and the problem statement is repeated here for convenience. The problem is represented in a unit domain of $0 \leq x \leq 1$ as

$$\frac{\partial \phi}{\partial t} + u\frac{\partial \phi}{\partial x} - k\frac{\partial^2 \phi}{\partial x^2} + c\phi = f\,, \tag{18}$$

where ϕ is the solution variable and u, k, and c are the convection speed, diffusivity, and source coefficient, respectively.They are all assumed to be one. The source term is modeled as $f = \alpha(x-1)e^{-t}$ with $\alpha = 2$. The exact solution of the model with the above given values takes the form

$$\phi_{\mathrm{exact}}(x,t) = x^2 e^{-t}\,.$$

The exact solution was used to evaluate the accuracy of the proposed RPC-K method in comparison to the RPC and UK methods. The methods were evaluated on a set of CFD solutions to the problem. The CFD model is based on a fourth-order finite-volume method.[12]

In the full polynomial chaos study we considered four random parameters, u, k, c, α; that is $n = 4$. The values prescribed above were used as the mean value for each parameter and each has a variation of 10%. The order of the PC expansion is $m = 3$. Our previous sensitivity study[1] found that the parameters c and α could be set to their mean values to reduce the number of input parameters to two for the RPC method.

The results of the RPC, UK, and RPC-K methods are shown in Fig. 1, which compares the predicted mean value from the three methods on the reduced model to the exact solution. Sixteen CFD simulations were used to build the reduced models. The exact solution is given in solid black, the mean predicted from the RPC method in solid green, the mean predicted from the UK method in solid blue, and the mean predicted from the RPC-K method is solid red.

Figure 1 shows that the RPC and UK methods give great accuracy for the solution with maximum errors no greater than 1.57%. The RPC-K method does not give as accurate of a solution as it continuously overpredicts the solution with an average error of 4.29% and a maximum error of 4.36%. Table 1 lists the average and maximum percent errors for the three methods. The error in the RPC-K method is believed to be caused by the low number of design points for the model which leads to difficulty in fitting the metamodel's hyperparameters. This can be adjusted with either more simulations or by adjusting the initial estimates for the maximum likelihood estimation. We will investigate this issue in a follow-up study. Nevertheless, the RPC-K method is more efficient and robust.

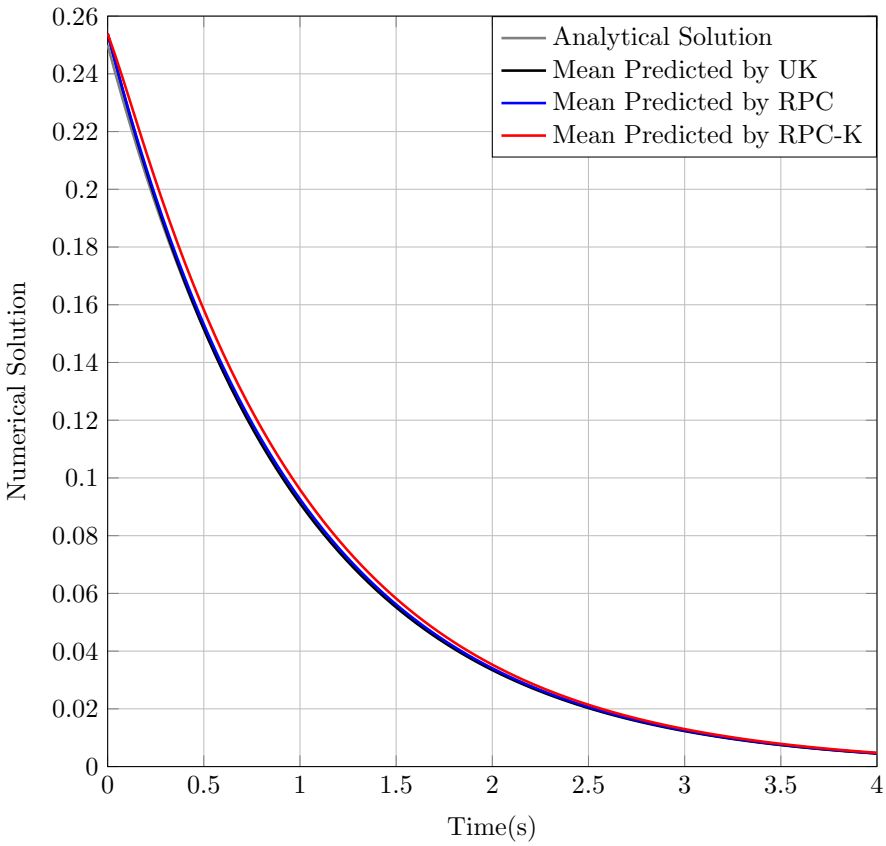


Figure 1: Predicted mean of the convection-diffusion-reaction problem of the three methods versus the analytical solution.

Table 1: The average and maximum errors for the convection-diffusion-reaction problem

	UK	RPC	RPCK
Avg error	1.04%	0.77%	4.29%
Max error	1.57%	1.57%	4.36%

IV. Results and Discussion

The RPC-K model is non-intrusive and can be used with any CFD software. For this research, the CFD++ solver from Metacomp Technologies[13] was used. It uses a second-order, upwind, finite-volume method. The governing equations used are the Reynolds-Averaged Navier Stokes (RANS) equations, and the gas is assumed calorically perfect. Steady-state RANS simulations were performed using the Gauss-Seidel relaxation method accelerated with an algebraic multigrid. The Spalart-Allmaras turbulence model was used in all simulations for its good performance and computational cost for aerospace engineering applications.

We used the same computational configuration for the engine nacelle inlet case as that in our previous work.[1] This allows a direct comparison between the results for the RPC and RPC-K methods. For convenience, the schematic of the engine nacelle geometry is shown in Fig. 2, and we reiterate some details of the case setup from our previous study. The rotor and stator blades were not modeled in this work to save computational cost and are shown only for reference. The geometry was meshed using an unstructured grid generated by Pointwise®'s T-Rex automated mesh generation tool. Data was collected at $X = 0$, denoted by the red line, where the rotor blade would intersect the hub (red dot). The boundary conditions can be referred to our previous study.[1]

The design of experiments includes seven parameters, $n = 7$, including the freestream Ma_∞, AOA, the corrected mass flow rate ($\dot{m}$), the ratio of the turbulent eddy viscosity to the molecular viscosity (μ_t/μ), the

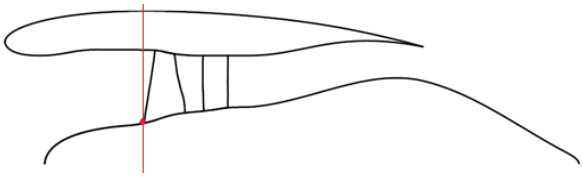


Figure 2: The location of the diameter measurement.

diameter of the nacelle inlet, the length of the nacelle, and the air density. We are interested in their influence on the flow separation along the nacelle lip and inlet distortion. The mean values for the 7 parameters are given in Table 2. A uniform distribution is assumed for each parameter with a variation of 15%. For cases with geometric changes, we scale the overall geometry based on the geometry of the baseline case. The baseline geometry has a maximum external diameter of nacelle, $D_{\text{ref}} = 0.68$ m, at $X = 0$ and a length, $L_{\text{ref}} = 0.93$ m, from the leading edge of the nacelle casing to the trailing edge. Note that, for the sensitivity analysis, the spatial location was normalized to [0, 1] to accommodate the cases for geometric variations.

Table 2: The mean values for the 7 parameters

Set	Mach	AOA (°)	$\dot{m}$ (kg/s)	μ_t/μ	D/D_{ref}	L/L_{ref}	Density (kg/m^3)
1	0.25	26	24	4	1	1	1.225
2	0.25	15	22	4	1	1	1.225

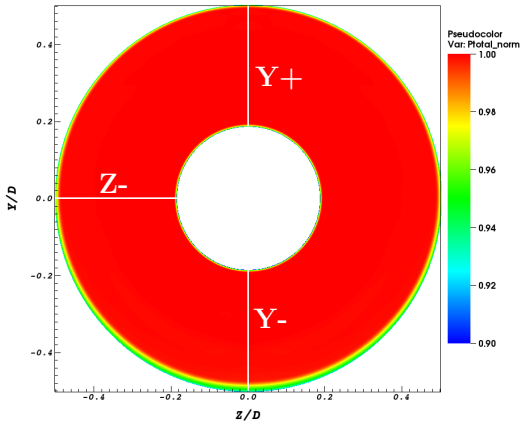


Figure 3: The total pressure distortion at $X = 0$ for the flow-through case at AOA=26^o.

Our previous studies[1,14] examined the important flowfield dynamics at $X = 0$ as shown by the red solid line in Fig. 2. For this particular plane, Fig. 3 represents a typical contour showing the distortion pattern (location, range, and magnitude). The distortion magnitude is measured using a normalized total pressure. The figure also shows three solid white lines, where the pressure and temperature profiles are used to compare solutions obtained by the RPC and RPC-K methods.

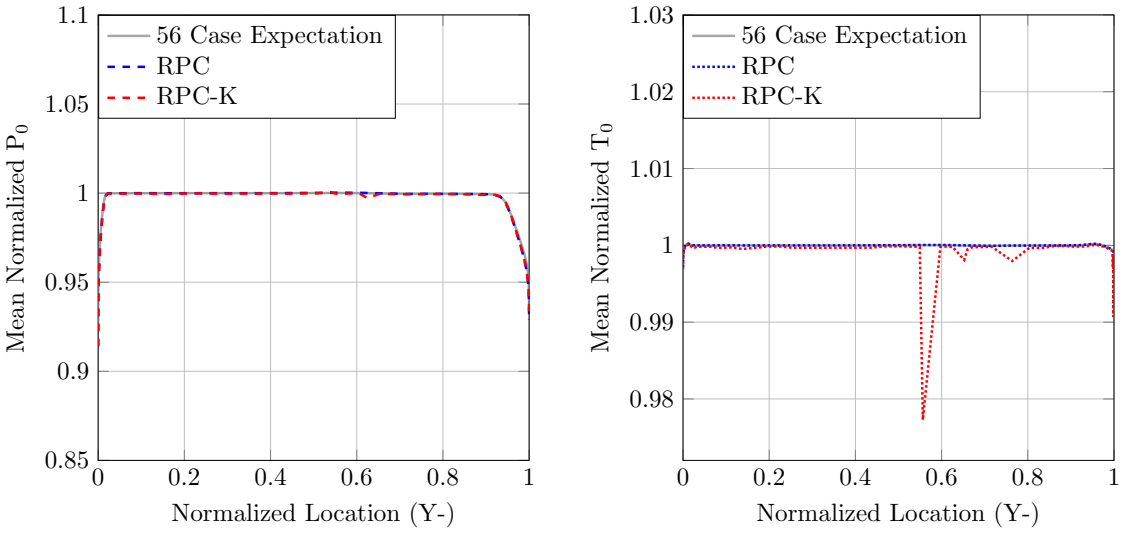
For the present study, we start with the RPC results from our previous work, where the input parameters were reduced from seven to five based on their impact. The molecular viscosity and air density were found to have little effect on the output parameters from a sensitivity study detailed in our previous work.[1] The methods are evaluated on this reduced parameter set as we are most interested in finding which method recovers the true solution most accurately from this sparser parameter set. Each method was trained using 56 CFD simulations as this is the minimum number needed to fit a PC expansion for five parameters using an expansion of order three. The RPC and RPC-K models were then evaluated along the three profiles at $X = 0$. These profiles are described by 134 spatial locations along the radial direction. The metamodels were evaluated at 10,000 random design configurations for each point on these profiles to find a predicted mean and standard deviation.

Two sets of CFD simulations were used. Data in set 1 contains the lower AOA cases with smooth flows

on all profiles. Data in set 2 contains the high AOA cases with separated flow on the Y- profile. Our previous study[1] found the separated flow region on the Y- profile for high AOA problems to have the greatest error when the number of parameters is reduced.

IV.A. Set 1

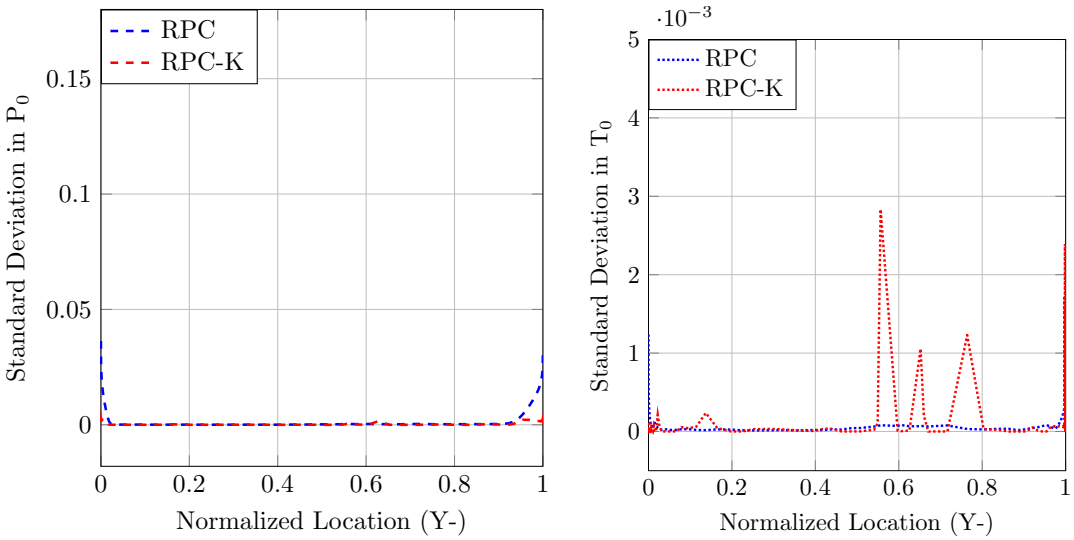
The total pressure, P_0, and total temperature, T_0, are plotted along each profile in non-dimensional form. The pressure and temperature values are non-dimensionalized versus the simulation's freestream P_0 and T_0 values. The locations along each profile are also non-dimensionalized to account for varying geometry with values of 0 corresponding to the hub wall of the engine and values of 1 at the nacelle inner wall. The color legend is solid light gray for the 56 case mean, blue for the RPC method, and red for the RPC-K method. The line styles for P_0 and T_0 are dashed and dotted, respectively.



(a) Mean normalized P_0 on Y-

(b) Mean normalized T_0 on Y-

Figure 4: Mean normalized profiles of P_0 and T_0 on Y- (set 2).



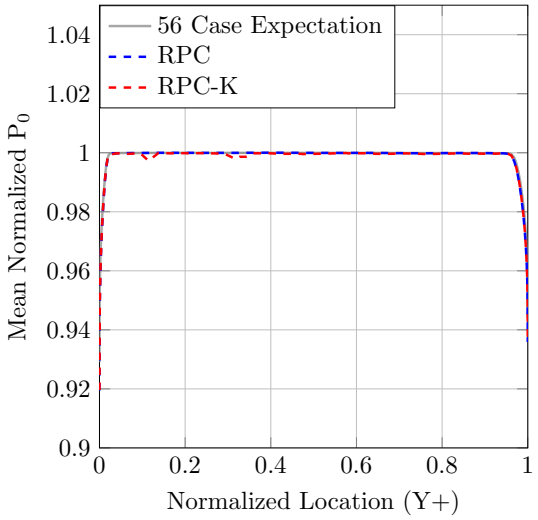
(a) Standard deviation in P_0 on Y-

(b) Standard deviation in T_0 on Y-

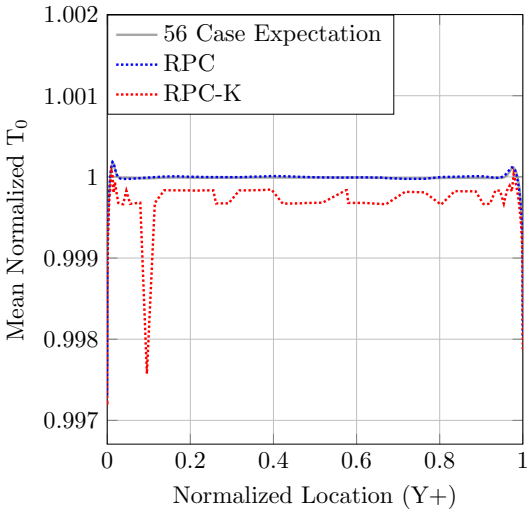
Figure 5: Standard deviations in P_0 and T_0 on Y- (set 2).

Figures 4a and 4b show the mean P_0 and T_0 profiles along Y- for the first set while Figs. 5a and 5b show their standard deviations. For P_0, the methods give good results. While for T_0, the RPC-K method does

not show smooth results with discontinuities near 0.57, 0.63, and 0.77.

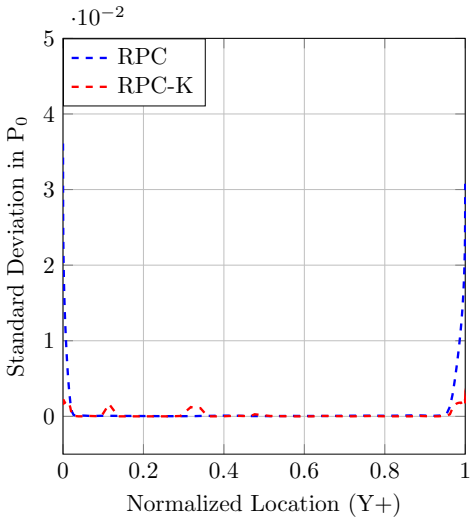


(a) Mean normalized P_0 on Y+

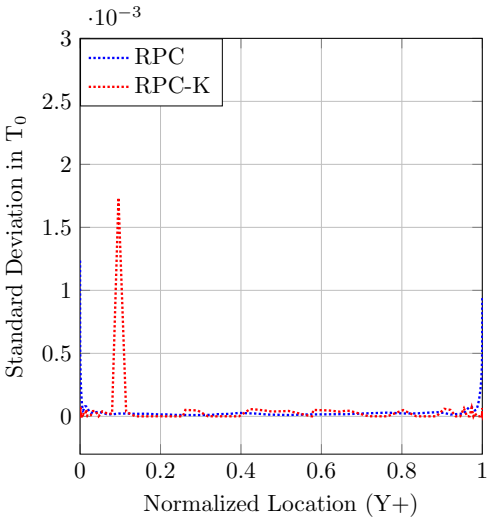


(b) Mean normalized T_0 on Y+

Figure 6: Mean normalized profiles of P_0 and T_0 on Y+ (set 2).



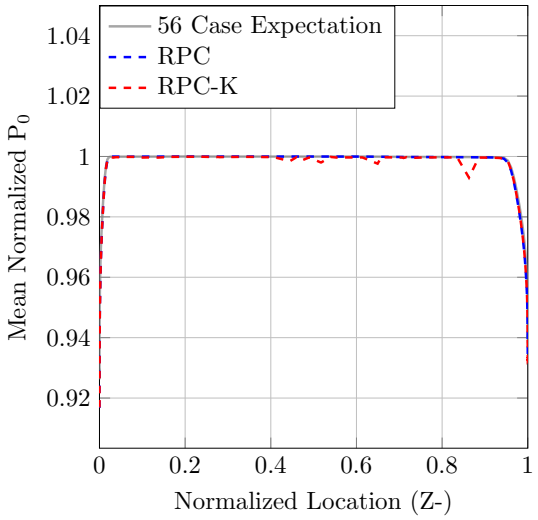
(a) Standard deviation in P_0 on Y+



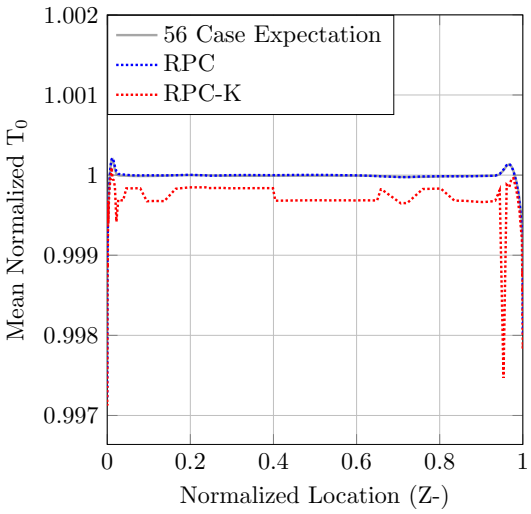
(b) Standard deviation in T_0 on Y+

Figure 7: Standard deviations in P_0 and T_0 on Y+ (set 2).

Figures 6a and 6b show the mean P_0 and T_0 profiles along Y+ while Figs. 7a and 7b show their standard deviations. Results for the Y+ profile are similar to those for the Y- profile. For T_0, the RPC-K method shows a discontinuity near 0.1 and also consistently underpredicts the solution, though by less than 0.05%.

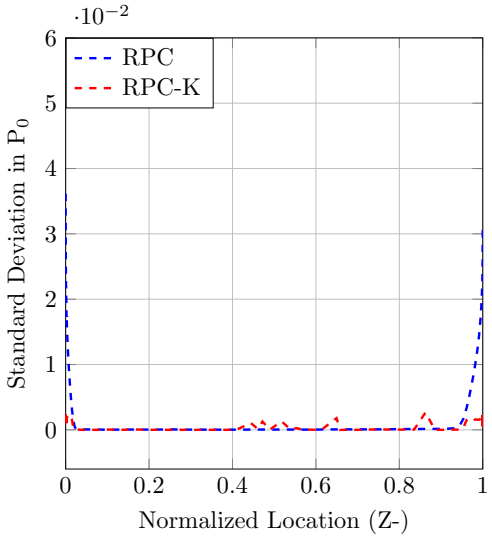


(a) Mean normalized P_0 on Z-

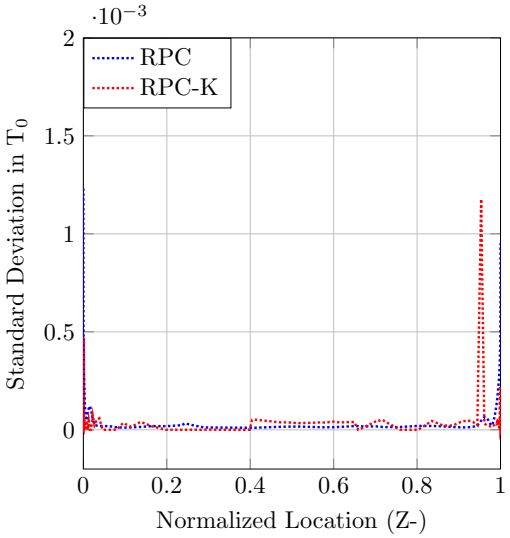


(b) Mean normalized T_0 on Z-

Figure 8: Mean normalized profiles of P_0 and T_0 on Z- (set 2).



(a) Standard deviation in P_0 on Z-



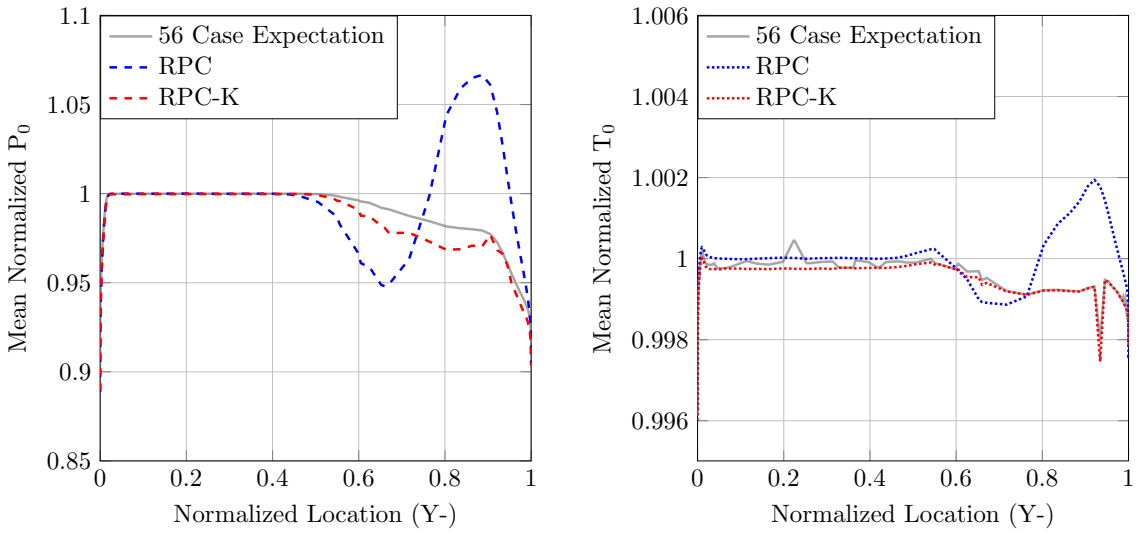
(b) Standard deviation in T_0 on Z-

Figure 9: Standard deviations in P_0 and T_0 on Z- (set 2).

Figures 8a and 8b show the mean P_0 and T_0 profiles along Z- while Figs. 9a and 9b show their standard deviations. For P_0, the methods give good results with the RPC-K method giving a small discontinuity near 0.85. The RPC-K method again shows a discontinuity for T_0 near 0.95 with a small underprediction bias.

IV.B. Set 2

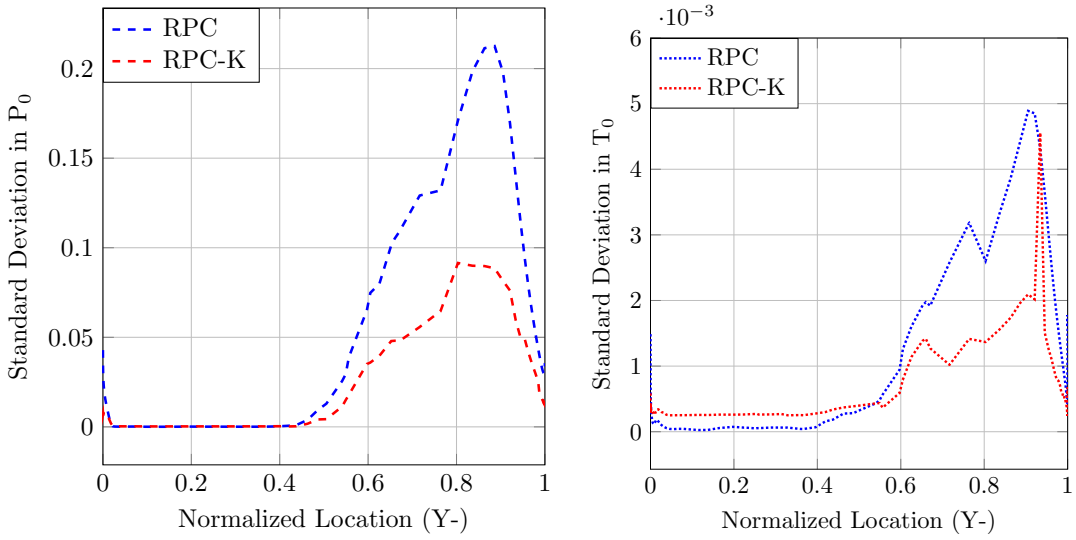
The total pressure, P_0, and total temperature, T_0, are plotted along each profile in non-dimensional form. The pressure and temperature values are again non-dimensionalized versus the simulation's freestream P_0 and T_0 values. The locations along each profile are also non-dimensionalized to account for varying geometry with values of 0 corresponding to the hub wall of the engine and values of 1 at the nacelle inner wall. The color legend is solid light gray for the 56 case mean, blue for the RPC method, and red for the RPC-K method. The line styles for P_0 and T_0 are dashed and dotted, respectively.



(a) Mean normalized P_0 on Y-

(b) Mean normalized T_0 on Y-

Figure 10: Mean normalized profiles of P_0 and T_0 on Y-.



(a) Standard deviation in P_0 on Y-

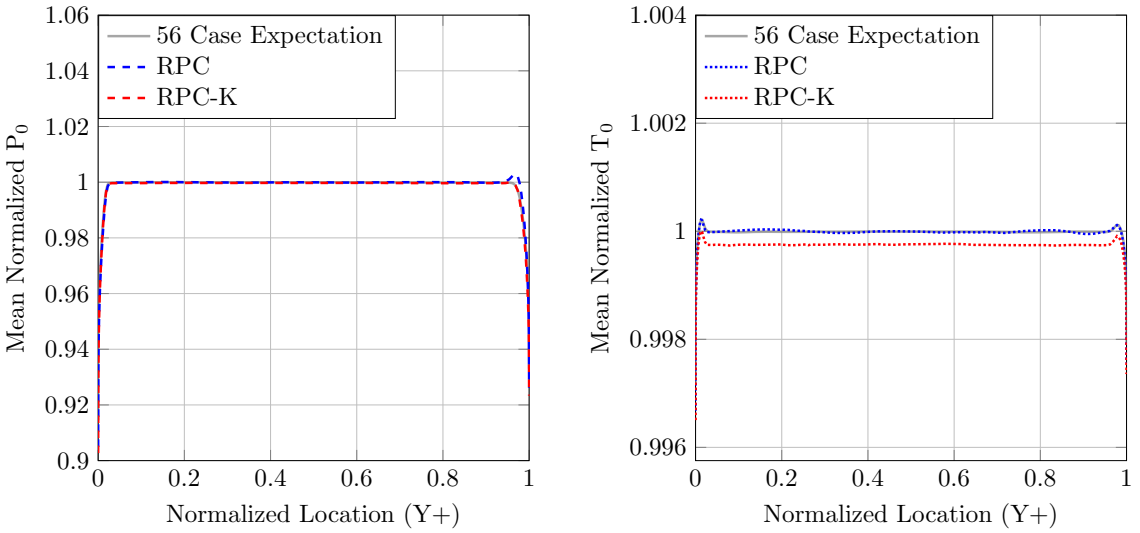
(b) Standard deviation in T_0 on Y-

Figure 11: Standard deviations in P_0 and T_0 on Y-.

Figures 10a and 10b show the mean P_0 and T_0 profiles along Y- for the second set while Figs. 11a and 11b show their standard deviations. It can be seen in Fig. 10a that the RPC method deviates from the trend greatly near the normalized locations 0.7 and 0.9. Notably, it gives a normalized total pressure values greater than one which is non-physical since there is no mechanism to increase the total pressure within the model. This deviation is more pronounced in P_0 than T_0. The RPC-K method gives similar results as the 56 case average with a greater loss in P_0 from normalized locations between 0.5 and 0.9. It also produces a less smooth solution. The RPC-K method gives better results than the RPC method as it eliminates the large fluctuations near 0.7 and 0.9 and does not produce non-physical values. Figure 10b shows that again the RPC method gives larger fluctuations at locations of 0.7 and 0.9 than the RPC-K method. The RPC-K method underpredicts the solution across the domain and also shows a discontinuity at a normalized location near 0.92. This discontinuity is visible in the case average data and persists even when the model is evaluated 100,000 times at the location.

Figure 11a shows the standard deviations given by the two models for P_0 along the Y- profile. The RPC method produces near zero standard deviation along the smooth portion of the flow and larger standard

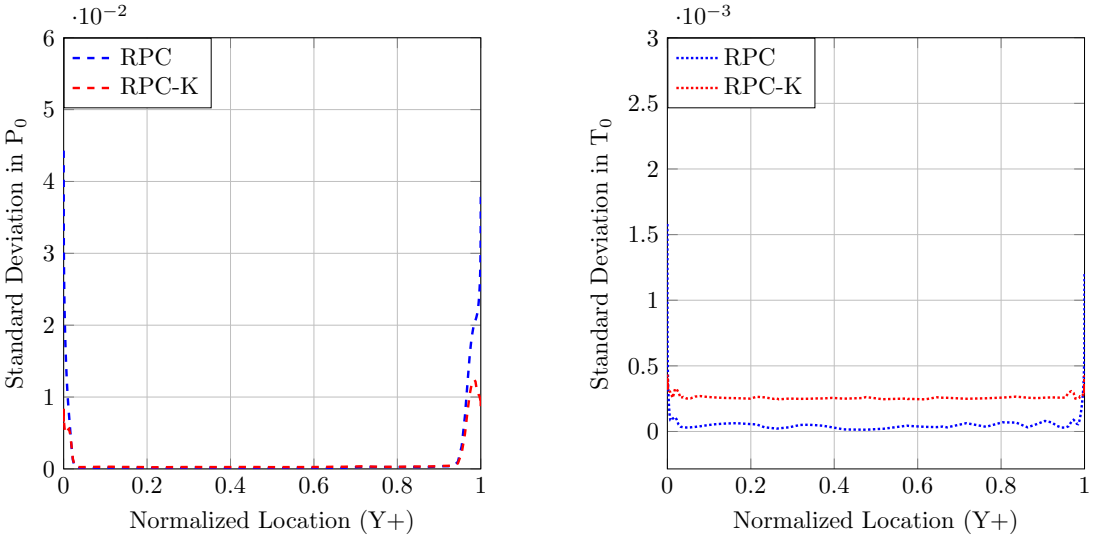
deviation in the separated region peaking near 0.2 at the normalized location 0.9. The RPC-K model predicts a standard deviation pattern the same as the RPC method, though it peaks near a value of 0.09. This is less than 50% of the value for the RPC method. For T_0, the standard deviation values are much lower. The RPC method gives near zero standard deviation for the smooth region and a value up to 5×10^{-3} for the separated flow. The RPC-K method gives non-zero standard deviation values for the smooth region as it underpredicts the solution. Again, it has lower standard deviations than the RPC method in the separated region, except for the discontinuity at 0.92. Though, this value is no higher than the standard deviation for the RPC method at this location.



(a) Mean normalized P_0 on Y+

(b) Mean normalized T_0 on Y+

Figure 12: Mean normalized profiles of P_0 and T_0 on Y+.



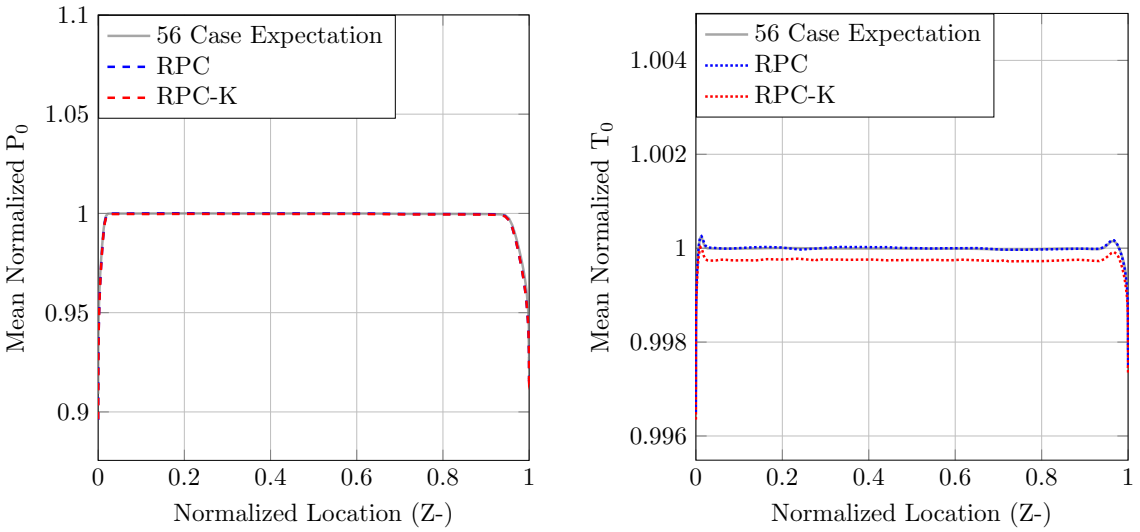
(a) Standard deviation in P_0 on Y+

(b) Standard deviation in T_0 on Y+

Figure 13: Standard deviations in P_0 and T_0 on Y+.

Figures 12a and 12b show the mean P_0 and T_0 profiles along Y+ while Figs. 13a and 13b show their standard deviations. Figures 12a and 12b show the improvement in fit of the methods as the solution becomes smooth. The RPC method is able to accurately fit the output data with only minor deviations near both walls. The RPC-K method fits the solution through the entire domain. For T_0, the RPC method accurately fits the data. Again, the RPC-K method underpredicts the solution across the domain. For the standard

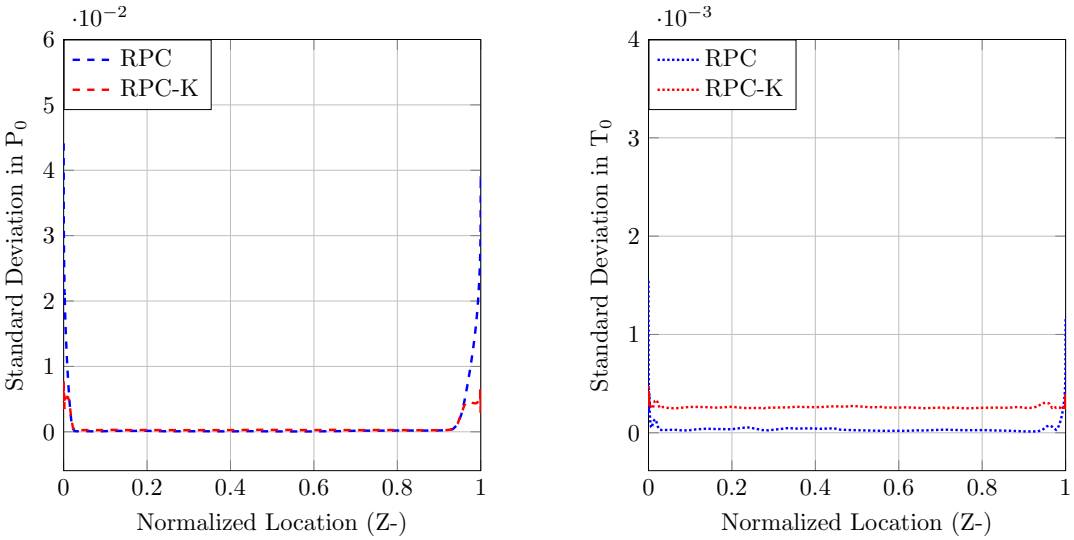
deviations for both P_0 and T_0, both methods give low values for the main flow with increasing deviation near the walls. The RPC-K method gives higher standard deviation values for T_0 than the RPC method because of its underprediction.



(a) Mean normalized P_0 on Z-

(b) Mean normalized T_0 on Z-

Figure 14: Mean normalized profiles of P_0 and T_0 on Z-.



(a) Standard deviation in P_0 on Z-

(b) Standard deviation in T_0 on Z-

Figure 15: Standard deviations in P_0 and T_0 on Z-.

Figures 14a and 14b show the mean P_0 and T_0 profiles along Z- while Figs. 15a and 15b show their standard deviations. The results for the Z- profile are similar to those for the Y+ profile except that they are slightly smoother. Again, the RPC-K method underpredicts T_0 across the domain.

V. Concluding Remarks and Future Work

The RPC-K method constructs a more accurate metamodel than the RPC method alone based on the test cases in this study. The Kriging component of the model is able to eliminate the large variances present in low-order PC methods. The UK method gives lower errors for the fitted profiles in our study, though it simply recovers the mean of the simulation data. This, along with its higher standard deviation values,

means the method is more likely to struggle to predict off-design values. Both the UK and RPC-K return these results without additional simulations required than the RPC method.

Future work will determine whether a sequential application of the RPC and UK methods or the combined RPC-K method can improve the accuracy and efficiency significantly. We plan to test various sampling methods on multiple types of engineering flows to evaluate their relative accuracy and stability in general. Additionally, we propose to develop an adaptive algorithm for both methods to determine its effectiveness at alleviating their sensitivity to the initial sample set. This method will also allow more accurate estimation of which CFD simulations will be of more importance to uncertainty quantification, further reducing the computational costs of design optimization studies.

References

[1]Gao, X., Wang, Y., Xie, N., Spotts, N., Roy, S., and Prasad, A., "Fast Uncertainty Quantification in Engine Nacelle Inlet Design Using a Reduced Dimensional Polynomial Chaos Approach," AIAA 2016-0298, 54th AIAA Aerospace Sciences Meeting, 2016.

[2]Webster, R. and Oliver, M. A., *Geostatistics for Environmental Scientists*, John Wiley and Sons Ltd., West Sussex, England, 2007.

[3]Kawai, S. and Shimoyama, K., "Kriging-model-based uncertainty quantification in computational fluid dynamics," *32nd AIAA Applied Aerodynamics Conference*, No. 2014-2737, AIAA, Atlanta, GA, June 2014.

[4]Schöbi, R., Sudret, B., and Wiart, J., "Polynomial-Chaos-Based Kriging," *International Journal for Uncertainty Quantification*, Vol. 5, No. 2, 2015, pp. 171–193.

[5]Rabitz, H. and Alis, O. F., "General formulation of high-dimensional model representations," *Journal of Math. Chem*, Vol. 50, No. 2-3, 1999, pp. 197–233.

[6]Xiu, D., "Fast numerical methods for stochastic computations: A review," *Commun. Comput. Phys.*, Vol. 5, No. 2-4, Feb. 2009, pp. 242–272.

[7]Eldred, M. S., "Recent advance in non-intrusive polynomial-chaos and stochastic collocation methods for uncertainty analysis and design," *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conf.*, May 2010, pp. 1–37, n Proc.

[8]Saltelli, A., Tarantola, S., and Campolongo, F., "Sensitivity analysis as an ingredient of modeling," *Statistical Science*, Vol. 15, No. 4, 2000, pp. 377–395.

[9]Sobol, M., "Sensitivity estimates for nonlinear mathematical models," *Mathematical Modeling and Computational Experiment*, Vol. 55, No. 1-3, Feb 2001, pp. 271–280.

[10]Dubourg, V., *Adaptive surrogate models for reliability analysis and reliability-based design optimization*, Ph.D. thesis, Blaise Pascal University, Clermont-Ferrand, France, 2011.

[11]Santer, T. J., Williams, B. J., and Notz, W. I., *The Design and Analysis of Computer Experiments*, Springer, 2003.

[12]Gao, X., Guzik, S. M. J., and Colella, P., "Fourth Order Boundary Treatment for Viscous Fluxes on Cartesian Grid Finite-Volume Methods," AIAA 2014-1277, 52nd AIAA Aerospace Sciences Meeting, 2014.

[13]"http://www.metacomptech.com," 2015.

[14]Spotts, N., *Unsteady Reynolds-averaged Navier-Stokes simulations of inlet flow distortion in the fan system of a gas-turbine engine*, Master's thesis, Colorado State University, 2015.