THESIS

DEVELOPMENT OF REDUCED POLYNOMIAL CHAOS-KRIGING METAMODEL FOR UNCERTAINTY QUANTIFICATION OF COMPUTATIONAL AERODYNAMICS

Submitted by

Justin Weinmeister

Department of Mechanical Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2018

Master's Committee:

Advisor: Xinfeng Gao

Stephen Guzik
Sourajeet Roy
Goldino Alves

ABSTRACT

DEVELOPMENT OF REDUCED POLYNOMIAL CHAOS-KRIGING METAMODEL FOR UNCERTAINTY QUANTIFICATION OF COMPUTATIONAL AERODYNAMICS

Computational fluid dynamics (CFD) simulations are a critical component of the design and development of aerodynamic bodies. However, as engineers attempt to capture more detailed physics, the computational cost of simulations increases. This limits the ability of engineers to use robust or multidisciplinary design methodologies for practical engineering applications because the computational model is too expensive to evaluate for uncertainty quantification studies and off-design performance analysis. Metamodels (surrogate models) are a closed-form mathematical solution fit to only a few simulation responses which can be used to remedy this situation by estimating off-design performance and stochastic responses of the CFD simulation for far less computational cost. The development of a reduced polynomial chaos-Kriging (RPC-K) metamodel is another step towards eliminating simulation gridlock by capturing the relevant physics of the problem in a cheap-to-evaluate metamodel using fewer CFD simulations. The RPC-K metamodel is superior to existing technologies because its model reduction methodology eliminates the design parameters which contribute little variance to the problem before fitting a high-fidelity metamodel to the remaining data. This metamodel can capture non-linear physics due to its inclusion of both the long-range trend information of a polynomial chaos expansion and local variations in the simulation data through Kriging. In this thesis, the RPC-K metamodel is developed, validated on a convection-diffusion-reaction problem, and applied to the NACA 4412 airfoil and aircraft engine nacelle problems. This research demonstrates the metamodel's effectiveness over existing polynomial chaos and Kriging metamodels for aerodynamics applications because of its ability to fit non-linear fluid flows with far fewer CFD simulations. This research will allow aerospace engineers to more effectively take advantage of detailed CFD simulations in the development of next-

generation aerodynamic bodies through the use of the RPC-K metamodel to save computational cost.

ACKNOWLEDGEMENTS

This work would not have been completed without the help of all those around me. First and foremost I must thank Dr. Xinfeng Gao for pushing me along in research until I was able to find this topic and progress on my own. The assistance in class, in meetings, and on papers is the driving force behind the quality of this work. I would also like to thank my committee; Dr. Dino Alves, Dr. Sourajeet Roy, and Dr. Stephen Guzik for their assistance throughout my career by helping me identify weak areas and fix them.

I would also like to thank the CFD lab for making work easier between lunch, coffee club, and all the other necessary breaks. I must also thank my roommates and the rest of the CSU cross country team for putting up with my work hours and work talk. Finally, I would like to thank my parents and siblings for their unwavering support of my passions and assistance in getting me to where I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

As the complexity of new engineering designs increases, the demand for statistical data of key design variables for uncertainty quantification (UQ) does as well. Uncertainty quantification is a quantitative study of a problem's uncertainties in order to understand and reduce them. This UQ data is especially important when using robust and reliability design methodologies [2]. This UQ data is often generated from propagating the known or estimated probability distributions of the design parameters through a model, known as a forward problem [3]. The forward problem model can be physical testing, empirical formulations, or computer simulations. Critically, generating a large amount of statistical data for complex problems always comes with a significant cost. To ameliorate this cost, a metamodel (surrogate model) can be used to approximate the model's response at off-design points [2, 4–10].

While this data can be generated in many ways, the present work focuses on using computer simulations to estimate design problem responses for aerodynamics problems. For the aerodynamics problems of interest in this thesis research, the computational tool of choice is computational fluid dynamics (CFD) solvers. These codes solve a form of the partial differential equations which govern fluid flow, the Navier-Stokes equations. These equations are famously difficult to solve analytically, so many versions and fidelities of solvers exist for solving fluid dynamics problems under a variety of conditions. Additionally, while computational power has been increasing, the complexity of the codes used for these simulations has also increased in order to capture more physics [5, 11–13]. Thus, simulations become computationally expensive. Metamodels can assist engineering design of these aerodynamics problems by fitting a computationally cheap mathematical model to estimate the responses of these CFD simulations. These estimates can be used to estimate the problem's uncertainty and its behavior at off-design points.

Generally, there are two ways to generate the UQ data for a computer experiment, intrusive and non-intrusive. Intrusive methods act by modifying the underlying governing equations so as to

"perturb" them with some distribution so that the response can be interpreted stochastically. Non-intrusive methods process the output data of a standard model which has been run at multiple input conditions dependent on the underlying probability distributions of the parameters. The stochastic data is then estimated from the set of deterministic responses. While intrusive methods can be quite efficient and reliable, they are rarely used so existing software does not need modification and the difficulty in re-formulating the governing equations [14]. In this thesis work, non-intrusive methods are employed.

For non-intrusive studies, the computer experiment can be broken down into the key steps of selecting a forward problem model, choosing a sampling strategy, running the simulations, and then processing the data. Since computer experiments have deterministic instead of stochastic responses, the stochastic response must be estimated [3]. The easiest method for this is Monte Carlo (MC) simulation which involves selecting a forward problem model and running it many times with random inputs chosen according to their parameter distributions. Then, the responses of the forward problem are considered stochastically [14–16]. This method, while easy to implement, is computationally expensive and has a slow convergence rate. Therefore, new tools are required to make more effective use of the data from each computer simulation.

The use of computer simulation for engineering design has been evaluated by several agencies and always found to be one of the critical components to maintaining leadership in engineering design, driving significant research. This is true at the international level [17], across the entire National Science Foundation (NSF) [18], and for the National Aeronautics and Space Administration's (NASA) goals for CFD [12]. In all of these reports, effective uncertainty quantification (UQ) is necessary for best leveraging simulation data for new design methodologies. This requires balancing the number of problems investigated and the amount of data generated for each.

While the data generated from UQ is used in robust and reliability-based designs [2], the metamodels frequently used in the UQ process are necessary in multidisciplinary design optimization (MDO) as well. These metamodels are necessary in MDO as the optimization methods in most architectures make many function calls to the separate physical models in a problem in order to

evaluate new design points [7, 19]. Without metamodels, the computational cost of full CFD simulations would prohibit any attempt at MDO, limiting new globally optimal designs. Fortunately, integration methods exist for metamodeling technologies [20]. While MDO structures are not the focus of this work, an optimization property must be considered when building metamodels. The "no free lunch" theorem of optimization states that no single optimization method is superior to all others in all cases [21]. In other words, it is unwise to pursue one optimization strategy above all others to make a single, master optimization method. This is true even when fitting metamodels to models treated as a black-box [22, 23]. However, the advantages gained from automating the design optimization process means that it is not always wise to consider all possible methods for a given problem.

Because of this, only a few UQ methods are considered for these sparse aerodynamic problems. For metamodels, they can be generally categorized as polynomial methods, Kriging, and neural networks at a high level. The high training costs of current neural networks prevents their use in sparse applications, though single layer applications which use radial basis functions have been used effectively before as a sort of cross between neural networks and Kriging [24]. Most polynomial methods used to be known as response surface methodologies, though the advent of generalized polynomial chaos (PC) theory has in general replaced other polynomial forms with the specialized orthogonal polynomials [4–6, 8, 10]. Kriging methods exist in several forms and predict responses using the problem's covariance data and mean [4–6, 10]. Polynomial chaos methods are able to predict long range trend information and are easy to compute whereas Kriging methods provide great local accuracy and are exact interpolators. Both methods are suitable for multi-fidelity and sensitivity analysis which are popular new features of metamodels [9]. Recently, Roland Schöbi and Bruno Sudret [25] introduced the combined method, polynomial chaos-Kriging (PC-K), which uses a PC expansion as the trend function in universal Kriging. This combined method promises the benefits of both metamodel strategies in an effort to make a more robust method for UQ and optimization work. The method can capture the long range trends of the data through the use of a PC expansion while simultaneously exploiting the covariance data of the problem to refine

local estimates at off-design points. It also recovers the true value of the deterministic solutions used to construct the metamodel. Additionally, the PC expansion can be quickly evaluated to get the statistical moments of the problem and perform a sensitivity analysis of the problem. This combination of characteristics should suit the PC-K method to a wide variety of UQ studies.

For CFD UQ studies, several aspects are important. First, the computer codes must be verified and validated before they are used as the forward model. This is an important step in CFD studies since numerical methods can struggle to predict certain non-linear physics of fluid flows [26]. Once the simulation is validated, a reliable UQ method, such as Monte Carlo, PC expansions, or Kriging, must be determined [14]. However, the recently developed PC-K metamodeling methods have had very few validation cases, focusing mainly on test functions and structural reliability problems [10, 25, 27–31]. Very recently, the method has been applied to two low-fidelity aerodynamics problems on an airfoil [32, 33] and more advanced, non-aerodynamic fluids simulations [34, 35]. In the present work, the PC-K method is combined with a sensitivity index-based model reduction strategy for PC expansions, reduced PC (RPC) from Prasad and Roy [8], to create a reduced PC-K (RPC-K) metamodel for UQ on sparse aerodynamics CFD problems.

In our work, the new RPC-K metamodeling method is analyzed on three separate problems to determine if it is superior to previous methods for sparsely sampled problems. First, the method is verified and validated on a convection-diffusion-reaction (CDR) partial differential equation which has an analytical and CFD solution. Through this, the metamodel is evaluated versus both an analytical and Monte Carlo solution. The problem is then tested on the National Advisory Committee for Aeronautics (NACA) 4412 airfoil. This airfoil is solved using a Reynolds-averaged Navier-Stokes (RANS) CFD simulation on a 2D grid based on validation data from NASA Langley Research Center's Turbulence Modeling Resource (TMR) [1]. This validation case is based off of experimental data originally collected by Donald Coles and Alan J. Wadcock [36,37] in 1979 using a flying hotwire. For this case, the low-Mach number, 2D RANS solution can be solved relatively quickly. The test case also has strong non-linear effects in the separation seen at the trailing edge at high angles-of-attack. The RPC-K metamodel is also evaluated on an aircraft engine nacelle case

which simulates takeoff under a strong crosswind for the design, possibly leading to separated airflow at the nacelle lip. This CFD simulation is based off of research conducted by Nathan Spotts [38] using RANS simulations to predict the airflow in aeroturbine engines.

The thesis is organized as follows. In chapter 2, a literature review is conducted on the relevant methodologies for this study to both show the state-of-the-art metamodeling strategies for UQ studies and place this thesis research within the discipline. Chapter 3 presents the methods used in the present study, including PC expansions, Kriging, and RPC-K, along with the sensitivity analysis and error estimation methods. Chapter 4 presents the validation of the RPC-K metamodel on the CDR problem. Chapter 5 describes the problem configurations and discusses the metamodeling results for the NACA 4412 airfoil and the aircraft engine nacelle. Finally, chapter 6 provides the concluding remarks and suggestions for future work.

Chapter 2

Literature Review

First, the existing literature is reviewed to cover the state-of-the-art PC and Kriging metamodeling technologies since they are the foundational elements in the new metamodeling method developed in this thesis. In section 2.1 the original development and current state of non-intrusive, regression-based polynomial chaos expansions are detailed. Then, in section 2.2 the development of Kriging methods is reviewed. Their current application in metamodeling for aerospace applications is also highlighted. The full depth of PC and Kriging methods can not be comprehensively reviewed in this chapter due to the large body of literature that exists for both these methods. Section 2.3 lays out the development of the RPC-K metamodel. This section includes the development of sensitivity analysis methods, least angle regression, experimental design, and error measures which are important parts of the newly developed RPC-K metamodel in this work. Finally, section 2.4 emphasizes the motivation of the development and application of the RPC-K metamodel.

2.1 Polynomial Chaos Expansions

Polynomial chaos expansions are a finite set of polynomial basis terms and corresponding coefficients from an infinite span which describe a random process. They were originally developed by Norbert Wiener [39] in 1938 as a homogeneous chaos which describes independent, normally distributed data using Hermite polynomials. As terms are added to the expansion, it converges in the L_2 sense for processes with finite second-order moments. This means that the expansion converges to the true process mean and variance. While it can be applied to any process, it converges optimally (exponentially) for normally distributed or Gaussian processes. This is because the weighting function for Hermite polynomials is the same as that for a Gaussian process. The expansion can then be applied to multivariate problems where the modeled distribution is a tensor product of independent variables.

The homogeneous chaos was then expanded by Xiu and Karniadakis [40] into generalized polynomial chaos or Wiener-Askey chaos for various distribution types. For a given distribution and its matching Askey orthogonal polynomial, a generalized PC expansion built using Askey polynomials converges optimally. Since then, generalized PC expansions have been employed both intrusively and non-intrusively to represent many stochastic processes [41, 42]. For non-intrusive methods, once the polynomial basis is chosen, the coefficients then need to be fit. This is commonly done with either projection or regression methods [43]. The present work considers the regression approach as it simplifies the sampling and data-fitting of the metamodel. This is because regression methods typically only require $2(p+1)$ simulations and a well-conditioned data matrix [44, 45]. Here $p+1$ is the number of terms in the expansion. Even when using regression-based PC expansions, several implementations exist and must be considered [42].

Over time, PC has been expanded for more general applications, but it has also been critiqued for its limitations. First, traditional expansions have been trimmed to create sparse PC expansions which include fewer terms [42]. Traditionally, PC expansions are truncated at either a total degree or maximum degree for each univariate component. Sparse expansions either use hyperbolic truncation, which removes higher-order interaction terms, or methods that keep only the significant terms in the expansions by measuring their impact on the solution response [42]. A common method used to choose these sparse terms is least angle regression (LARS) which selects the model parameters which are most correlated with the output [42,46–49]. This thesis work also reduces the size of PC expansions by using the reduced PC (RPC) method [8, 50]. The RPC method works by fitting the PC expansion to only the parameters that have the greatest impact on the outputs. Applications of the method have been demonstrated by Prasad and Roy [8, 50] for electrical engineering and by Gao et al. [51] for aerodynamics.

Polynomial chaos has also been the subject of some criticism. This includes the lack of proof for their convergence [52] and their shortfalls from a statistical perspective [53]. Specifically, Anthony O'Hagan [53] has expressed doubt for PC's ability to fully quantify the uncertainty of a random process compared to traditional methods, such as Monte Carlo. Despite this criticism, PC

expansions remain popular because they are computationally efficient, relatively easy to construct, and they allow for very easy calculation of the statistical moments of the metamodel [9].

2.2 Kriging

The development of Kriging is credited to Georges Matheron [54] who based his research on Danie Krige's research on gold grade distributions in South Africa [55]. Kriging, also known as Gaussian process regression, helped found the field of geostatistics and has since been incorporated into other statistical applications. Kriging uses distance-weighted averages of existing data to predict off-design points as an interpolator. Fundamentally, Kriging estimates the deviation of an estimation point from a mean value based on the deviation of nearby points. The deviation is estimated using the empirical data of the processes' covariance. Kriging comes in three versions: simple Kriging, which estimates deviations from an unknown mean; ordinary Kriging (OK), which estimates deviations from a known mean; and universal Kriging (UK), which estimates deviations from a trend function [55]. Additional variations exist to estimate values of averaged blocks, correlated variables, and other extensions which have seen limited applications in computational studies.

While the literature on Kriging is extensive, some elements that are critical for this thesis research are noted. In 1983, Olivier Dubrule [56] found a closed form solution to calculate the leave-one-out cross validation (LOOCV) error of a Kriging surrogate. Then in 1984, Armstrong [57] presented the fundamental difficulty in creating a UK metamodel. The metamodel must simultaneously fit a trend function and the underlying covariance of the data. They interact and modify each other. Because of this, the assumption of stationarity of the data is often violated as the estimates of the trend and covariance can be inaccurate. Regardless, UK is still popularly used today [55, 58]. To make this fitting process as easy as possible, various correlation functions are considered.

The correlation functions in Kriging are fit to the empirical data. (Sometimes, the covariance function, which is a scaled and inverted version of the autocorrelation function, is referred to

instead.) To visualize this fit, the data is often plotted on a semivariogram. An example semivariogram is shown in Fig. 2.1 which plots example experimental data as black circles with a bounded linear covariance function fit to the data. The x-axis is the lag distance of a single parameter and the y-axis is the semivariance, γ, which is one-half of the variance of the simulation data. The plotted semivariance values are calculated between any two data points. For example, γ_{12} would be one-half of the variance between the response of simulation one and two and plotted as a single point based on h_1, the lag distance between the points in parameter one's space. This variance is often called a covariance as it describes the variance of discrete data points. Generally, the semivariances (and thus the covariances) should increase in some given shape to a maximum value which is constant throughout the domain, called the sill. This maximum covariance is the process variance which is assumed constant throughout the domain through the stationarity assumption. Data with a trend violates this assumption because this covariance never has a true sill value [55]. The covariance functions fit to the data can be either bounded or unbounded. Bounded covari-

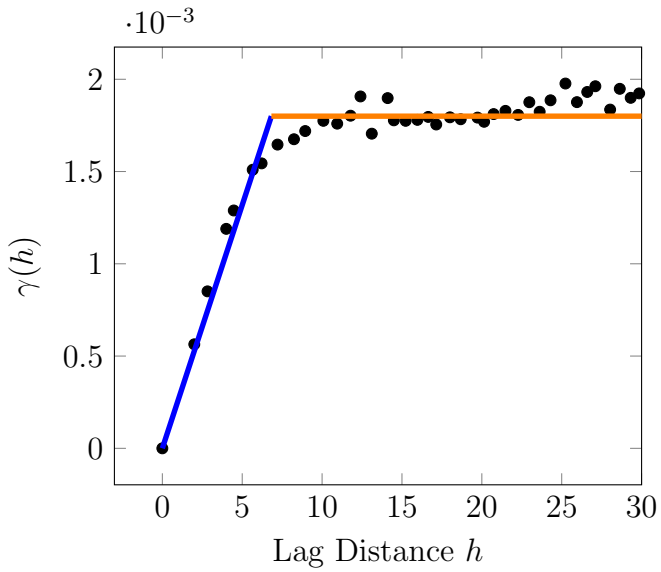


Figure 2.1: Example semivariogram with fitted bounded linear autocorrelation function.

ance functions have a finite lag distance at which the covariance reaches its sill while unbounded functions asymptotically approach it. In Fig. 2.1, a bounded linear covariance function is shown

which models the covariance as a linear increase from $h = 0$ to its maximum. From there, the covariance is constant at its sill value, $\gamma_{\max}$. The correlation function associated with this covariance function is modeled as $R = 1 - \gamma/\gamma_{\max}$. In geostatistics there is sometimes a resting variance value, called a nugget, which assumes a small variance even at lags of zero. For deterministic computer experiments, like the present study, this is not the case. In computer experiments, unbounded correlation functions are most common, including the Matérn, Gaussian-family, and cubic functions [55, 59–62]. Additionally, some recent research [63] has been conducted on non-stationary covariance functions. However, the methods have not yet been widely applied. Research on lifted-Brownian covariance functions has also been conducted, and it was found that they can describe the covariance of computer simulation data better than established covariance functions for geostatistical applications [64]. The process of fitting these covariance functions can be automated, though manual inspection is still conducted as the process is not rigorous [55].

When fitting these covariance functions, the key is to find the parameters of the function which minimize the error of the resulting Kriging model. These parameters of the covariance functions are known as the hyperparameters of Kriging and are generally smoothing terms and the lag ranges. The two commonly used methods are maximum likelihood estimation (MLE) and cross validation (CV) [3, 65, 66]. François Bachoc [67] explored both methods for estimating true covariance functions under uncertainty and misspecification. It was found that MLE is more likely to give the best estimate, although CV is more robust when the model is misspecified. Both methods are explored in the present research in order to deduce whether model misspecification was an issue for the aerodynamics problems investigated. These functions were optimized using either global or local optimization functions implemented in MATLAB®. Global methods are capable of finding global minima, but they require far more function calls than local methods. When using local methods, either the region of the global minima must be known or a local minima must suffice. The global methods explored in this thesis work include the functions *patternsearch*, *particleswarm*, *globalsearch*, *simulannealbnd*, and *ga* (genetic algorithm). Of these, simulated annealing and genetic algorithms have attracted significant interest recently. The local optimization functions used for

the present research was *fmincon* [68]. In general, MLE and local optimization work well for the problems in the present work as the functions are not complicated and global optimization does not yield a substantial benefit. Still, one of the main keys to efficiently finding the true covariance function which describes a process is to properly sample the problem regardless of the method used.

Multiple experimental design methodologies exist which are superior to the original Monte Carlo method used in computer experiments. Quasi-Monte Carlo methods are more efficient and are generally executed using quasi-random sequences like the Sobol sequence. Latin hypercube sampling (LHS) is another popular method which samples randomly from the problem distribution in areas of equal probability. Both of these methods are space-filling designs which are superior for computer experiments to the traditional boundary sampling methods used in physical testing [3]. Of these two methods, quasi-random sequences have the additional benefit of being more well suited to adaptive sampling methods as they can add any number of new points to an existing design while maintaining the experimental design structure. Latin hypercube sampling does not have this benefit as the sampling strategy no longer is consistent across regions of equal probability when an additional point is added [28, 69]. These experimental design methods conflict with the best methods for PC as PC expansions are typically best estimated using quadrature rules whereas Kriging needs as much covariance data as possible. Because of this, random sampling methods are superior to the more regular designs seen outside of Kriging [55].

For Kriging models created on sparsely sampled data, several methods have been explored to make more efficient use of the simulation data [3, 10]. Generally, these sparse simulation sets can be effectively analyzed for only a few key design parameters because they provide limited information. Thus, the model can fit to a subset of the originally considered parameters or a reduced set of transformed parameters. Some methods include partial least squares [70], active subspace [71–73], and sensitivity studies [74–77]. Partial least squares and active subspace methods are similar to principal component analysis where the key dimensions of a problem are identified and the problem is fit to these dimensions which may or may not be the actual dimensions (parameters) of the

problem. These key dimensions are computed as a linear combination of the existing dimensions. In sensitivity methods, the sensitivity of the response to each input parameter is estimated from the relative variance due to each parameter. From this information, the input parameters of the problem which contribute very little variance to the problem are culled from the design space as the loss of information (variance) from their removal is offset by the improved fitting of the important parameters [74–76]. These sensitivity methods are used in our study to better fit sparse data instead of dimension rotation methods as the dimensions for both Kriging and PC must be simultaneously considered.

2.3 Reduced Polynomial Chaos-Kriging

The formal process of introducing PC expansions as the trend function of a UK metamodel was first presented by Roland Schöbi and Bruno Sudret [27]. The details and results of using the method are covered in a series of papers which explored the best way to combine the two methods for various types of problems [10, 25, 27–31] (e.g., analytical functions, structural reliability, and geotechnical problems). Generally, PC-K methods can be solved in the same way as any UK metamodel except with the PC expansion being fit using linear regression. The method usually diverges from UK when selecting one of n PC expansions from the LARS algorithm for the trend function of the metamodel. Schöbi et al. [28] describes two general methods for selecting the best metamodel from the n choices, sequential and optimal PC-K. In sequential PC-K, the best PC expansion is selected from a candidate set and is used in the PC-K metamodel. In optimal PC-K, a PC-K metamodel is constructed for each expansion and the best PC-K model selected. The optimal method can give better results, though the computational cost is greater [28]. This is mainly due to the cost of fitting the Kriging hyperparameters for each model. These optimal PC-K metamodels can show superior performance to the sequential method because correlation information is used in the least squares regression for estimating the polynomial coefficients [28].

The major benefit that the PC-K metamodels provide over either separate metamodel is the inclusion of both long-range trend information and short-range covariance data. Also, because

Kriging methods are interpolators, the predicted response values of design points match the responses used to construct the metamodel. This is a desirable trait for computer experiments since they are deterministic. This also means that off-design responses at points near simulations that have already been run can be more accurately predicted. Further, the analysis of the underlying variance structure of the data can provide useful information for analysis [3, 10, 55]. Furthermore, the PC component of the combined method provides easy-to-compute trend information from a computationally cheap metamodel which is easy to evaluate for statistical information. Polynomial chaos expansions have the superior characteristic of known convergence for various distributions of uncertainty and a large body of literature on their construction over other polynomial trend functions [10, 40, 44]. The key to exploiting the benefits of both is to ensure that the PC expansion accurately represents the data to the available fidelity, and the Kriging assumptions are not violated to the point that the quality of the metamodel is degraded.

The RPC-K metamodel was first presented by Weinmeister et al. [78] in 2017 based on the deficiencies of RPC for sparsely sampled, non-linear problems, as found in the work of Gao et al. [51]. The RPC-K method was developed to combat the over-fitting issues found in the RPC method by incorporating Kriging information. This initial work found the method showed some improvement, but the over-fitting issue still persisted even with further modifications [79]. This thesis includes the methods developed to counter-act this over-fitting, but first some of the important elements in the RPC-K method must be reviewed.

2.3.1 Sensitivity Analysis

The cornerstone of the RPC-K method is the sensitivity analysis used to reduce the problem to its high-variance dimensions. For the RPC-K metamodel, Sobol indices are used. In 1993 Ilya Sobol [74] introduced the concept of being able to decompose an integrand into separate summable dimensions. This is done by expressing the model in a high-dimensional model representation which is a sum of the mean of a model, the main effects, and each level of the interacting effects [80]. In 2001 Sobol [76] expanded this idea into the creation of global sensitivity indices

which can be applied to non-linear equations. These indices give a simple compact form to estimate the relative impact of each of the dimensions of a problem that can be simply computed from Monte Carlo and quasi-Monte Carlo methods. The method has since been expanded so that the sensitivity indices can be estimated from a sparse sampling of a problem [81]. In 2000, Saltelli [75] described many of the ways that sensitivity indices can be used in the model-building and evaluation processes. Importantly, the key parameters which have the greatest global impact on the model can be identified. In this thesis work, the indices are calculated from a PC metamodel [8, 82]. The sensitivity indices are evaluated based on an analysis of variance (ANOVA) concept which says that the variance of the problem can be decomposed into the components of each parameter. For total effects, every term in the PC expansion which is a product of a particular parameter is included in the computation of the variance for that parameter. For main effects, only the univariate terms for a parameter are summed. Using the concept of sparsity of effects, which states that low-order terms carry the most impact on an output, the main effects can be further simplified to the first-order effects only. These three measures provide decreasing levels of accuracy on the estimation of the true sensitivity indices of the problem.

Sensitivity indices are the backbone of the model reduction strategies employed in this thesis research based on work by Prasad and Roy [8,50]. This method works by using either the first-order or main effects estimations of the sensitivity indices to identify the un-important parameters of a problem before many simulations are performed. Using these estimations of the relative impact of the dimensions of the problem, the un-important dimensions are culled with minimal loss in accuracy for the resulting metamodel. After this, fewer simulations are required to fit the resulting metamodel for the problem.

The development of sensitivity indices has continued, including the development of new and efficient estimators for sparse problems [83]. Many of the more advanced features of sensitivities are not considered in the present work as the aerodynamics problems have fairly few dimensions for computer experiments and parameters which are clearly the most impactful. For a comprehen-

sive review of the state-of-the-art, readers can consult the work by Iooss [77]. Once the problem has been reduced using a sensitivity analysis, the sparse expansion must be chosen.

2.3.2 Least Angle Regression

The selection of the most appropriate metamodel for a problem is difficult and can be a chief issue in the automation of the metamodeling process. While Kriging models are generally assumed to have one form once the autocorrelation function is chosen, a PC expansion can have many forms depending on the basis terms chosen. In order to chose the most impactful dimensions, least angle regression (LARS) is employed to select the basis terms in order of decreasing correlation with the output response [46]. The algorithm is related to two forward selection methods, least absolute shrinkage and selection operator (Lasso) and forward stagewise. However, it is more computationally efficient than both as the number of steps in the LARS algorithm is exactly equal to the number of covariates considered in the problem. The method works by identifying the covariate (in this case the basis term) most correlated with the response and progressing along the vector describing that term until another covariate is equally correlated. The progression towards the response is measured by the increase in the coefficient for the selected covariate. At the point of equal correlation, the direction of movement becomes equiangular between the previous covariate, the new covariate, and the response. The process is then repeated until all covariates are selected. This process is more efficient than the aforementioned forward selection methods because it exploits the information of the problem to know exactly how far in each direction the coefficients need to be increased. For more advanced discussions on this original work, readers can refer to the discussion which follows the original paper [46]. A review of the method as it compares to the other L_1 regression methods is covered by Hesterberg [48].

The LARS algorithm has since been modified to include the considered covariates in a hierarchical method, so the terms show heredity [47, 84]. For a PC expansion, this means that after the mean term, only the first-order basis terms are considered. The higher-order terms can not be added to the regression problem until their lower-order terms are chosen. Additionally, interaction

terms are not chosen until the main effect terms which compose it are chosen. Multiple versions of hierarchical or heredity methods can be employed, and they can generally be classified as strong or weak. A strong heredity adheres to the above methodology while weak heredity only requires that one of the covariates of an interaction term is included before it is added to the problem [47]. These concepts are modified in this thesis work to create a new hierarchical method specific to these sparsely sampled problems. Primarily, the LARS methodology adopted for choosing the sparse PC expansion comes from the work of Blatman [49] which was adapted by Schöbi in the development of PC-K [28]. This hierarchical method works well in concert with the reduction of the problem to create a very sparse polynomial expansion. Next, the forward problem must be sampled in the most efficient manner to fit this sparse expansion.

2.3.3 Experimental Design Methods

An effective experimental design (sampling strategy) allows for the construction of a good metamodel by reducing the number of simulations which need to be run across the entire input parameter space. A good design chooses the samples in a way such that they are representative of the problem, allow for accurate fitting of the model, and the computation of classic tools, such as sensitivity indices. Where computer experiments differ from physical experiments is in the need for space-filling designs [3, 45, 85]. These designs are necessary in order to generate the UQ data for deterministic codes. For these reasons, Monte Carlo (MC) methods are still superior in terms of absolute convergence than any other method. It is only the large computational cost of MC methods which prevents their use in these applications. Some of the alternatives include: quasi-Monte Carlo or quasi-random sampling, which uses a quasi-random sequence such as a Sobol sequence; Latin hypercube sampling (LHS), which divides the domain into areas of equal probability and samples randomly in each area; K-means clustering, which uses MC sampling to establish regions of equal probability randomly in the domain; orthogonal arrays, which randomly sample at grid points within the problem domain; and uniform sampling, which creates a full factorial grid of the problem [3, 4]. Of these methods, quasi-random sampling and LHS designs are the most common

with Sobol sequences and minimax LHS the most popular versions for each, respectively. These methods have been further refined to provide optimal versions for metamodel construction [86,87]. These improved designs have also been expanded for sparse, adaptive problems.

Significant previous research has gone into developing sampling strategies that work well with adaptive schemes because these adaptive schemes often provide the lowest computational demand on a problem. Because of this, the present research investigates these methods to ensure their compatibility with the RPC-K metamodel. The need for new adaptive strategies has been driven by optimization algorithms and a need for criteria which take into account both the uncertainty in the problem domain and the location of design points already evaluated [88]. These adaptive algorithms also consider the type of metamodel being used for the problem and slowly converge to an optimal design. However, recent work has tried to tie adaptive sampling strategies into the model selection process itself. This way, if a model is misspecified, the algorithm can correct for it [89]. Despite this, fixed experimental designs and metamodels still dominate the field. For the present work, it was important to identify methods which work well with sparse sampling for both PC expansions and Kriging [90]. All of these methods have been developed to create an effective metamodel; however, one last key to the metamodeling processes is devising an estimate of the accuracy of the model.

2.3.4 Error Measures

Calculating the error associated with a metamodel, especially sparsely-constructed ones, is a chief difficulty in the metamodeling process since generally no "truth" exists. There are several error measures to consider. The most important error measure is the generalization error. This is the error of the model when it is applied to predict a response not within the design data used to construct the model. There is also a testing error which is a form of generalization error that is computed from a test set of data if one is set aside. Another common error term is the training error. This is the error that exists between the model and the data used to construct it. Ideally, one has a data set large enough that it can be divided into a training, validation, and test set. The

training set is used to calculate the model and the training error. The validation error is used to asses the model selection process, and the test set is used to estimate the generalization error. While both testing and training measures are monitored, the model chosen is the one which minimizes the generalization error [45]. This is the correct metamodel based on the idea of the bias-variance decomposition for mean squared errors.

For a mean squared error, the error can be broken down into a bias, variance, and irreducible error. The bias error is squared and represents the error from under-fitting a model while the variance is a measure of over-fitting. There is also always an irreducible error which occurs as the stochastic process varies around its mean value. An under-fitted model does not accurately predict off-design points because it does not account for the effects of all the parameters accurately. As more terms are added to the model, the model can account for these deviations. However, as terms are added, the model becomes more sensitive to the training data and over-fits the data. This occurs as the model attempts to perfectly predict every design point, thus leading to large oscillations between data points. Depending on the size of the data set and model, this over-fitting can occur quickly or slowly. In large problems, several measures exist to limit model complexity earlier than normally predicted by the mean squared error alone. The most common are the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) which both penalize the mean squared error based on the number of terms. In sparse problems, these additional criterion are generally not necessary as the mean squared error increases rapidly [45].

For sparse problems, several error measures can be used. However, all of the data must generally be used to construct the metamodel as not enough data exists to set aside for error prediction. Typically, to estimate the generalization error, cross validation is used. For small data sets, leave-one-out cross validation (LOOCV) is the preferred implementation [91]. This method works by leaving one data point out of the training set and calculating the prediction error for it. The process is then repeated for every training point and averaged. This method is well established and works well for most metamodels, though research continues on alternative methods to determine

over-fitting for sparsely-constructed metamodels [92]. These error measures help close the RPC-K metamodeling processes by giving feedback on the quality of the model constructed.

2.4 Motivation

State-of-the-art metamodels give engineers the ability to process statistical data for computer experiments more effectively than using brute force Monte Carlo methods, though existing methods are still not mature enough to meet demands. Specifically in aerodynamic CFD simulations, the computational cost of uncertainty quantification studies can still be cost prohibitive. Because of the non-linear physics being modeled, high-fidelity, high-cost simulations are used. As such, only a certain number of simulations can be performed, giving a design space which has only been sparsely sampled. Traditional PC expansions can be used to fit a metamodel to data for very little computational cost, and they allow easy computation of the problem statistics and sensitivity indices. However, they require larger data sets to fit the response surface accurately to the data. Meanwhile, Kriging methods fit this data very well, but they can have large uncertainties in regions with no sampling because they assume no underlying trend of the data. Combining a PC expansion with UK solves this problem because Kriging is used to help the PC expansion give more accurate estimates of the data when it is not well fit. Additionally, the model reduction makes the metamodel's estimates more accurate because it only has to fit the most important parameters of the model. The RPC-K metamodel in this thesis work provides engineers and CFD practitioners the ability to use the high-fidelity CFD simulations in robust and multi-disciplinary design methodologies through its efficient generation of UQ data. Ultimately, new aerodynamic designs will be more reliable and efficient with incorporating the UQ data generated from the RPC-K metamodel.

Chapter 3

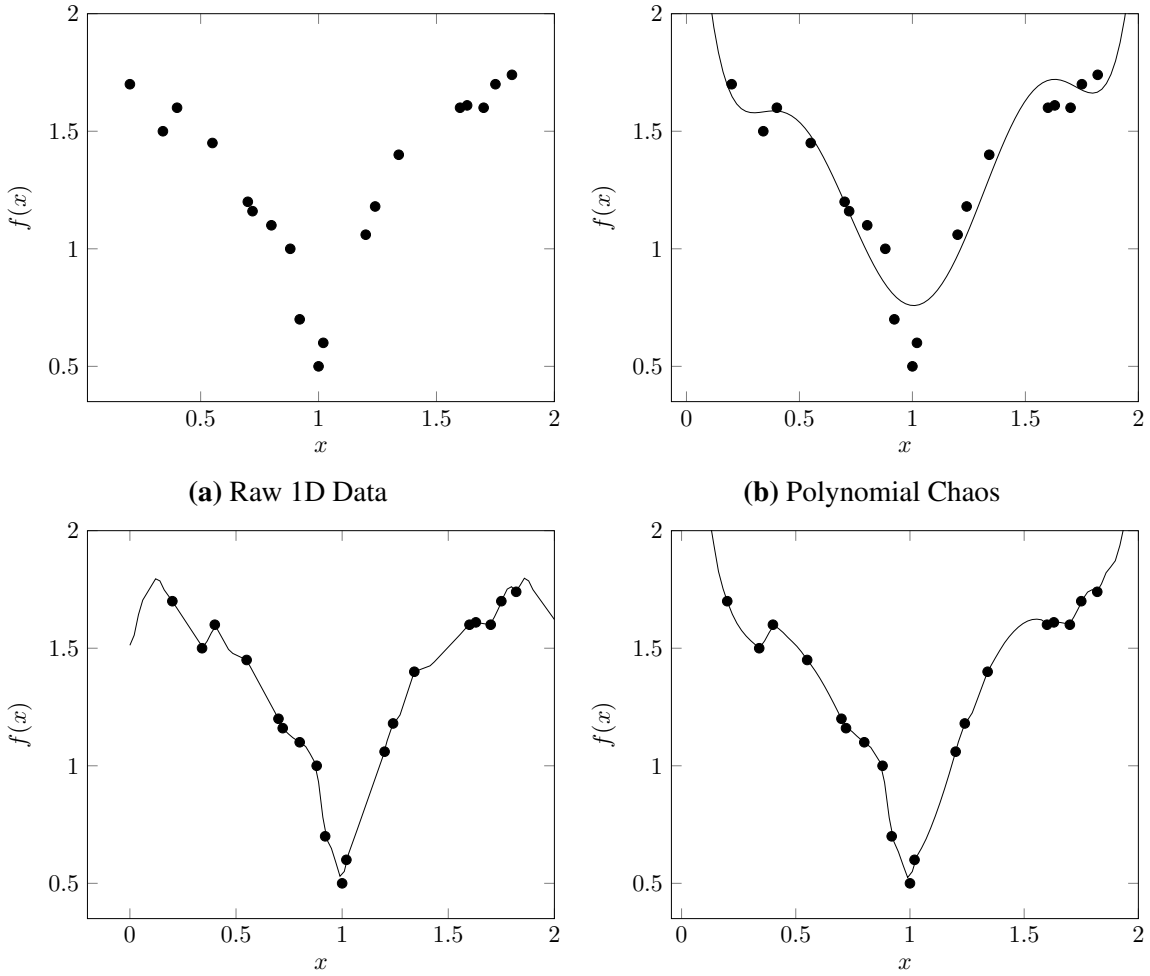
Reduced Polynomial Chaos-Kriging Methods

In this chapter the mathematical formulation of the RPC-K metamodel is presented. Sections 3.1 and 3.2 present the mathematical details of the PC and Kriging methods, respectively, since both are the major components of the new RPC-K metamodel. Then, in section 3.3 the two methods are combined to present the RPC-K formulation. Also discussed in this section are the sensitivity analysis procedure, least angle regression, and error measures. Finally, section 3.4 describes some additional characteristics of the new RPC-K metamodel which must be considered when it is used on reduced and sparsely-sampled problems. It also discusses the full methodology of a RPC-K workflow for its application by an engineer.

First, the general behavior of the metamodels is illustrated in Fig. 3.1. The metamodels are fit to example one-dimensional data with an input parameter x and response $f(x)$ shown in Fig. 3.1a. Then, an example 6th degree PC expansion is fit to the data in Fig. 3.1b which shows how PC metamodels accurately capture long-range trend information, but they do not necessarily interpolate through the known data points. Figure 3.1c shows the fit of an ordinary Kriging model on the data. Kriging exactly interpolates through the known data points, but it produces rough results and carries no long-range trend information. However, if the trend for a Kriging metamodel is approximated using PC, the result is improved as shown in Fig. 3.1d. This metamodel exactly interpolates through all of the known data points, smooths the solution, and retains long-range trend information. Therefore, a combined PC-K metamodel is desirable. The details on how these models are constructed are presented next.

3.1 Polynomial Chaos

The original Wiener PC expansions selected a set of Hermite polynomials from a full homogeneous chaos to describe a Gaussian process [39]. This has since been expanded to Wiener-Askey chaos by Xiu and Karniadakis [40] to include multiple distribution types and orthogonal polyno-



(a) Raw 1D Data

(b) Polynomial Chaos

(c) Ordinary Kriging

(d) Polynomial Chaos-Kriging

Figure 3.1: Examples of PC, OK, and RPC-K metamodels on 1D data.

mials in what is now known as generalized PC. A PC model can be expressed as

$$\mathrm{U}(\boldsymbol{\omega}) = \sum_{k=0}^{p} \beta_k \phi_k(\boldsymbol{\omega}), \tag{3.1}$$

where $\mathrm{U}(\boldsymbol{\omega})$ is the modeled response variable based on the standard random parameters $\boldsymbol{\omega}$ and there are a total of $p+1$ terms defined in Eq. (3.3). The vector of polynomial basis terms (ϕ_k) is $\boldsymbol{\phi}$ and the corresponding coefficients (β_k) vector is $\boldsymbol{\beta}$. There are a total of $p+1$ terms in the expansion and k is the multi-index for the terms. The basis consists of multivariate orthogonal polynomials given by

$$\phi_k(\boldsymbol{\omega}) = \prod_{i=1}^{n} \phi_{k_i}^{j}(\omega_i), \tag{3.2}$$

where $\phi_{k_i}^{j}(\omega_i)$ is the univariate polynomial of the i-th variable of $\boldsymbol{\omega}$ possessing a degree j based on the specific multivariate term. There are many ways to truncate the expansion, though the most common way is to limit the multivariate terms to a total degree. With this truncation, the total number of terms in Eq. (3.1) can be found from

$$p + 1 = \frac{(n+m)!}{n!m!}, \tag{3.3}$$

where n is the number of parameters i.e. the size of vector $\boldsymbol{\omega}$ and m is the total degree. For example, the multivariate terms for a two-parameter problem limited to the second degree are given in Table 3.1. A result of Eq. (3.3) is that the number of basis terms scales factorially with

Table 3.1: Multivariate polynomial terms for example expansion.

Term #	ω_1	ω_2
1	ω_1^0	ω_2^0
2	ω_1^1	ω_2^0
3	ω_1^2	ω_2^0
4	ω_1^0	ω_2^1
5	ω_1^0	ω_2^2
6	ω_1^1	ω_2^1

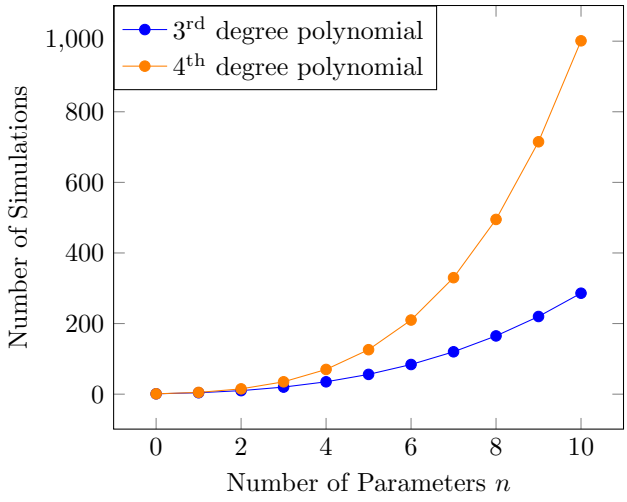


Figure 3.2: Effects of curse of dimensionality for PC.

the number of input parameters and the total degree of expansion. This factorial scaling is worse than even exponential scaling, and it is known as the curse of dimensionality. One result of the curse of dimensionality is that the number of basis terms required to fit the data quickly becomes computationally intractable. For example, Fig. 3.2 shows that while ten parameters can be fit with around 300 basis terms when truncated using 3rd order polynomials, it takes around 1000 when truncating at 4th order terms.

The orthogonal polynomials are selected from several classes depending on the probability density function they are representing. For Gaussian distributions, the probabilist's form of the Hermite polynomials are used while for uniform distributions, Legendre polynomials are used. A few of the common continuous distribution forms are reproduced in Table 3.2 [40]. These poly-

Table 3.2: Some orthogonal polynomial types and their associated distribution.

Distribution	Polynomial	Support
Gaussian	Hermite	$(-\infty, \infty)$
gamma	Laguerre	$[0, \infty)$
beta	Jacobi	$[a, b]$
uniform	Legendre	$[a, b]$

nomials belong to the Askey scheme where each type has a weight function equal to the weight

function of its corresponding probability distribution and satisfies a three-term recurrence relationship. For these distributions, the expansions created from the associated polynomials converge in the L_2 sense if the random parameters are independent and have finite statistical moments. An orthogonal polynomial, $\phi(\omega)$, satisfies the relation

$$\int_S \phi_n(\omega)\phi_m(\omega)w(\omega)d\omega = h_n^2\delta_{nm}, \tag{3.4}$$

where $n, m \in \mathcal{N}$ and $\mathcal{N} = [0, 1, 2, ..., N]$, or the exact degrees of the polynomials, and N is the number of data points. Here S is the support of the measure which has a weight $w(\omega)$ and δ_{nm} is the Kronecker delta. The polynomials are orthonormal when $h_n = 1$. This relation is the inner product in the Hilbert space and the weighting function is

$$w(\boldsymbol{\omega}) = \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}}, \tag{3.5}$$

when the random process is Gaussian [40, 44].

For non-intrusive expansions, the PC expansion is completed by estimating the value of the coefficients. This can be accomplished through either projection or regression methods. Regression methods are often used for their simplicity and convergence [42]. Regression minimizes the mean squared error of the coefficient estimates and gives the solution form

$$\hat{\beta} = (\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{U}, \tag{3.6}$$

where $\mathbf{F}$ is the data or Vandermonde matrix given as

$$\mathbf{F}_{ij} = \phi_j(\boldsymbol{\omega}_i), \tag{3.7}$$

for data points $i = 1...N$ and basis terms $j = 0...p$.

One of the benefits of using PC expansions is the ability to easily get the statistical moments of the expansions where the mean and standard deviation are given as

$$\mathrm{E}(\mu) = \beta_0, \tag{3.8}$$

$$\mathrm{E}(\sigma^2) = \sum_{i=1}^{p} \beta_i^2, \tag{3.9}$$

where E represents the mathematical expectation of the quantity. Much more information on PC expansions is provided in the literature [40, 42, 43]. However, this thesis work is not intended to provide an extensive review of PC expansions. Rather, it emphasizes these basic concepts to facilitate the understanding of the new RPC-K metamodel developed in the present work.

3.2 Kriging

Fundamentally, PC-K metamodels are a special form of Kriging method. All Kriging (Gaussian process) methods use variations of the basic linear regression estimator $\hat{\mathrm{U}}(\boldsymbol{\omega})$ which is defined as [59]

$$\hat{\mathrm{U}}(\boldsymbol{\omega}) - m(\boldsymbol{\omega}) = \sum_{i=1}^{N} \lambda_i [\mathrm{U}(\boldsymbol{\omega}_i) - m(\boldsymbol{\omega}_i)]. \tag{3.10}$$

In Eq. (3.10), $\boldsymbol{\omega}$ is the input parameter vector for an estimation point, $\boldsymbol{\omega}_i$ is the location vector for the neighboring data point i, N is the number of data points used for estimation, $\hat{\mathrm{U}}(\boldsymbol{\omega})$ is the estimated value at $\boldsymbol{\omega}$, $\mathrm{U}(\boldsymbol{\omega}_i)$ is the value at $\boldsymbol{\omega}_i$, $m(\boldsymbol{\omega})$ is the expected mean of $\mathrm{U}(\boldsymbol{\omega})$, $m(\boldsymbol{\omega}_i)$ is the expected mean of $\mathrm{U}(\boldsymbol{\omega}_i)$, and λ_i is the Kriging weight for datum $\mathrm{U}(\boldsymbol{\omega}_i)$ when used to estimate $\mathrm{U}(\boldsymbol{\omega})$.

$\mathrm{U}(\boldsymbol{\omega})$ can be treated as a random field with mean $m(\boldsymbol{\omega})$, giving a residual of $R(\boldsymbol{\omega}) = \mathrm{U}(\boldsymbol{\omega}) - m(\boldsymbol{\omega})$, also treated as a random field. In simple Kriging, the known residuals are used to calculate an unknown mean. In ordinary Kriging (OK), a known mean is used and in universal Kriging (UK) a known trend function is used for the mean. The Kriging weights are solved from a system of linear equations derived by minimizing the variance of the estimated value. These equations depend on the covariances of the data points used for calculations. First, however, a correlation

equation is fitted to the data. This correlation equation is known as the autocorrelation function as it describes the correlation of the data with itself based on the distance between two points in the parameter space. It is scaled to the covariance values using the process variance. This correlation equation is then used for all future calculations as the exact covariance between known points and the estimation point is unknown and must be estimated. This correlation equation is constructed so it is only dependent on the distance between points, known as lag.

3.2.1 Basic Formulation

Kriging is a best linear unbiased predictor (BLUP). The bias of a predictor is given as the difference between the true value and the predictor's expected value of a variable at that point. An unbiased predictor does not introduce global bias which would consistently over- or under-predict values. A consistent predictor, or one which converges to zero error as the number of data points increases, is asymptotically unbiased. The BLUP is "best" in the sense that it has the least amount of variance for an unbiased predictor; however, other biased predictors may exist with less variance. Generally, unbiased predictors with slightly greater variance are the best predictor for a problem. The error of the model is defined as the difference between the true value and the estimate as shown in Eq. 3.11.

$$\text{Err} = \mathrm{U}(\boldsymbol{\omega}) - \hat{\mathrm{U}}(\boldsymbol{\omega}) \tag{3.11}$$

However, since the true value of the variable is un-observable, terms derived from it are actually estimates, though that notation is dropped here for simplicity. The tools used to estimate the error are presented in section 3.3. When determining the Kriging weights, the goal is to minimize the variance of the unbiased predictor by

$$\sigma_E^2(\boldsymbol{\omega}) = \text{Var}[\hat{\mathrm{U}}(\boldsymbol{\omega}) - \mathrm{U}(\boldsymbol{\omega})]. \tag{3.12}$$

The unbiased constraint for this predictor can be expressed mathematically as

$$\mathrm{E}[\hat{\mathrm{U}}(\boldsymbol{\omega}) - \mathrm{U}(\boldsymbol{\omega})] = \mathrm{E}[\hat{\mathrm{U}}(\boldsymbol{\omega})] - \mathrm{U}(\boldsymbol{\omega}) = 0, \tag{3.13}$$

where E is the mathematical expectation.

The random field is then decomposed into residual and trend components, so the residual is treated as a random field with a stationary mean of zero across the domain, $\mathrm{E}[R(\boldsymbol{\omega})] = 0$. Further, the variance of the random field is held constant across the domain. This means the autocorrelation of the residuals is also stationary. As such, the data is assumed to have weak-sense stationarity. This means that the covariance is solely a function of the lag between points, $\boldsymbol{h}$, and not their location in the parameter space. These assumptions can usually be verified by plotting the covariances of the data. The resulting covariance relationships are

$$\begin{aligned}
\mathrm{Cov}[R(\boldsymbol{\omega}), R(\boldsymbol{\omega}+\boldsymbol{h})] &= \mathrm{E}[(R(\boldsymbol{\omega}) - \mathrm{E}[R(\boldsymbol{\omega})])(R(\boldsymbol{\omega}+\boldsymbol{h}) - \mathrm{E}[R(\boldsymbol{\omega}+\boldsymbol{h})])] \\
&= \mathrm{E}[(R(\boldsymbol{\omega}) - 0)(R(\boldsymbol{\omega}+\boldsymbol{h}) - 0)] \\
&= \mathrm{E}[R(\boldsymbol{\omega}) \cdot R(\boldsymbol{\omega}+\boldsymbol{h})] = C_R(\boldsymbol{h}) = C(\boldsymbol{h}).
\end{aligned} \tag{3.14}$$

Here C_R is the covariance function for the residuals determined empirically from semivariogram data. It is assumed that the covariance function for the residuals is equal to the covariance function for the data, $C_R = C$. All covariances in Kriging methods are calculated by Eq. (3.14).

3.2.2 Universal Kriging

In UK, or Kriging with a trend, Eq. (3.10) can be rewritten as

$$\hat{\mathrm{U}}(\boldsymbol{\omega}) = m(\boldsymbol{\omega}) + \sum_{i=1}^{N} \lambda_i [\mathrm{U}(\boldsymbol{\omega}_i) - m(\boldsymbol{\omega})]. \tag{3.15}$$

In UK, the mean ($m(\boldsymbol{\omega})$), is taken to be a polynomial trend. For example, for a first-order trend on a two-parameter domain, the mean would be represented as

$$m(\boldsymbol{\omega}) = m(\omega_1, \omega_2) = \beta_0 + \beta_1 \omega_1 + \beta_2 \omega_2, \tag{3.16}$$

where β_0, β_1, and β_2 are coefficients determined from a regression of the data. This example model can also be expressed as

$$m(\omega) = \sum_{k=0}^{2} \beta_k \omega_k, \tag{3.17}$$

where β_k are the coefficients, ω_k are the variables (ω_0 is defined as 1, so β_0 is a constant), and there are three coefficients in the equation. Notice how this form is very similar to that of PC, except that in PC the basis parameters ω_k are re-cast into multivariate orthogonal basis terms of the form $\phi_k(\boldsymbol{\omega})$ to assure convergence. While UK allows any form of polynomial for the trend function, PC-K uses PC expansions as the trend for UK and the following math will be presented using this notation.

This modification allows Eq. (3.10) to be rewritten so that the BLUP becomes

$$\hat{\mathrm{U}}_{\mathrm{UK}}(\boldsymbol{\omega}) = \sum_{i=1}^{N} \lambda_i^{\mathrm{UK}} \sum_{k=0}^{p} \beta_k \phi_k(\boldsymbol{\omega}). \tag{3.18}$$

The unbiased constraint then takes the form

$$\begin{aligned} \mathrm{E}[\hat{\mathrm{U}}_{\mathrm{UK}}(\boldsymbol{\omega}) - U(\boldsymbol{\omega})] &= \mathrm{E}[\hat{\mathrm{U}}_{\mathrm{UK}}(\boldsymbol{\omega})] - \mathrm{E}[\mathrm{U}(\boldsymbol{\omega})] \\ &= \mathrm{E}[\sum_{i=1}^{N} \lambda_i^{\mathrm{UK}} \sum_{k=0}^{p} \beta_k \phi_k(\boldsymbol{\omega})] - \sum_{k=0}^{p} \beta_k \phi_k(\boldsymbol{\omega}) \\ &= \sum_{i=1}^{N} \lambda_i^{\mathrm{UK}} \sum_{k=0}^{p} \mathrm{E}[\beta_k \phi_k(\boldsymbol{\omega})] - \sum_{k=0}^{p} \beta_k \phi_k(\boldsymbol{\omega}). \end{aligned} \tag{3.19}$$

Then, requiring the weights to sum to one and assuming the unbiased constraint, the following form is obtained as

$$\begin{aligned} \sum_{i=1}^{N} \lambda_i^{\mathrm{UK}} \sum_{k=0}^{p} \mathrm{E}[\beta_k \phi_k(\boldsymbol{\omega})] &= \sum_{k=0}^{p} \beta_k \phi_k(\boldsymbol{\omega}) \\ \sum_{i=1}^{N} \lambda_i^{\mathrm{UK}} \phi_k(\boldsymbol{\omega}) &= \phi_k(\boldsymbol{\omega}). \end{aligned} \tag{3.20}$$

The Lagrange equation is then calculated as

$$L = \sigma_E^2(\omega) + 2\sum_{k=0}^{p} \mu_k \left(\sum_{i=1}^{N} \lambda_i^{\mathrm{UK}} \phi_k(\boldsymbol{\omega}) - 1\right). \tag{3.21}$$

Equation (3.21) is then minimized, and the resulting equation can be expressed in matrix notation as

$$\mathbf{K}\boldsymbol{\lambda} = \mathbf{k}, \tag{3.22}$$

where $\mathbf{K}$ is the Kriging data matrix, $\boldsymbol{\lambda}$ is the vector of Kriging weights and Lagrange parameters, and $\mathbf{k}$ is the correlation vector for the estimation point.

A sample expansion of the matrix form when solving for a point in an $[\omega_1, \omega_2]$ domain using a linear fit for the mean and using five neighboring points for estimation is used to illustrate the mathematical form. The Kriging data matrix $\mathbf{K}$ is composed of the correlation matrix, $\mathbf{R}$, the Vandermonde matrix, $\mathbf{F}$, and a zeros matrix $\mathbf{0}$ as shown in Eq. (3.23).

$$\mathbf{K} = \begin{bmatrix} \mathbf{R} & \mathbf{F} \\ \mathbf{F}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \tag{3.23}$$

For the components (R) of the correlation matrix ($\mathbf{R}$), the notation $R(h_{21})$ indicates the autocorrelation between data points 2 and 1 based on the fitted autocorrelation function while $R(0)$ is the autocorrelation for zero lag which is assumed to be zero in our models. The notation $\phi_2(\boldsymbol{\omega}^1)$ indicates the second basis term based on the input parameters for data point one. The subscript and superscript e on the right-hand side indicates the estimation point. It can be seen that in addition to the five weights for the Kriging model, there are three Lagrange parameters, one for each basis term in the PC expansion.

$$\begin{bmatrix} R(0) & R(h_{12}) & R(h_{13}) & R(h_{14}) & R(h_{15}) & 1 & \phi_1(\boldsymbol{\omega}^1) & \phi_2(\boldsymbol{\omega}^1) \\ R(h_{21}) & R(0) & R(h_{23}) & R(h_{24}) & R(h_{25}) & 1 & \phi_1(\boldsymbol{\omega}^2) & \phi_2(\boldsymbol{\omega}^2) \\ R(h_{31}) & R(h_{32}) & R(0) & R(h_{34}) & R(h_{35}) & 1 & \phi_1(\boldsymbol{\omega}^3) & \phi_2(\boldsymbol{\omega}^3) \\ R(h_{41}) & R(h_{42}) & R(h_{43}) & R(0) & R(h_{45}) & 1 & \phi_1(\boldsymbol{\omega}^4) & \phi_2(\boldsymbol{\omega}^4) \\ R(h_{51}) & R(h_{52}) & R(h_{53}) & R(h_{54}) & R(0) & 1 & \phi_1(\boldsymbol{\omega}^5) & \phi_2(\boldsymbol{\omega}^5) \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ \phi_1(\boldsymbol{\omega}^1) & \phi_1(\boldsymbol{\omega}^2) & \phi_1(\boldsymbol{\omega}^3) & \phi_1(\boldsymbol{\omega}^4) & \phi_1(\boldsymbol{\omega}^5) & 0 & 0 & 0 \\ \phi_2(\boldsymbol{\omega}^1) & \phi_2(\boldsymbol{\omega}^2) & \phi_2(\boldsymbol{\omega}^3) & \phi_2(\boldsymbol{\omega}^4) & \phi_2(\boldsymbol{\omega}^5) & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} R(h_{e1}) \\ R(h_{e2}) \\ R(h_{e3}) \\ R(h_{e4}) \\ R(h_{e5}) \\ 1 \\ \phi_1(\boldsymbol{\omega}^e) \\ \phi_2(\boldsymbol{\omega}^e) \end{bmatrix}$$

The estimate of a point (Eq. (3.18)) is obtained by using the weights calculated from Eq. (3.22). The variance of the estimation point is solved by

$$\sigma_E^2(\omega) = \sigma^2 - \boldsymbol{\lambda}'\mathbf{k}. \tag{3.24}$$

3.3 Polynomial Chaos-Kriging

Let a computational model $\mathcal{M}$ be defined which maps an input vector of n dimensions, $\boldsymbol{\omega}$, to an output, U. The random vector $\boldsymbol{\omega}$ has a joint probability density function of $f_{\boldsymbol{\omega}}$ which is a tensor product of the n independent probability density functions. When the computational model is the PC-K metamodel, it takes the form

$$\mathcal{M}(\boldsymbol{\omega}) = \mathrm{U}(\boldsymbol{\omega}) \approx \mathrm{U}^{\text{PC-K}}(\boldsymbol{\omega}) = \sum_{i \in \mathcal{A}}^{p} \beta_i \phi_i(\boldsymbol{\omega}) + \sigma^2 Z(\boldsymbol{\omega}, \vec{\xi}). \tag{3.25}$$

Here, $\sum_{\alpha \in A}^{p} \beta_\alpha \phi_\alpha(\boldsymbol{\omega})$ is the trend defined by $|\mathcal{A}|$ multivariate orthonormal polynomial basis terms, $\phi_\alpha(\boldsymbol{\omega})$, and their coefficients, β_α, indexed by the multi-index, α. When a dense polynomial expansion is used, the active set of basis terms $\mathcal{A}$ is equal to the $p+1$ multivariate polynomials constructed from the individual polynomial expansions. When a sparse set of terms is selected from the LARS algorithm, the active set is defined by the minimum error model selected by LARS. The random

process $Z(\boldsymbol{\omega}, \vec{\xi})$ is defined by the autocorrelation function which is fit to the input parameters based on their probability density functions, $\vec{\xi}$, and σ^2 is the process variance.

The complex forms of the Kriging equations are then simplified into matrix form and solved. The coefficients are solved for using linear regression as

$$\boldsymbol{\beta} = \left(\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{U}, \tag{3.26}$$

and the process variance is solved for as

$$\sigma^2 = \frac{1}{N}(\mathbf{U} - \mathbf{F}\boldsymbol{\beta})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{U} - \mathbf{F}\boldsymbol{\beta}), \tag{3.27}$$

where again $\mathbf{R}$ is the autocorrelation matrix, $\mathbf{F}$ is the Vandermonde matrix, $\boldsymbol{\beta}$ is the vector of coefficients, $\mathbf{U}$ is the output response of the evaluated design points, and N is the number of evaluated design points [28]. Once the Kriging metamodel has been constructed, the estimated mean at a new design point can be given as

$$\mu_{\hat{\mathrm{U}}}(\boldsymbol{\omega}) = \boldsymbol{f}(\boldsymbol{\omega})^{\mathrm{T}}\boldsymbol{\beta} + \mathbf{r}(\boldsymbol{\omega})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{U} - \mathbf{F}\boldsymbol{\beta}), \tag{3.28}$$

and the variance at this design point as

$$\sigma^2_{\hat{\mathrm{U}}}(\boldsymbol{\omega}) = \sigma^2\left(1 - [\boldsymbol{f}(\boldsymbol{\omega})^{\mathrm{T}}\mathbf{r}(\boldsymbol{\omega})^{\mathrm{T}}]\begin{bmatrix}\mathbf{0} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{R}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{f}(\boldsymbol{\omega}) \\ \mathbf{r}(\boldsymbol{\omega})\end{bmatrix}\right), \tag{3.29}$$

where $\boldsymbol{f}(\boldsymbol{\omega})$ is the polynomial basis evaluated at the new design point defined by $\boldsymbol{\omega}$ and $\mathbf{r}(\boldsymbol{\omega})$ is the correlation vector evaluated between the new design point and $\boldsymbol{\omega}$.

If the PC-K metamodel is not sufficiently accurate for the problem, the experimental design can be adaptively refined until the metamodel is sufficiently accurate. While it is possible to refine the problem globally by increasing the number of samples used in the experimental design, it is more computationally efficient to sample based on the estimated accuracy of the metamodel. To do

so, a criterion needs to be defined. For this, the criterion produced by Kawai and Shimoyama [69] was adapted for a PC-K metamodel. The criterion is given as

$$\mathrm{Crit}(\boldsymbol{\omega}) = \left(\left| \frac{\partial \hat{\mathrm{U}}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} \right| \Delta\omega + D_{\hat{\mathrm{U}}}(\boldsymbol{\omega}) \right) \sigma^2_{\hat{\mathrm{U}}}(\boldsymbol{\omega}) \mathrm{PDF}(\boldsymbol{\omega}). \tag{3.30}$$

Here $\frac{\partial \hat{\mathrm{U}}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} = \left[\frac{\partial \hat{\mathrm{U}}(\boldsymbol{\omega})}{\partial \omega_1}, \frac{\partial \hat{\mathrm{U}}(\boldsymbol{\omega})}{\partial \omega_2}, ..., \frac{\partial \hat{\mathrm{U}}(\boldsymbol{\omega})}{\partial \omega_n} \right]^{\mathrm{T}}$ is the vector of gradients of the predicted mean, $\hat{\mathrm{U}}(\boldsymbol{\omega})$, with respect to the input variables, $\boldsymbol{\omega}$. It can be derived analytically as

$$\frac{\partial \hat{\mathrm{U}}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} = \left(\frac{\partial \mathbf{r}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} \right)^{\mathrm{T}} \mathbf{R}(\boldsymbol{\theta})^{-1} (\mathbf{U} - \mathbf{F}\boldsymbol{\beta}), \tag{3.31}$$

where $\frac{\partial \mathbf{r}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}}$ is an $N \times n$ matrix with (i, j) entries

$$\left[\frac{\partial \mathbf{r}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} \right]_{i,j} = R'(\omega_j - \omega_j^i) \times \prod_{\substack{k=1 \\ k \neq j}}^{n} R(\omega_k - \omega_k^i). \tag{3.32}$$

Again, N is the number of data points. It is multiplied by $\Delta\omega = \min_{\mathrm{i=1,2,...,n}} |\boldsymbol{\omega} - \boldsymbol{\omega}^{\mathrm{i}}|$ to keep the units consistent with $D_{\hat{\mathrm{U}}}(\boldsymbol{\omega})$. The term $D_{\hat{\mathrm{U}}}(\boldsymbol{\omega})$ is defined as

$$D_{\hat{\mathrm{U}}}(\boldsymbol{\omega}) = |\hat{\mathrm{U}}(\boldsymbol{\omega}) - \hat{\mathrm{U}}_{\mathrm{pre}}(\boldsymbol{\omega})|, \tag{3.33}$$

which is the difference in the estimate at the design point of the current sample set and the previous set of $N-1$ points. For the first iteration, $D_{\hat{\mathrm{U}}}(\boldsymbol{\omega})$ is set to zero. The value $\sigma^2_{\hat{\mathrm{U}}}$ is taken from Eq. (3.27) and $\mathrm{PDF}(\boldsymbol{\omega})$ is taken from the probability density functions of the uncertainties in the input variables.

This criterion was used as it was found to work well for Kriging due to its inclusion of both uncertainty and gradient information [69]. This method is superior to just gradient-based methods as gradients are generally fit well after a few refinement steps, leaving large regions unrefined afterwards. However, the high-gradient regions can have low estimated uncertainty but not be fit accurately, so throwing away this information completely is unwise. Finding the criterion for a

potential new sample location is not easy as many points in the problem domain must be evaluated and sometimes the criterion averaged over multiple metamodels fit to several output variables, times, or locations. One method is to take the global maximum criterion. Another is to find the mode of the maximum criteria from all the considered metamodels, although these methods may not be optimal.

3.3.1 Least Angle Regression

For a case with $n = 8$ design parameters and maximum degree of $m = 3$ there are a total of 165 terms in the polynomial basis. This large expansion size can be trimmed down by using a model selection algorithm to select only those terms which have the greatest effect on the response output. Least angle regression estimates this effect by finding the terms in order of most correlation with the response output [46].

The process can be visualized in two dimensions in Fig. 3.3. First, the correlation of ω_1 and ω_2 with U is determined from the initial estimate $\hat{\mu}_0$. In the case of Fig. 3.3, ω_1 is most correlated with the output because moving in its direction better describes the output than moving in the direction of ω_2. To move in this direction, the value of the coefficient for ω_1 is increased. The coefficient for ω_1 is increased until the prediction point $\hat{\mu}_1$ is reached. At this point, both ω_1 and ω_2 are equally correlated with U, so both their coefficients are increased together until the estimate equals the output response. What makes LARS unique from other methods is the ability to accurately determine the amount the coefficient of a term needs to be increased until the next term is equally correlated, so that only one step is needed per term. For higher dimensional problems, this procedure is repeated for each additional term.

The mathematical formulation is presented next. First, these basis terms need to be transformed into valid covariates for the LARS algorithm to sort by standardizing them with zero mean and unit variance. In addition, the output response needs to be normalized to zero mean. These basis terms exist for all N simulations, so the $N \times 1$ covariate vectors are defined as $\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, ...\boldsymbol{\phi}_p$ where $\boldsymbol{\phi}_0 = [\phi_0^1, \phi_0^2, ...\phi_0^N]^{\mathrm{T}}$. The N output responses, $\mathbf{U}$, can be estimated with a linear model as $\hat{\mathbf{U}}$.

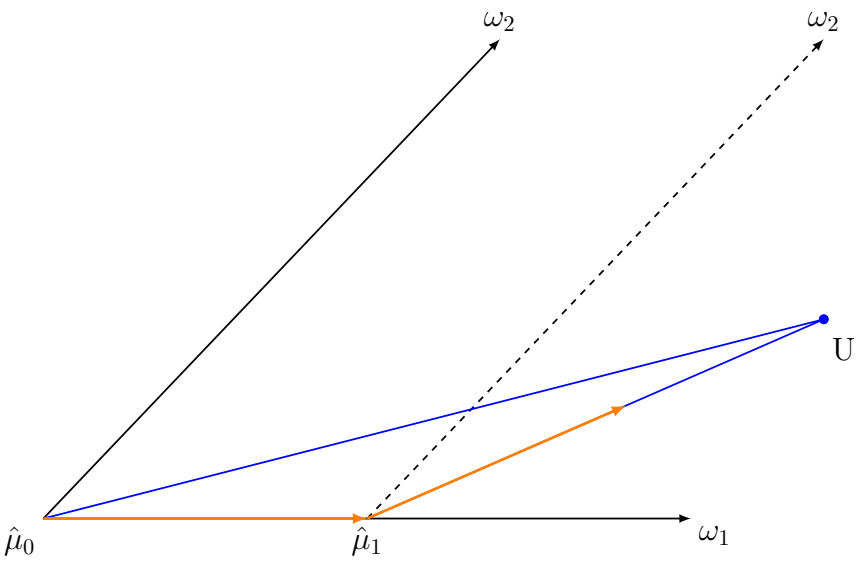


Figure 3.3: Visualization of LARS in 2D.

Then, the problem can be expressed as

$$\mathbf{U} \approx \hat{\mathbf{U}} = \sum_{i=0}^{p} \boldsymbol{\phi}_i \hat{\beta}_i = \boldsymbol{\Phi}\hat{\boldsymbol{\beta}}, \tag{3.34}$$

where $\boldsymbol{\Phi} = [\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, ...\boldsymbol{\phi}_p]$ and $\hat{\boldsymbol{\beta}} = [\hat{\beta}_0, \hat{\beta}_1, ...\hat{\beta}_p]$ which are the estimated regression coefficients for each basis term. The model estimates associated with these coefficients have a total squared error of

$$S(\hat{\boldsymbol{\beta}}) = \sum_{i=1}^{N} (\mathrm{U}_i - \hat{\mathrm{U}}_i)^2, \tag{3.35}$$

and an absolute coefficient norm of

$$T(\hat{\boldsymbol{\beta}}) = \sum_{i=0}^{p} |\hat{\beta}_i|. \tag{3.36}$$

The general optimization problem is to minimize the error, $S(\hat{\boldsymbol{\beta}})$, under some complexity bound $t \geq T(\hat{\boldsymbol{\beta}})$. In general, the best bound for t is hard to determine as the amount of complexity a problem can handle depends upon the type of model and the size of the experimental design.

The first response estimates can be given as $\hat{\mathbf{U}}_0$ and are assigned a value of zero. Then the current correlations of the covariates are calculated as

$$\hat{\mathbf{c}} = \mathbf{\Phi}^{\mathrm{T}}(\mathbf{U} - \hat{\mathbf{U}}_{\mathcal{A}}), \tag{3.37}$$

where the subscript $\mathcal{A}$ denotes the set of included covariate terms. Initially, the active set is empty and no covariates have been chosen. The most highly-correlated covariate is chosen first because when the estimate is moved in this direction, it converges the most quickly. This direction is defined by the equiangular vector, $\mathbf{u}_{\mathcal{A}}$. The estimate for the response is moved in this equiangular direction by increasing the coefficient for the respective covariates. This continues until another covariate is equally correlated with the response. At this equal correlation point, the new covariate is added to the active set and the new estimate of the response is given as

$$\hat{\mathbf{U}}_{\mathcal{A}} = \hat{\mathbf{U}}_{\mathcal{A}-1} + \hat{\gamma}\mathbf{u}_{\mathcal{A}}, \tag{3.38}$$

where $\hat{\gamma}$ is the distance traveled, or step size, in the equiangular direction. The step size can be calculated as

$$\hat{\gamma} = \min^{+}_{i \in \mathcal{A}^c} \Big\{ \frac{\hat{C} - \hat{c}_i}{A_{\mathcal{A}} - a_i}, \frac{\hat{C} + \hat{c}_i}{A_{\mathcal{A}} + a_i} \Big\}, \tag{3.39}$$

where $\min^{+}$ indicates that the minimum is taken of the positive components only, $\hat{C}$ is the maximum correlation, $A_{\mathcal{A}}$ is a covariate term, and a_i is the ith inner product term [46]. When a new covariate is added, $\mathbf{u}_{\mathcal{A}}$ is re-computed so that it is equiangular between the covariates in the active set and the output response. This new equiangular direction is still the direction which most quickly approaches the response as the coefficients of the active covariate terms are increased. The process is repeated until all covariate terms are included. The method is deemed optimal as the step size can be reliably computed for each covariate, and the number of step sizes is exactly equal to the number of covariates [46]. This improves upon the least absolute shrinkage and selection operator (Lasso) and forward stagewise regression techniques which do not compute and move in the *least angle* direction. Thus, they take more computations.

It is generally advised to construct models so that low-order basis terms are added before their respective higher-order basis terms because the resulting models are more physically realistic and

easier to interpret [46–48]. Additionally, because of the sparsity of effects principle, low-order basis terms carry most of the process variance [80]. One possible method to achieve this was originally proposed by Efron et al. [46]. This method runs the LARS algorithm on only the first-order basis terms first, then on second-order basis terms, and so on, until a desired model is reached. Although this method builds the metamodel in a logical order, it either restricts the metamodel to having all parameters represented equally at lower-orders or requires restricting the basis terms of each order of the metamodel separately. To get around these constraints, the heredity of the basis terms can be exploited. This means higher-order basis terms are only added to the metamodel if all of its lower-order, dependent basis terms are in the metamodel. For strong heredity, the correlation of a higher-order basis term which does not have its dependent terms in the metamodel is calculated as the average of the higher-order and dependent basis terms [47]. This method penalizes higher-order basis terms which have dependent terms with low correlation from joining the model. Unfortunately, if a higher-order basis term is chosen, all dependent basis terms are added simultaneously. This may cause difficulties in very small data sets when adding multiple basis terms could increase the error significantly. Instead, a hierarchical LARS method which only adds one covariate at a time is implemented in this thesis research. It only chooses basis terms which have all dependent basis terms already included in the active set.

The hierarchical LARS method at first only considers the first-order basis terms of the problem. This is accomplished by selecting covariates for the active set from a candidate set. The candidate set is the subset of eligible basis terms from all available basis terms. So at first, this is only first-order basis terms. The terms in the candidate set are chosen by heredity, so the lower-order versions of a basis term must be chosen first. As such, once a first-order term has been chosen, its respective second-order term is added to the candidate set. This process is repeated so that all higher-order terms are not included in the candidate set until their lower-order, dependent basis terms are included in the active set. Interaction terms are only added to the candidate set once the basis terms of each component have been added to the active set to the equal order of the interaction basis term. This follows the strong heredity principle [47]. For example, the interaction

term $\phi(\omega_1^2, \omega_2^3)$ will only be included in the candidate set once the basis terms $\phi(\omega_1^2)$ and $\phi(\omega_2^3)$ have been added to the active set. If any of basis terms added to the candidate set have a correlation higher than the current max, those terms are added without stepping in any direction. This process is repeated until all covariates are chosen.

The LARS algorithm is used to find the order of importance for the covariates so the sparse PC expansions are constructed with the most impactful terms only. A separate PC expansion can be built for each basis size from 1 to $p + 1$. Each of these expansions includes the most correlated basis terms for that size. For example, a PC expansion with five terms is built from the first five basis terms chosen through the LARS algorithm. These terms describe the data better than any other combination of five terms. For each of these $p + 1$ expansions, a PC-K metamodel can be constructed and its generalization error found. Then, the PC-K metamodel used in the final UQ study is the one with minimal error. This methodology follows the optimal PC-K method described by Schöbi [28]. The error measures used for this comparison are prescribed next.

3.3.2 Error

When constructing a metamodel, the generalization error is used to evaluate the fit of the model. One of the most efficient ways to approximate the generalization error of a sparse metamodel is through the leave-one-out cross validation (LOOCV) measure. For a problem with N design points, the method calculates the metamodel with $N-1$ points. Then the left-out design point is predicted from the resulting $N-1$ metamodel and the mean squared error calculated for this estimate. This process is then repeated for all N design points and a resulting average mean squared error is calculated [3, 93]. The LOOCV error is the special form of leave-k-out cross validation with $k = 1$. Leave-one-out differs from leave-k-out cross validation because LOOCV is exhaustive in calculating the error for every design point and has analytical solutions. The LOOCV error for a PC-K metamodel is expressed as

$$\mathrm{Err}_{\mathrm{LOO}} = \frac{1}{N} \sum_{i=1}^{N} \left(\mathrm{U}_i - \mathrm{U}_{-i}^{(\text{PC-K})}(\boldsymbol{\omega}_i) \right)^2, \tag{3.40}$$

where $\mathrm{U}_{-i}^{(\text{PC-K})}(\boldsymbol{\omega}_i)$ is the PC-K metamodel constructed from the simulation data set minus point i evaluated at point i.

The analytical forms for this error exist for both PC and PC-K methods, allowing the error to be calculated quickly and efficiently. The leave one out error for a PC metamodel can be calculated as

$$\mathrm{Err}_{\mathrm{LOO}}^{(\text{PC})} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mathrm{U}_i - \mathrm{U}^{(\text{PC})}(\boldsymbol{\omega}_i)}{1-h_i}\right)^2, \tag{3.41}$$

where h_i is the ith diagonal term in $\mathbf{h} = \mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}$. This form requires the construction of only the full PC metamodel. Polynomial chaos-Kriging does not have a different compact form, but work by Olivier Dubrule [56] found a way to calculate $\mathrm{U}_{-i}^{(\text{PC-K})}(\boldsymbol{\omega}_i)$ without computing the N individual metamodels. The mean estimate is calculated as

$$\mathrm{U}_{-i}^{(\text{PC-K})}(\boldsymbol{\omega}_i) = -\sum_{j=1,j\neq i}^{N}\frac{\mathbf{B}_{ij}}{\mathbf{B}_{ii}}\mathrm{U}_j = -\sum_{j=1}^{N}\frac{\mathbf{B}_{ij}}{\mathbf{B}_{ii}}\mathrm{U}_j + \mathrm{U}_i, \tag{3.42}$$

where the new matrix $\mathbf{B}$ is given by

$$\mathbf{B} = \begin{bmatrix} \sigma^2\mathbf{R} & \mathbf{F} \\ \mathbf{F}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}^{-1}. \tag{3.43}$$

In Eq. (3.43), σ^2 is again the Kriging process variance, $\mathbf{R}$ is the correlation matrix, and $\mathbf{F}$ is the Vandermonde matrix from the full PC-K metamodel. These analytical forms of the LOOCV error can quickly provide accurate estimates of the accuracy for PC and PC-K metamodels. However, LOOCV errors are not perfect. Cross validation estimators, while nearly unbiased, are always slightly biased as the training set does not include all data points. This bias can become an issue in special cases if not accounted for, such as in binary distributions with equal probability. Additionally, the variance of the estimators can be large and the estimator with the lowest error may not always be the best. Finally, cross validation does not account for the appropriateness of the models used for the problem. The unbiased constraint is generally not a concern; however, the

other two issues are pertinent in metamodeling. For this thesis work, only PC-K metamodels are being investigated, so model appropriateness issues are not considered. Additionally, while the lowest error model is not always best, this is a result of the bias-variance decomposition of mean squared error estimates and not specific to LOOCV.

The mean squared error is used in the metamodel selection as the Euclidean norm is often the easiest to minimize. The total, or L_0 norm, is frequently the norm which should be minimized, but its discrete nature makes it mathematically hard to solve. The next closest norm, the L_1, can be minimized but requires sampling to do so as it is not uniform. This leaves the L_2 norm, or mean squared error, as the best norm to minimize [45]. Various criterion have been developed that can assist with metamodel selection. Two of the most popular ones are the Akaike information criterion (AIC) and Bayesian information criterion (BIC) which are a function of the mean squared error and the number of terms in the model. This is useful in problems with many data points as the mean squared error, after initially decreasing, stays relatively flat for a long time before increasing as over-fitting occurs. These criteria help penalize metamodels with a large number of terms, so metamodels with fewer terms and nearly the same mean squared error are chosen instead. This is generally preferred. However, in sparse problems, the mean squared error rapidly increases after the minimum as over-fitting occurs much earlier. This limits their usefulness. While the AIC and BIC criterion are not analyzed in this thesis work, the training error is analyzed as it can provide useful information for the model. Now that the full PC-K metamodeling process is complete, the process of reducing the problem for the application of a RPC-K metamodel can be easily presented.

3.4 Reduced Polynomial Chaos-Kriging

3.4.1 Sensitivity Analysis

The cornerstone of the development of the RPC-K metamodel is the effective use of sensitivity analysis in order to reduce the number of parameters in the metamodel. In order to reduce the number of parameters in the PC-K metamodeling of a problem, a sensitivity study must be conducted to quantify their impacts. Following Sobol's 1993 work [74], a problem can be represented as

a high-dimensional model representation (HDMR). This represents the problem as a hierarchical superposition of functions that describe the interactions of the random parameters. The HDMR representation is given as

$$\mathrm{U}(\boldsymbol{\omega}) = \mathrm{U}_0 + \sum_{i=1}^{n} \mathrm{U}_i(\omega_i) + \sum_{i<j} \mathrm{U}_{ij}(\omega_i, \omega_j) + \cdots + \mathrm{U}_{12\ldots n}(\omega_1, \omega_2, \ldots, \omega_n). \tag{3.44}$$

Here, U_0 is the mean value, $\mathrm{U}_i(\omega_i)$ represents the contribution of ω_i to U by itself (main effect), and $\mathrm{U}_{ij}(\omega_i, \omega_j)$ represents the impact of the interaction of ω_i and ω_j. When using a PC expansion to represent this form, the total order is limited to m based on a specified error tolerance [94]. From this decomposed form, an analysis of variance (ANOVA) can be performed and sensitivity indices constructed. Different indices can be calculated. The total sensitivity index for a parameter i is the sum of the square of every term and interaction of i. The main effects index is calculated simply from the main effects of that index. Using the principle of sparsity of effects, we can cut the HDMR expansion after the main effects term, and use this cut-HDMR to represent the problem as

$$\mathrm{U}(\boldsymbol{\omega}) \approx \mathrm{U}_0 + \sum_{i=1}^{n} \mathrm{U}_i(\omega_i). \tag{3.45}$$

Then, the second terms of Eq. (3.45) can be expressed as

$$\mathrm{U}_i(\omega_i) = \mathrm{U}(\boldsymbol{\omega})|_{\boldsymbol{\omega}^0 \setminus \boldsymbol{\omega}_i} - \mathrm{U}_0, \tag{3.46}$$

where $\mathrm{U}(\boldsymbol{\omega})|_{\boldsymbol{\omega}^0 \setminus \boldsymbol{\omega}_i}$ represents the response at the mean values of $\boldsymbol{\omega}$ except with ω_i varying [80]. This gives a compact form to express the variation due to a single parameter, i. When using a PC expansion, the left-hand side can be solved for as

$$\mathrm{U}_i(\omega_i) = \sum_{k=1}^{m} \beta_i \phi_i(\boldsymbol{\omega}). \tag{3.47}$$

Using normalized polynomials, this simplifies to

$$\mathrm{U}_i(\omega_i) = \sum_{k=1}^{m} \beta_i. \tag{3.48}$$

The variance of parameter i is then given as

$$\sigma_i^2 = \sum_{k=1}^{m} \beta_i^2. \tag{3.49}$$

Equation (3.49) shows that the main effects variance for parameter i is estimated by the sum of the square of the PC coefficients for the main effects basis terms of i. Again, because of the sparsity of effects principle, this estimate is a good estimate of the total variance due to parameter i. To estimate the variance of the whole problem, we construct a PC expansion. Since the variance of a PC expansion is given in Eq. (3.9) as the sum of the squares of the $i = 1 \ldots p$ coefficients, we can build a simple PC expansion for a problem which includes only main effects terms to estimate the problem variance. From this sparse expansion, the individual sensitivity indices for parameters $i = 1 \ldots n$ are estimated using Eq. (3.49). Then the problem variance is decomposed and estimated as

$$\sigma^2_{\text{True}} \approx \sigma^2_{\text{PC}} = \sigma_i^2 + \sigma_2^2 + \ldots + \sigma_n^2. \tag{3.50}$$

This allows the definition of sensitivity indices for the parameters given as

$$S_i = \frac{\sigma_i^2}{\sigma^2_{\text{PC}}}. \tag{3.51}$$

This produces a sensitivity index for every model parameter i bounded between [0,1] where zero indicates that the parameter has no effect on the output and one indicates that the parameter is the only one which impacts the output [76]. This process identifies those parameters which have very little effect on the output and can be removed from the study with minimal loss in accuracy [75]. If multiple times, outputs, or locations are considered, the average of the sensitivity indices can be used to the same effect. When these parameters are removed, a reduced problem is considered which is computationally more efficient to evaluate and provides sufficient accuracy [8,50,51]. The

RPC metamodel is then defined as the PC expansion over the most impactful parameters. Likewise, the RPC-K metamodel is the specific PC-K metamodel constructed using the RPC polynomial basis.

3.4.2 RPC-K Characteristics

The construction of the metamodel using this reduced basis presents several unique challenges that must be addressed. Most importantly, the total variation of the problem is not accurately modeled because the metamodel is not accounting for various sources of uncertainty. In the parlance of Donald Rumsfeld, the sensitivity analysis gives a list of *known unknowns*. While the exact impact of these parameters on the problem variance is unknown, their relative impact compared to the considered parameters is known. Another result of this model reduction is that the metamodel becomes lower-fidelity. When some parameters are not included in the metamodel, it does not necessarily converge to the true solution even if the number of data points and basis terms is increased. Finally, the RPC-K metamodel suffers less from over-fitting because it allows for greater over-sampling of the problem, but it does not remove the danger of under-fitting the model or having outliers skew the resulting metamodel for sparsely-sampled problems.

The full workflow for using the RPC-K metamodel is presented in Fig. 3.4. First, the input parameters for the problem are chosen and the sampling method determined. Once the design points are known, they are run by the forward model. In the present work, a CFD model is used and it is highlighted in blue in the flowchart. However, the non-intrusive method can be applied equally well to finite element analysis (FEA), experimental data, or any other model. Then, the sensitivity analysis is performed for obtaining the RPC metamodel. This is the critical difference between the RPC-K and other PC-K metamodels. Additionally, the Kriging hyperparameters must be determined before the full RPC-K metamodel is constructed. Then, a Monte Carlo analysis can be run on the metamodel to determine estimates of the solution variable and its variation throughout the solution space. If these estimates are sufficiently accurate, the process is stopped. If the values are not under an acceptable tolerance, further points can be added to the design space and run in

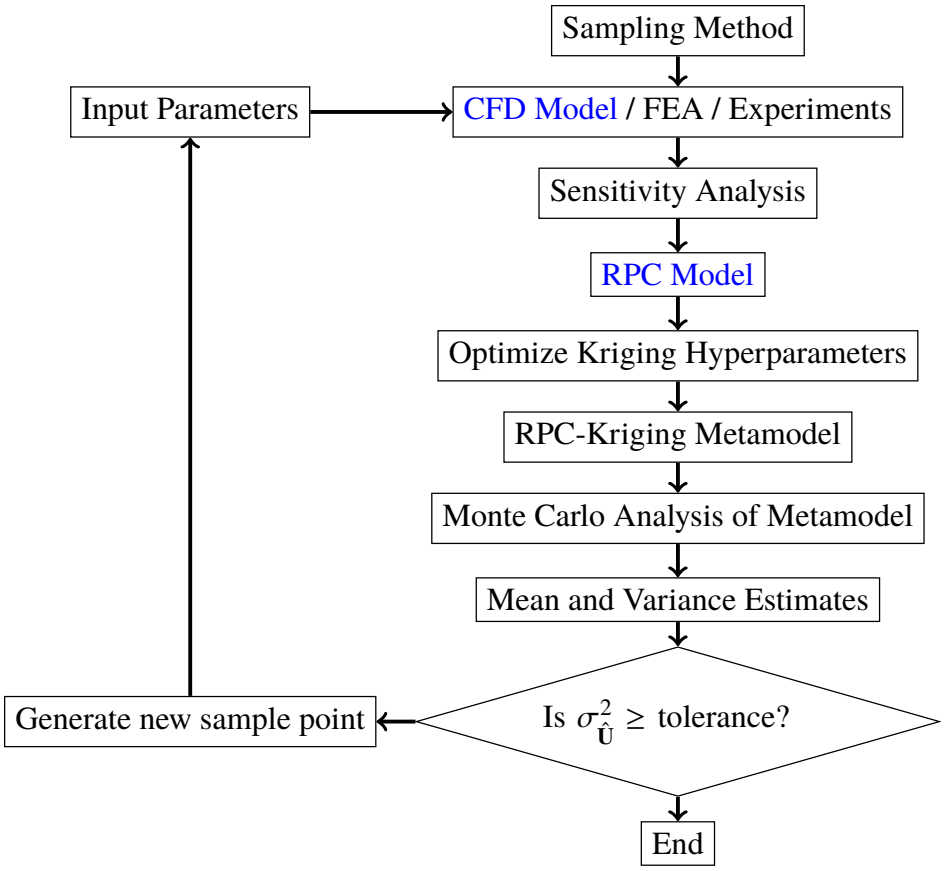


Figure 3.4: RPC-K methodology flowchart.

the model until the metamodel is sufficiently accurate. This metamodel can then be considered complete.

Chapter 4

Validation

The new RPC-K metamodel is first validated on a convection-diffusion-reaction (CDR) model problem. The one-dimensional transient CDR problem is governed by

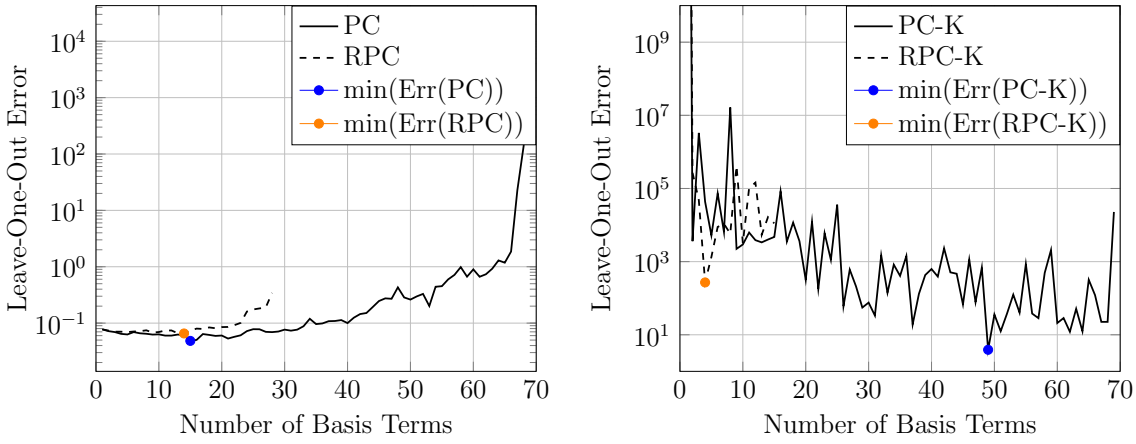
$$\frac{\partial \phi}{\partial t} + u\frac{\partial \phi}{\partial x} - k\frac{\partial^2 \phi}{\partial x^2} + c\phi = f, \tag{4.1}$$

in a unit domain from $0 \leq x \leq 1$ with $f = \alpha(x-1)e^{-t}$. Here ϕ is the solution variable, u is the convection speed, k is the diffusivity, and c and α are coefficients for the source terms. The true values of u, k, c, and α are one, one, one, and two, respectively. The inlet is a Dirichlet boundary condition with $\phi = 0$ since $x = 0$. The outlet is a Neumann boundary condition which varies with time, $\phi(1,t) = 2xe^{-t}$. The analytical solution to the problem is

$$\phi_{\text{exact}}(x,t) = x^2 e^{-t}. \tag{4.2}$$

The problem is assumed to have $n = 4$ parameters and they are viewed as uncertain. To calculate the statistical data of the problem under uncertainty, 10,000 Monte Carlo simulations are used. Each parameter is assumed to have normally distributed uncertainty with a standard deviation of 10% of its mean. The solution is analyzed at $x = 0.5$ from time $t = 0$ to 4 seconds. A sensitivity study [51] was conducted and found the parameters u and k to be the most impactful, so the reduced model is constructed with the $\hat{n} = 2$ parameters. The hierarchical LARS algorithm is applied to the PC, RPC, PC-K, and RPC-K metamodels with a maximum degree of four and the leave one out errors are plotted against the number of basis terms in the polynomial function. Results are shown for the case with 70 model evaluations using Sobol sequence sampling. Figure 4.1 plots the LOOCV error for the metamodels versus the number of terms chosen by the hierarchical LARS algorithm. Figure 4.1a shows the error progression for the PC and RPC metamodels and Fig. 4.1b

shows that for the PC-K and RPC-K metamodels. It can be seen that the PC and RPC metamodels have strong trends in the LOOCV error progression and do not have much variation while the PC-K and RPC-K metamodels do.



(a) Polynomial chaos metamodels.

(b) Polynomial chaos-Kriging metamodels.

Figure 4.1: LOOCV error progression of ϕ for varying basis sizes.

To explore this behavior, multiple types of polynomial bases are investigated in the models instead of the standard probabilists' version of Hermite polynomials. This included the physicists' version, normalized versions, and a simple basis $(x, x^2, x^3, ...)$. The results indicate that better results can exist among varied basis types. Unfortunately, no apparent trend exists for the various basis types. The models have only a narrow range of accuracy and mostly under- or over-predict the variation of the solution. The fluctuations in the LOOCV error for the PC-K metamodels persist through all forms. This is shown in Fig. 4.2 which plots the LOOCV error for the physicists' version of Hermite polynomials, normalized probabilists' Hermite polynomials, and a simple basis. To rule out the addition of the correlation information in the least squares regression for the PC-K metamodels as the cause of the variability, the cases are also considered using the coefficients chosen from plain least squares regression. The variability persists and is likely the result of an inconsistent correlation structure at varying times. Additionally, the basis size of the PC-K metamodels is always chosen to be a large number of terms, and the results seem to indicate that

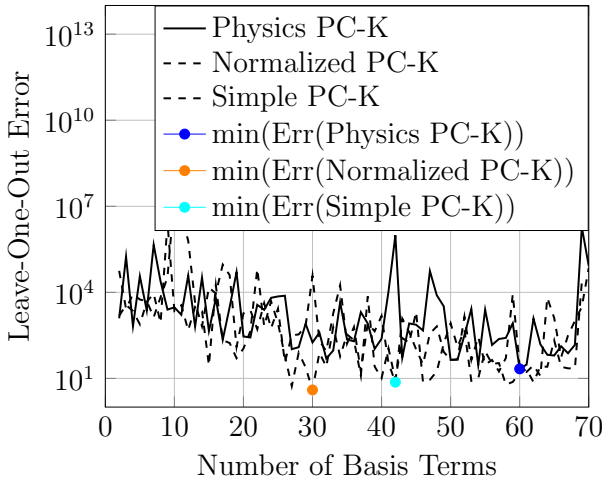
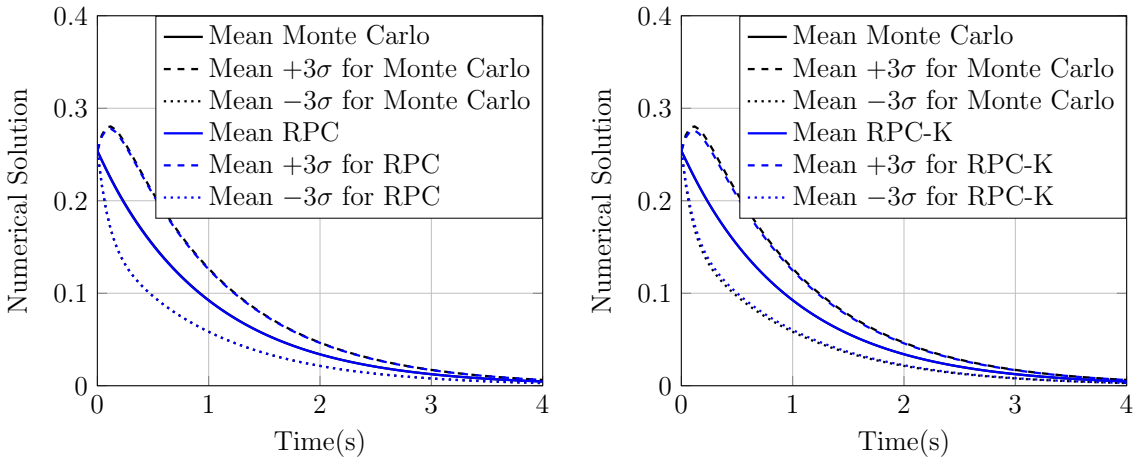


Figure 4.2: LOOCV error for multiple polynomial types.

the models are suffering from over-fitting that the LOOCV error does not show. The root cause of these difficulties in fitting the CDR problem is believed to be the result of the smooth behavior of the function with little variation in the outputs at later time steps. This means that the problem requires more manual interaction than the airfoil and engine nacelle studies. Regardless, when the models do work, the results are excellent as seen in Fig. 4.3.

The mean, $+3\sigma$, and -3σ values of the RPC metamodel are compared to the Monte Carlo data in Fig. 4.3a. It can be seen that the RPC metamodel data agrees well with the Monte Carlo data with an average mean squared error of 2.52×10^{-5} for the mean and 9.8×10^{-3} for the standard deviation. This helps validate the effectiveness of using the hierarchical LARS algorithm and LOOCV error to chose the best RPC metamodel. The mean, $+3\sigma$, and -3σ values of the RPC-K metamodel given by its basis coefficients are compared to the Monte Carlo data in Fig. 4.3b along with the Monte Carlo data of the RPC-K metamodel. The average mean squared error of the mean is 2.13×10^{-7} and it is 1.36×10^{-2} for the standard deviation. The mean is very well predicted and the standard deviation is fairly close, though this accuracy is not robust due to the erratic nature of the LOOCV error of the metamodels.

The PC and PC-K type metamodels can accurately model the example physics problem and the sensitivity study used to reduce the parameter size of a problem is effective in reducing the computational cost while still maintaining sufficient accuracy. Additionally, we have learned through



(a) Results of RPC metamodel for CDR.

(b) Results of RPC-K metamodel for CDR.

Figure 4.3: Mean, $+3\sigma$, and -3σ values of the RPC and RPC-K metamodels compared to Monte Carlo results for the CDR problem.

this research that the construction of RPC-K metamodels is not straightforward due to the interacting effects of fitting data with both a least squares solution and correlation function. Additionally, depending on the conditioning of the data, manual manipulation may be required.

Chapter 5

Results and Discussion

The RPC-K metamodel is applied to two aerodynamics problems. The first problem is a two-dimensional Reynolds-averaged Navier-Stokes (RANS) solution to the NACA 4412 airfoil at low speed and a high angle of attack. The second problem is a RANS solution to study the inlet air of an aircraft engine at takeoff in a strong crosswind. Both of these cases have separated flow regions which we attempt to capture in our metamodel with only a sparse sampling of the problem. ICFD++ [95] is used for all of our CFD simulations.

5.1 Airfoil Problem Configuration

The NACA 4412 CFD cases are run as two-dimensional fully-turbulent cases with a far-field boundary at 100 times the chord length on C grids provided from NASA Langley Research Center's Turbulence Modeling Resource (TMR) [1]. The case is based on experimental wind tunnel data from Coles and Wadcock [36]; however, the CFD case differs from the experiment in that it is two dimensional. Because of this, it is a weak validation case for CFD. A grid convergence study conducted with grids of 113×33, 225×65, 449×129, 897×257, 1793×513 is shown in Fig. 5.2 where the coarsest grid has 65 points on the airfoil surface and the finest grid has 1025 points. The experimental data was collected at 50 points around the surface of the airfoil. For the computational study, the airfoil was modified to close to a sharp trailing point as opposed to the true NACA 4412 airfoil shape [1]. The schematic of the computational study is reproduced from TMR in Fig. 5.1.

The validation cases are run following the TMR setup with Riemann boundary conditions, an adiabatic solid wall for the airfoil, a Mach number (Ma) of 0.09 at an angle of attack (AoA) of 13.87°, a non-dimensional chord length (c) of one, an operating temperature (T) of 536°R (297.8 K), and a Reynolds number (Re) of 1.52 million based on the chord length. The freestream turbulence intensity is set to 0.086% and the turbulent to laminar viscosity ratio to 0.009 [1].

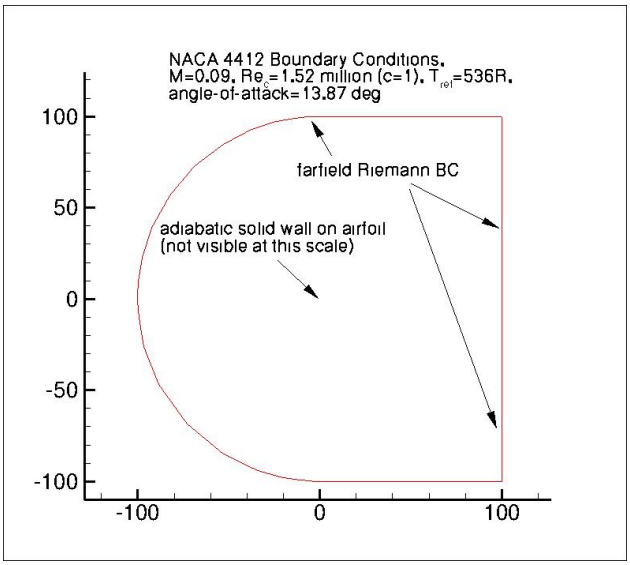


Figure 5.1: Schematic of NACA 4412 computational study, reproduced from the TMR setup [1].

To compute the additional values of the simulation, terms are derived as follows. The speed of sound is calculated as $a = \sqrt{\gamma R_a T}$ where γ is the ratio of specific heats of dry air which is 1.4 at 25°C, R_a is the specific gas constant for dry air which is calculated as $R/M = 287.2\ \mathrm{J/kg}$, and the temperature of the fluid is given as $T = 297.8\ \mathrm{K}$. The resulting speed of sound is $a = 345.6\ \mathrm{m/s}$ [96]. Then, following the CFL3D code used in the validation study, the dynamic viscosity is calculated using Sutherland's law which can be written as

$$\mu = \mu_{\text{ref}} \left(\frac{T}{T_{\text{ref}}} \right)^{3/2} \frac{T_{\text{ref}} + S}{T + S}, \tag{5.1}$$

where $T_{\text{ref}} = 273.11\ \mathrm{K}$ is a reference temperature, $\mu_{\text{ref}} = 1.716 \times 10^{-5}\ \mathrm{kg/ms}$ is the viscosity at the reference temperature, and $S = 111\ \mathrm{K}$ is Sutherland's constant [97]. The resulting dynamic viscosity is $\mu = 1.833 \times 10^{-5}\ \mathrm{kg/ms}$. Finally, the density can be calculated from the Reynolds number. The Reynolds number for the airfoil is defined as

$$Re = \frac{\rho q_{ref} c}{\mu}, \tag{5.2}$$

with $Re = 1,520,000$, $q_{ref} = 31.71\ \mathrm{m/s}$, and $c = 1$. This gives a density of $\rho = 0.8997\ \mathrm{kg/m^3}$. These values are used to specify and non-dimensionalize the problem.

The far-field outlet is specified as a simple back pressure with zero deviation from the freestream pressure. The nominal far-field inlet conditions in non-dimensional form are zero deviation from freestream pressure, freestream temperature, Mach 0.0892 for the x direction, and Mach 0.0214 for the y direction. These nominal Mach numbers are used to provide an effective angle of attack of 13.87° for the airfoil and are calculated as $Ma_x = Ma\cos(13.87^\circ)$ and $Ma_y = Ma\sin(13.87^\circ)$. Turbulence is initialized using a freestream turbulence level of 0.086% and a turbulent to laminar viscosity ratio of 0.009. The RANS shear stress transport (SST) k-ω turbulence model is used. The validation cases are run with the NASA codes CFL3D, FUN3D, and OVERFLOW while our cases are run using ICFD++ from Metacomp Technologies [95].

The grid convergence study on the five C mesh grids is used for validation of the model. The coefficients of pressure for all five grids and the experimental data are shown in Fig. 5.2. This shows that the simulation converges well using the given grids. The coarsest grid does not provide accurate results and differs from the other four grids. This is most likely due to the fact that the calculated y_1^+ values of the simulation are around 2.5 and thus wall functions are required for the simulation. The 225×65 grid also does not have a sufficiently fine grid to solve to the wall as its y_1^+ values are near 1.1. However, its results do agree better with the finer grids. The three finest grids all provide sufficient y_1^+ values to solve to the wall and show converged results. For robustness, the 897×257 grid is used in our study as in the reference validation case.

The experimental and simulation coefficient of pressure around the airfoil are compared in Fig. 5.3. The coefficients of lift (C_L) and drag (C_D) for the CFD cases are compiled in Table 5.1 and show ICFD++ compares favorably to the TMR results. The coefficients are calculated using the dimensional length of $0.9012\ \mathrm{m}$ for the airfoil chord.

For our RPC-K metamodeling study, 200 CFD simulations are run under uncertain operating conditions. The eight parameters Ma, AoA, T, Re, c, M, I, and μ_t/μ are assumed to be under normally distributed uncertainty of varying magnitude and their mean values and standard devia-

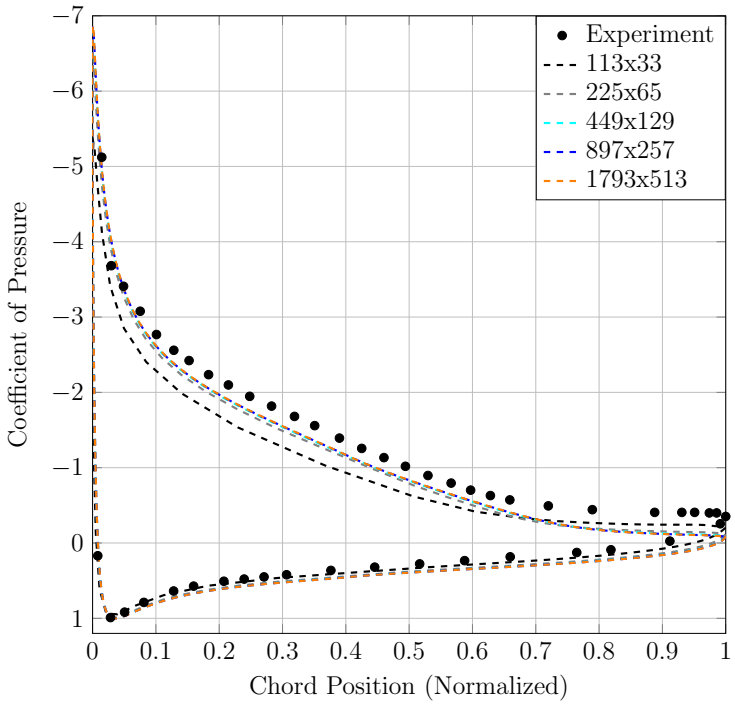


Figure 5.2: Coefficients of pressure around the NACA 4412 airfoil using five successively refined grids.

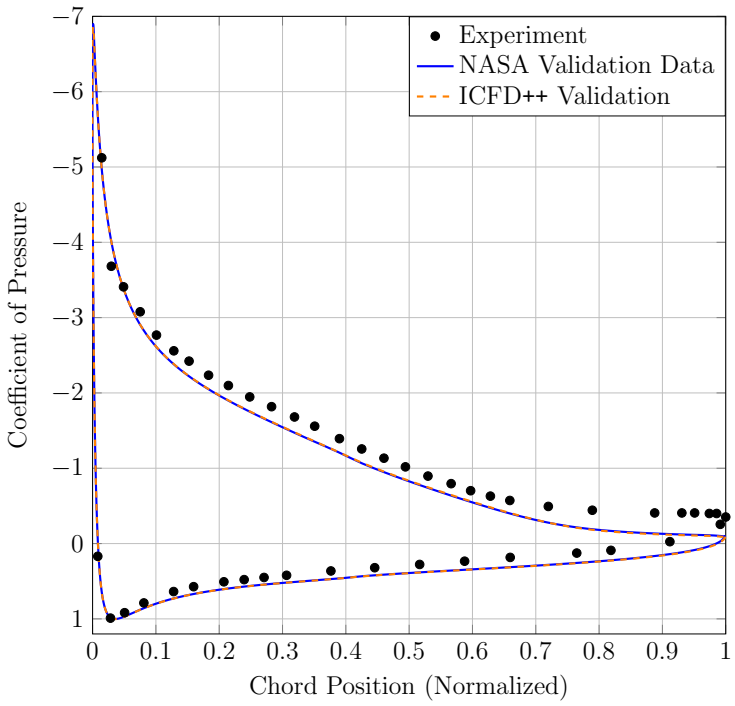


Figure 5.3: Comparison of experimental and computational results for coefficient of pressure.

Table 5.1: Coefficients of lift and drag from the validation cases of the NACA 4412 airfoil.

Case	C_L	C_D
CFL3D	1.616	0.0311
FUN3D	1.615	0.0320
OVERFLOW	1.621	0.0321
ICFD++	1.620	0.0320

Table 5.2: Parameter mean and standard deviation values for the NACA 4412 case.

Parameter	Ma	AoA($^\circ$)	T (K)	Re	c (m)	M (kg/kmol)	I	$\frac{\mu_t}{\mu}$
Mean	0.09	13.87	297.7778	1.52e6	0.9012	28.95	0.086%	0.009
Std. Dev.	0.003	0.4623	9.9259	25333	0.006	0.0096	0.0287%	0.003

tions are compiled in Table 5.2. The Sobol quasi-random sequence is used to sample the parameter values for the 200 simulations based on the parameters' statistical moments. Quasi-random sequences are effective for these types of studies as they converge to Monte Carlo results quickly since all sample points are distributed throughout the parameter space. They have the additional benefit over Latin hypercube sampling in that any additional number of points are easy to add to the sample [90].

5.2 Engine Nacelle Problem Configuration

The aircraft engine nacelle case is derived from research by Nathan Spotts [38] and Gao et al. [51]. The problem analyzes the airflow conditions of a new nacelle design for a representative aircraft engine under an extreme operating condition, takeoff under a strong crosswind. The low speed, high angle-of-attack condition can cause separation of the airflow at the fan or rotor station of the engine if the nacelle is sufficiently wide and short as is common in modern high-bypass turbofan engines. For computational simplicity, only half of the geometry is modeled and the geometry only includes the nacelle and engine hub. Figure 5.4 shows an axisymmetric schematic of the geometry with a red line indicating where profiles of the airflow are of interest. The profile lines are located at the fan station of the engine where the first rotor blade intersects the engine hub. The fan and stator blades shown in Fig. 5.4 are not modeled for computational simplicity and are shown only for clarity.

The case is solved as a steady-state RANS solution on an unstructured mesh using a second-order, upwind, finite-volume method in ICFD++ [95]. The fluid is assumed calorically perfect and the solution is sped up using Gauss-Seidel relaxation and algebraic multigrid. The Spalart-Allmaras turbulence model is used because of its wide adoption in aerodynamic flows. The far-field boundary is set with a static pressure, static temperature, and velocity based on the Mach number and angle of attack. Additionally, the mass flow rate is achieved by setting a mass flow boundary and having the solver adjust the static back pressure until the desired flow rate is achieved based on the work of Spotts [38].

The case was originally run to determine if uncertainty quantification could be conducted more quickly using the RPC metamodeling method over a full PC expansion [51]. The output variables of interest in this study are the total pressure, P_0, and the total temperature, T_0. However, only the results of total pressure are discussed for brevity due to the similarity of the temperature results. The uncertainty quantification study considers seven input parameters ($n = 7$); freestream Mach number (Ma), angle-of-attack (AoA), mass flow rate ($\dot{m}$), turbulent to laminar viscosity ratio (μ_t/μ), nacelle diameter (D), nacelle length (L), and air density (ρ). The diameter and length are handled in a non-dimensional form based on their reference lengths $D_{\text{ref}} = 0.68\,\text{m}$ at X = 0 and $L_{\text{ref}} = 0.93\,\text{m}$. These parameters are subject to uniform uncertainty of 15% of their mean. After a sensitivity study conducted by Gao et al. [51], the problem is reduced to five parameters ($\hat{n} = 5$) by setting the viscosity ratio and air density to their mean values since they contribute very little variance to the problem. The parameter's means, minimums, and maximums are shown in Table 5.3. The total pressure profiles are taken along three axes of the contour at the fan station, Y-, Z-, and Y+ shown in Fig. 5.5. Results are shown for the flow along the Y- profile only because

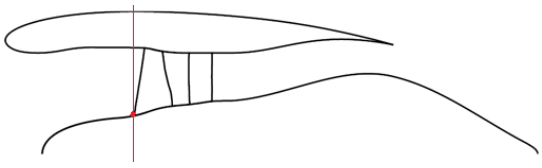


Figure 5.4: Schematic, axisymmetric view of the engine nacelle geometry with the profile location denoted by the red line.

the flow separates here under certain operating conditions and is thus the most difficult flow to construct a metamodel for due to the non-linear physics governing its behavior. In order to compare results from multiple nacelle sizes, results are plotted along a non-dimensional length, x, which starts at zero at the engine hub and ends at one at the nacelle's inner wall. The total pressure values are also non-dimensionalized against their freestream values.

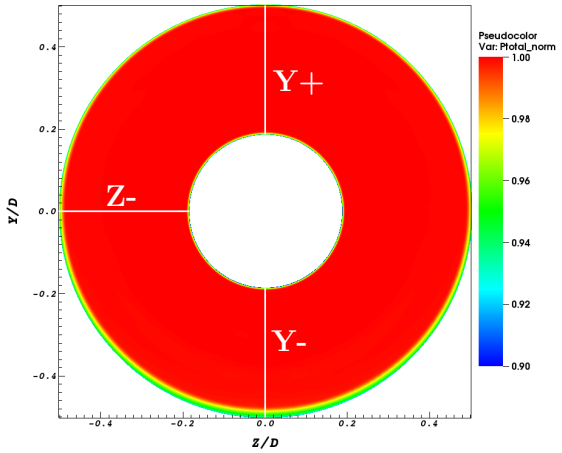


Figure 5.5: Examined profiles for engine nacelle case on a pseudocolor plot of total pressure.

Table 5.3: The mean values for the seven parameters in the engine nacelle study.

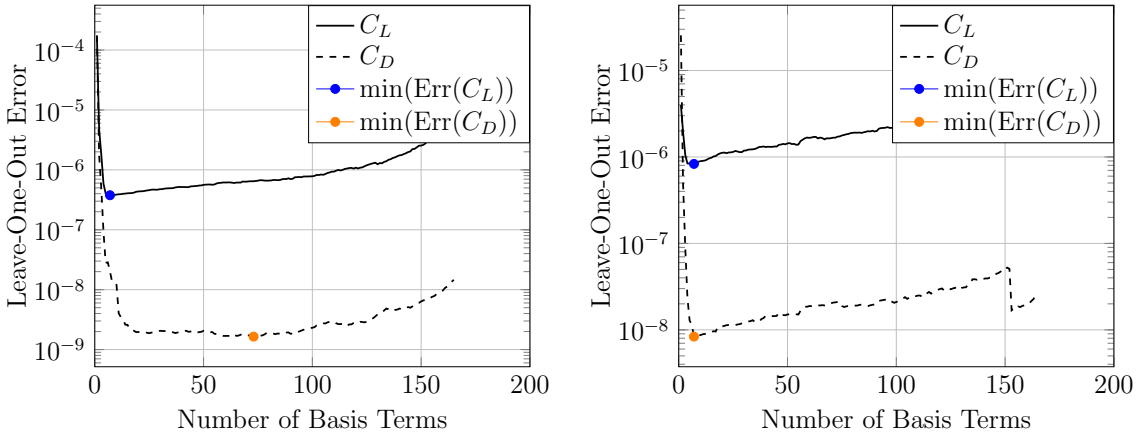
	Ma	AoA(°)	$\dot{m}$ (kg/s)	μ_t/μ	D/D_{ref}	L/L_{ref}	Density (kg/m^3)
Maximum	0.2875	29.9	27.6	4	1.15	1.15	1.225
Mean	0.25	26	24	4	1	1	1.225
Minimum	0.2125	22.1	20.4	4	0.85	0.85	1.225

5.3 Results and Discussion

5.3.1 NACA 4412 Airfoil

First, the lift and drag on the airfoil are analyzed. The LOOCV error is plotted against the number of basis terms chosen by the hierarchical LARS algorithm for a data set containing 200

points in Fig. 5.6 where Fig. 5.6a shows the PC results and Fig. 5.6b shows the PC-K results. It can be seen that only a few basis terms are needed to reduce the LOOCV error and that for the PC-K metamodels, both C_L and C_D are estimated best with only seven terms. The results are tabulated in Table 5.4 and show good agreement for both the mean, $+3\sigma$, and -3σ values for both metamodeling methods. However, it is more instructive to apply the RPC-K metamodel to a subset of the 200 CFD simulations and evaluate its performance since 200 CFD simulations can still be viewed as a large number.



(a) Polynomial chaos metamodel.

(b) Polynomial chaos-Kriging metamodel.

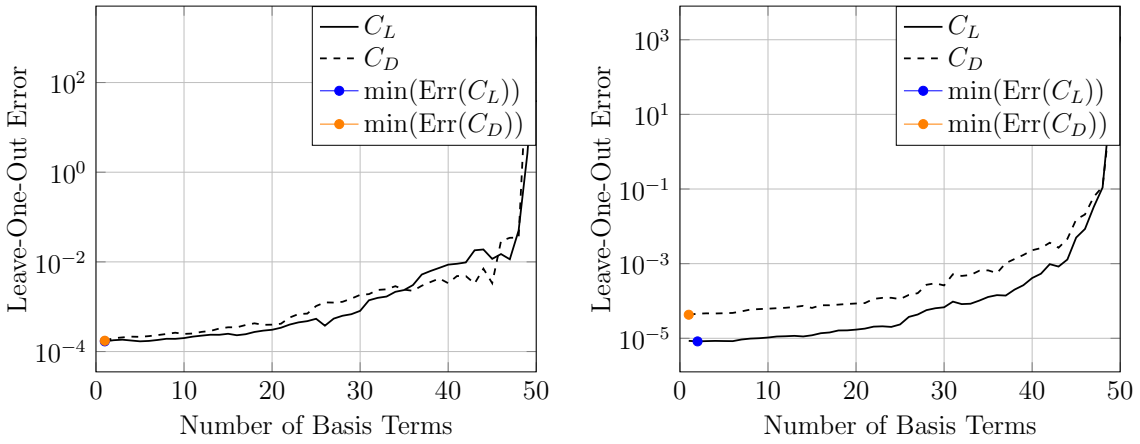
Figure 5.6: LOOCV error progression of C_L and C_D with 200 simulations.

Table 5.4: Mean and standard deviations for C_L and C_D using metamodels on a 200-simulation set.

Model	200 Simulations	PC	PC-K
Mean C_L	1.6168	1.6168	1.6167
$C_L + 3\sigma$	1.6662	1.6665	1.6664
$C_L - 3\sigma$	1.5675	1.5671	1.5670
Mean C_D	0.0323	0.0323	0.0323
$C_D + 3\sigma$	0.0395	0.0395	0.0395
$C_D - 3\sigma$	0.0251	0.0252	0.0252

It can be seen in Fig. 5.7 that both the PC and PC-K metamodels do not fit the data well when only using 50 CFD simulations. In fact, only the PC-K metamodel for C_L has a basis of more

than the mean value and it only selects three terms. Analyzing the tabulated data shows that the resulting models are not effective as they simply return the mean value with no standard deviation. It should be noted that the PC-K metamodel reverts back to an ordinary Kriging metamodel and still produces reasonably accurate results if a small Monte Carlo simulation is run on it. The statistical data are summarized in Table 5.5.



(a) Polynomial chaos metamodel. **(b)** Polynomial chaos-Kriging metamodel.

Figure 5.7: LOOCV error progression of C_L and C_D with 50 simulations.

Table 5.5: Mean and standard deviations for C_L and C_D using metamodels on a 50-simulation set.

Model	200 Simulations	PC	PC-K	PC-K Monte Carlo
Mean C_L	1.6168	1.6164	1.6167	1.6174
$C_L + 3\sigma$	1.6662	1.6164	1.6181	1.6615
$C_L - 3\sigma$	1.5675	1.6164	1.6152	1.5732
Mean C_D	0.0323	0.0323	0.0327	0.0324
$C_D + 3\sigma$	0.0395	0.0323	0.0327	0.0373
$C_D - 3\sigma$	0.0251	0.0323	0.0327	0.0275

This apparent failure of the methods can be explained by the curse of dimensionality. When working with larger parameter sets, even space-filling sampling methods only sample the boundary of the problem due to the curse of dimensionality. This issue is exacerbated with small data sets. Least squares regression fits the data based on minimizing the mean squared error of the problem

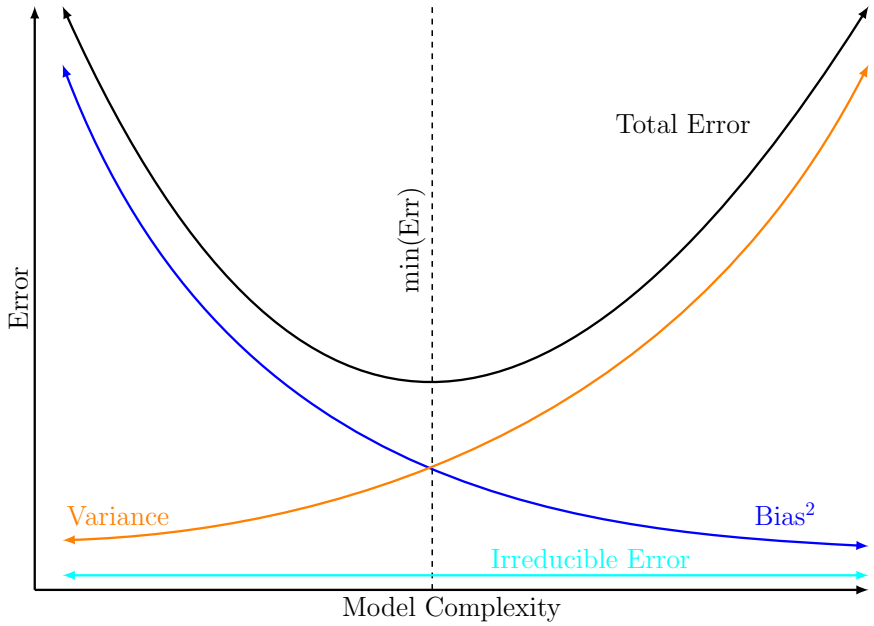


Figure 5.8: Bias-variance decomposition of mean squared error.

which can be decomposed using the bias-variance decomposition. This states

$$\text{Err}(x) = \sigma_\epsilon^2 + \text{Bias}^2(\hat{f}(x)) + \text{Var}(\hat{f}(x)), \tag{5.3}$$

or that the mean squared error is composed of bias error, variance error, and irreducible model error [93]. This is also illustrated in Fig. 5.8. Under-fitted models have a high bias error and over-fitted models have a high variance error. Normally, least squares regression selects a low error solution with minimal bias and variance error. In small data set, higher-parameter models, the bias error is not easily reduced by incorporating additional basis terms and the variance error grows rapidly. Thus, the least error solution is the one with just a mean term. This problem can not be fixed using a different error measure such as Akaike information criterion (AIC), Bayesian information criterion (BIC), or the coefficient of determination (R^2) as they are all based off of the mean squared error as well. However, the problem can be combated by manually selecting a meta-model from inspection, increasing the number of samples, or decreasing the number of parameters. Nevertheless, manually fitting the model is not only time consuming but requires some knowledge of the solution and the approximate statistical moments. This makes the method inadvisable. Fur-

ther, with the computational cost and additional time required to run additional CFD simulations, increasing the number of samples is generally impossible in practical application. This leaves reducing the number of parameters as a viable solution.

The number of parameters in the airfoil problem are reduced using a sensitivity study. Sobol sensitivity indices are constructed from the coefficients of a PC model with only main effects up to third-order. The sensitivity indices are given as

$$S_j = \frac{\sigma_j^2}{\sigma_u^2} = \frac{\sum_{i=1}^{m} (\hat{U}_{ki}^j)^2}{\sum_{j=1}^{n} \sum_{i=1}^{m} (\hat{U}_{ki}^{j2})}, \tag{5.4}$$

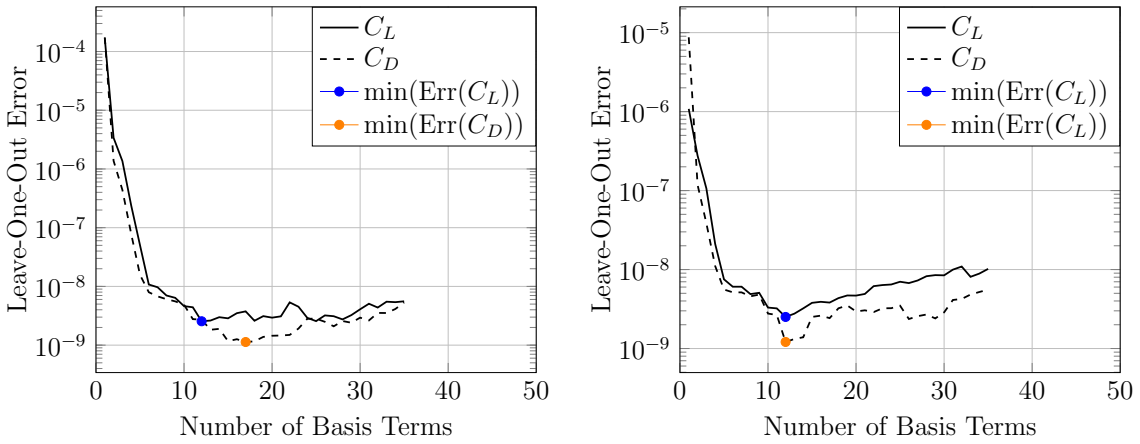
where $\sigma_u^2 \approx \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$ [51, 76, 82, 98]. This way the sensitivity indices add up to 1 where S_j is the estimated variance in the output due to parameter j. For the airfoil case, the coefficients are fit to the 50 airfoil cases with a basis including only main effects up to degree three for each parameter. Using the study, it can be seen that the angle of attack dominated the problem, and the problem can easily be reduced to the same four parameters for both the C_L and C_D as they show an order of magnitude more effect than the others. These indices are compiled in Table 5.6 and show that AoA, Ma, c, and Re are the retained parameters.

Table 5.6: Sensitivity indices for the airfoil parameters.

Parameter	Ma	AoA	T	Re	c	M	I	$\frac{\mu_t}{\mu}$
S_j for C_L	2.25e-4	0.9924	1.80e-5	0.0063	0.0011	1.45e-6	1.79e-6	7.31e-6
S_j for C_D	3.94e-5	0.9978	1.07e-5	0.0018	3.17e-4	1.44e-6	9.97e-7	5.44e-6

The new RPC and RPC-K metamodels are then fit to the 50 data set airfoil problem and the LOOCV errors are shown in Fig. 5.9. It can be seen that now the error is initially high and decreases with an increasing number of basis terms, as desired. The models generally are optimal at 12 basis terms, but the RPC metamodel for the C_D is optimal at 17 terms. The statistics of these improved models is shown in Table 5.7 which shows the dramatic improvement in the metamodels selected by the algorithm. The statistics of the RPC-K metamodel for C_L differ from those of the RPC model and that of the 200 cases. This can be explained by the fact that the RPC-K metamodel fits

the polynomial coefficients with correlation data, so it slightly skews the polynomial model. The resulting model can generally give more accurate predictions for individual points at the cost of the accuracy of the polynomial coefficients. A small Monte Carlo simulation of the C_L data results in similar standard deviation estimates for C_L.



(a) Reduced polynomial chaos metamodel.

(b) Reduced polynomial chaos-Kriging metamodel.

Figure 5.9: LOOCV error progression of C_L and C_D with 50 simulations on a reduced parameter set.

Table 5.7: Mean and standard deviations for C_L and C_D using metamodels on a 50-simulation set and reduced parameter set.

Model	200 Simulations	RPC	RPC-K	RPC-K Monte Carlo
Mean C_L	1.6168	1.6169	1.6168	1.6169
$C_L + 3\sigma$	1.6662	1.6668	1.6682	1.6661
$C_L - 3\sigma$	1.5675	1.5669	1.5655	1.5676
Mean C_D	0.0323	0.0323	0.0323	0.0323
$C_D + 3\sigma$	0.0395	0.0395	0.0395	0.0394
$C_D - 3\sigma$	0.0251	0.0252	0.0252	0.0253

The final RPC and RPC-K metamodels for the airfoil case demonstrate how the RPC-K metamodel can be applied to aerodynamics problems by utilizing sensitivity indices and LARS model selection to produce the best PC and PC-K metamodels for a resource-constrained problem.

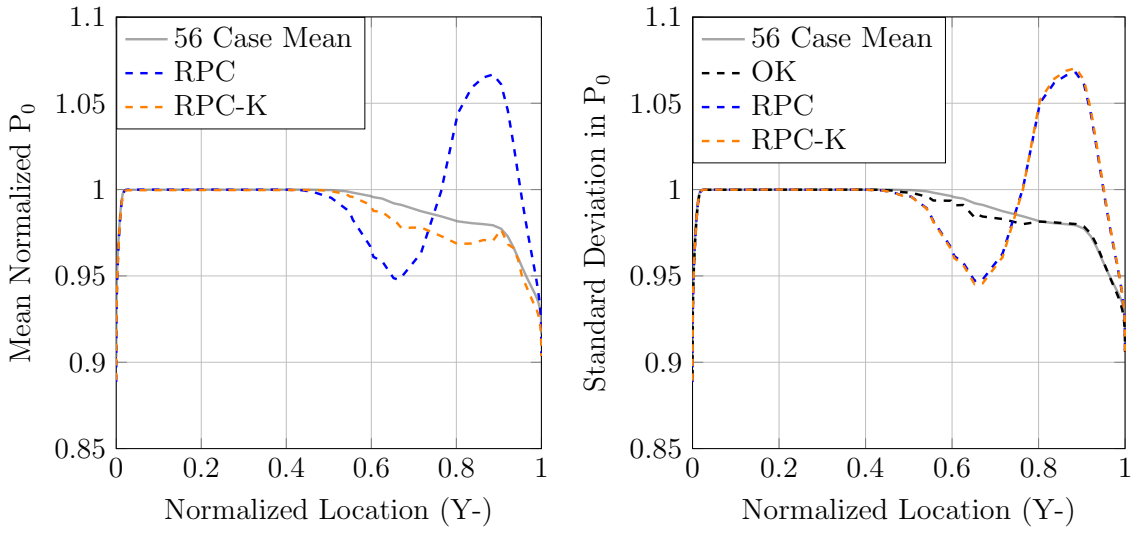
5.3.2 Aircraft Engine Nacelle

In order to make sure the results are consistent with the original RPC-K research [51], the PC expansion are fit to the five parameters Ma, AoA, $\dot{m}$, D/D_{ref}, and L/L_{ref} with a maximum exponent of degree three to give a total of 56 basis terms. Also to maintain consistency, 56 simulations are run to fit the models; however, Latin hypercube sampling is used over Gaussian quadrature techniques so that accurate correlation data can be calculated. Early research [78,79] for the aircraft engine nacelle showed a discrepancy between the results achieved using a correctly-formulated PC expansion versus an expansion which uses non-standardized variables for the RPC-K metamodel. The standardized variables are calculated as

$$\boldsymbol{\omega}_{\text{Standard}} = (\boldsymbol{\omega} - \overline{\boldsymbol{\omega}})/\boldsymbol{\sigma} \tag{5.5}$$

where $\overline{\boldsymbol{\omega}}$ is the vector of means for $\boldsymbol{\omega}$ and $\boldsymbol{\sigma}$ is the vector of standard deviations. They have a mean of zero and equal variation magnitudes. Using these polynomial expansions as the basis for the RPC-K metamodel constructed on the total pressure data for the engine nacelle data gives the results seen in Fig. 5.10. Also plotted are the results from using the ordinary Kriging metamodel. These results indicate that if the RPC metamodel is over-fit, just adding it to the more complex RPC-K metamodel and recalculating the coefficients with correlation information does not prevent over-fitting. This is because the over-fit PC expansion perfectly predicts every design point, so the Kriging deviation term $Z(\boldsymbol{\omega})$ has a value of zero. Unfortunately, the off-design points are still poorly predicted by the over-fit function.

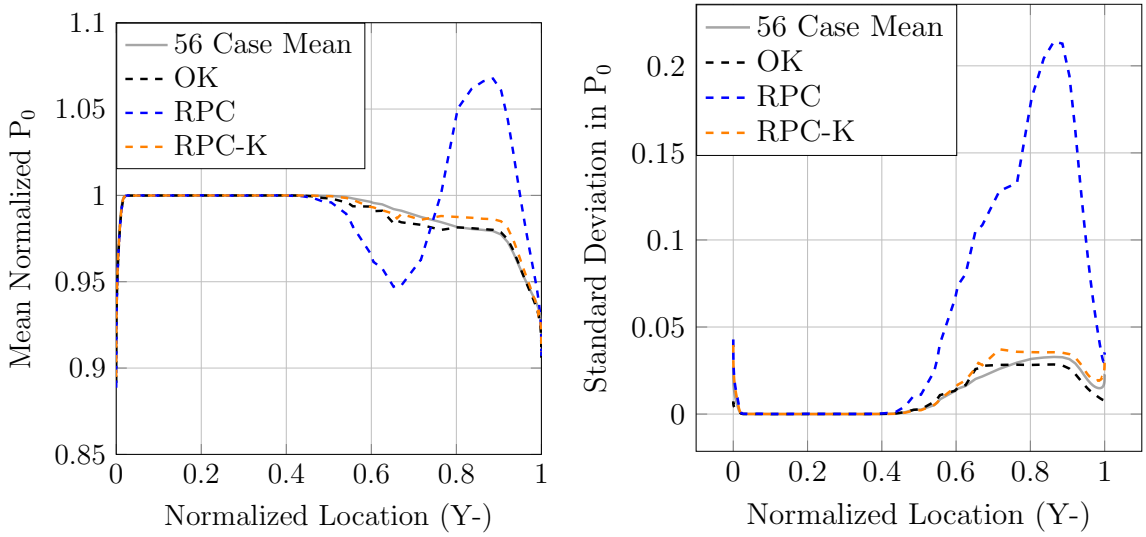
The non-standardized PC expansion provides a superior fit to the function because the PC expansion is not capable of perfectly fitting every design point and so has an active deviation term. However, this is not a suitable solution because it will not necessarily provide sufficient results in all cases. Importantly, if the true values of variables are used, poor results can be seen when a parameter has much larger values than others because it dominates the polynomial expansion. Additionally, many advantages of PC expansions can not be exploited. Still, the research repeatedly



(a) Non-standardized Variable Results **(b)** Standard Random Variable Results

Figure 5.10: Comparison of non-standardized and standard variable results.

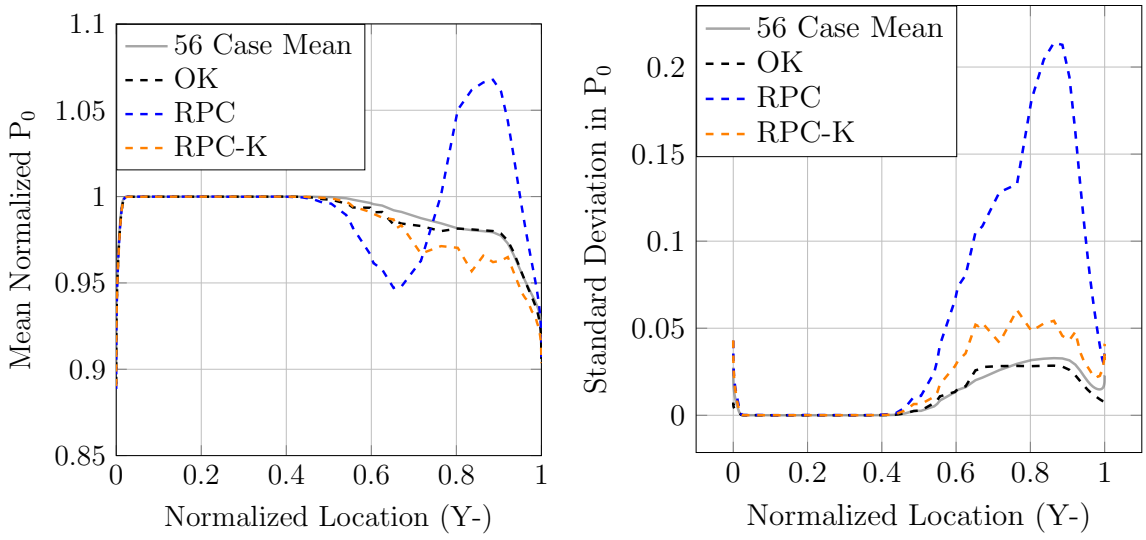
shows that PC expansions may not be the ideal polynomial expansion for sparse problems as they only guarantee convergence to the true solution as additional basis terms are added.



(a) 1st Degree Mean **(b)** 1st Degree Standard Deviation

Figure 5.11: Comparison of over-fitting solutions using first-order polynomials.

Regardless of the fact that PC expansions may not be the best global polynomial forms for sparsely sampled problems, they are utilized in this thesis research due to the lack of any systematic



(a) 2nd Degree Mean

(b) 2nd Degree Standard Deviation

Figure 5.12: Comparison of over-fitting solutions using second-order polynomials.

way to determine a better form for a particular problem and their advantages for sensitivity analysis. Using the PC form, multiple techniques to overcome the over-fitting issue are explored and it is found that a reduction in terms is the only viable solution [79]. This is illustrated in Fig. 5.11 and Fig. 5.12 which compare the results of using polynomials limited to a maximum of 1st and 2nd degree terms versus the 3rd order limit as shown in Fig. 5.10b. It can be seen that the greater over-sampling for the smaller polynomial bases prevents the over-fitting issues previously seen. However, the accuracy of the 1st degree expansion is not sufficient and that there are spurious oscillations in the data for the 2nd degree polynomials. These are especially pronounced in the standard deviation data. This indicates that a systematic way to determine an ideal number of terms which likely lie between a full first-order and second-order expansion is needed. This is accomplished using the LARS algorithm and the LOOCV error.

The leave-one-out error for varying basis sizes of both the original RPC and RPC-K methods on the engine nacelle problem are shown in Fig. 5.13. Again, these models are fit to the five parameters Ma, AoA, $\dot{m}$, D/D_{ref}, and L/L_{ref} with a maximum exponent of degree three for a total of 56 basis terms using 56 data points. It can be seen that the RPC metamodel has minimum error with 11 basis terms and the RPC-K metamodel with 11 as well. The behavior of the plot

shows that the model is likely under-fitting the data because there is no large drop in the error when initial basis terms are added.

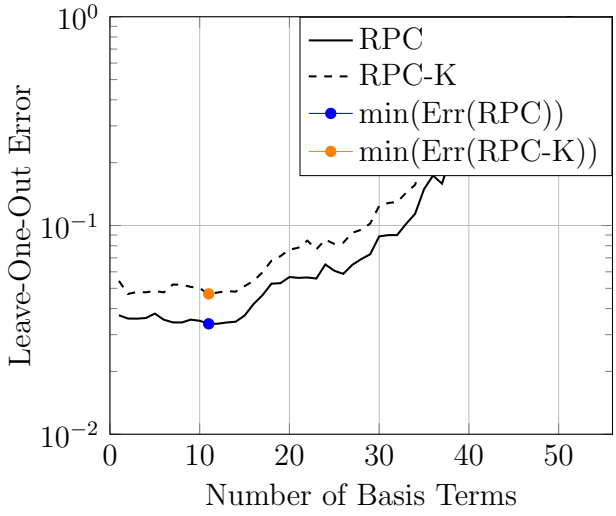


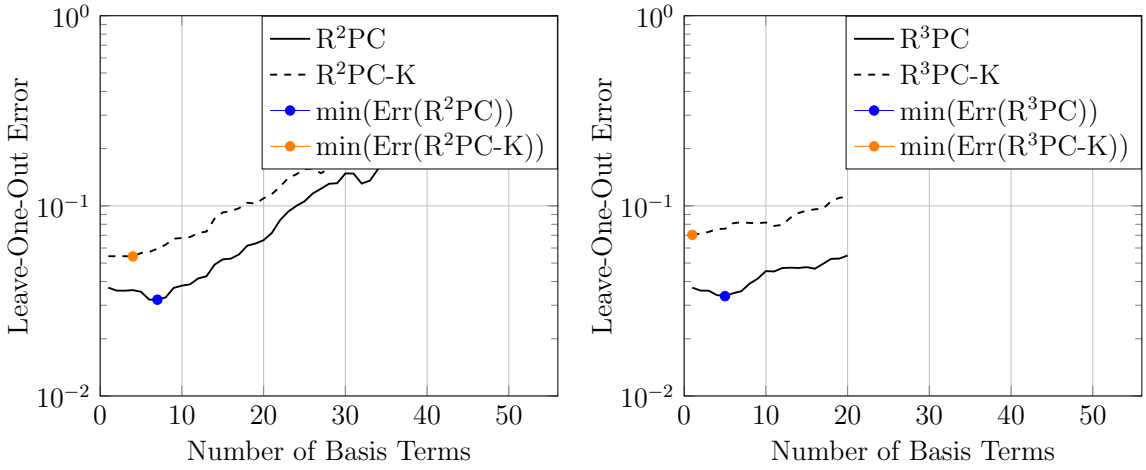
Figure 5.13: LOOCV error progression for RPC and RPC-K metamodels using hierarchical LARS algorithm to choose basis terms.

To combat this under-fitting, a further sensitivity study is carried out on the RPC-K metamodel to remove non-essential parameters. The sensitivity indices are listed in Table 5.8 for both the whole domain and the region of interest, the separated flow. It can be seen that the relative impact of the parameters varies from the smooth flow regions to the separated flow region, but the Mach number is always the least impactful dimension. As such, the twice reduced problem removes Ma from the parameter list and the resulting metamodels are indicated as R^{2}PC and R^{2}PC-K since they have a twice reduced parameter space. Then the problem is further reduced by removing $\dot{m}$ from the parameter space as it is the next least impactful parameter in the region of interest. The metamodels constructed from this small parameter set are indicated as R^{3}PC and R^{3}PC-K and are used to measure the efficacy of very small parameter sets.

Table 5.8: Sensitivity indices for the engine nacelle parameters.

Parameter	Ma	AoA	$\dot{m}$	D/D_{ref}	L/L_{ref}
S_j Total	0.0742	0.1784	0.2689	0.2984	0.1802
S_j Separation	0.0426	0.2550	0.1642	0.2837	0.2544

The LOOCV error progressions for the R^{2}PC and R^{2}PC-K metamodels are shown in Fig. 5.14a and the error progressions for the R^{3}PC and R^{3}PC-K metamodels are shown in Fig. 5.14b. As shown, the R^{2}PC and R^{3}PC metamodels do demonstrate some of the characteristic drop in error when the initial basis terms are added while the R^{2}PC-K and R^{3}PC-K metamodels do not. Altogether, reducing the number of parameters does not necessarily prevent under-fitting. Additionally, all of the models have lower LOOCV error with progressively fewer terms as the parameter space is reduced. The R^{2}PC metamodel has a minimum error with seven terms, the R^{3}PC with five terms, the R^{2}PC-K metamodel with four terms, and the R^{3}PC-K with one. The minimum errors for the metamodels are provided in Table 5.9 and show that all of the RPC metamodels have similar errors while the errors for the RPC-K metamodels steadily increases when parameters are removed.



(a) RPC and RPC-K metamodels using four parameters.

(b) RPC and RPC-K metamodels using three parameters.

Figure 5.14: LOOCV error progression for further reduced RPC-K metamodels for the engine nacelle problem.

Table 5.9: LOOCV error for the engine nacelle metamodels.

Model	RPC	RPC-K	R^{2}PC	R^{2}PC-K	R^{3}PC	R^{3}PC-K
min(Err)	3.38×10^{-2}	4.70×10^{-2}	3.21×10^{-2}	5.41×10^{-2}	3.35×10^{-2}	7.04×10^{-2}
Terms	11	11	7	4	5	1

The training errors for a metamodel can also be analyzed to determine how well it is fitting the data. An under-fit model has large training errors and an over-fit model has almost zero error. The training errors for the RPC, R^2PC, and R^3PC metamodels are shown in Fig. 5.15. It shows that the training error stays quite large for the RPC metamodels until the model ultimately becomes over-fitted when the number of terms is close to the number of data points. This indicates the models are not likely very accurate. Meanwhile, the training errors for the RPC-K metamodels are zero as Kriging metamodels interpolate data points. While this accuracy can not be extended to unknown data points, it does suggest that the RPC-K models are likely more accurate based on the additional correlation information for data points that are close in value to the known data points.

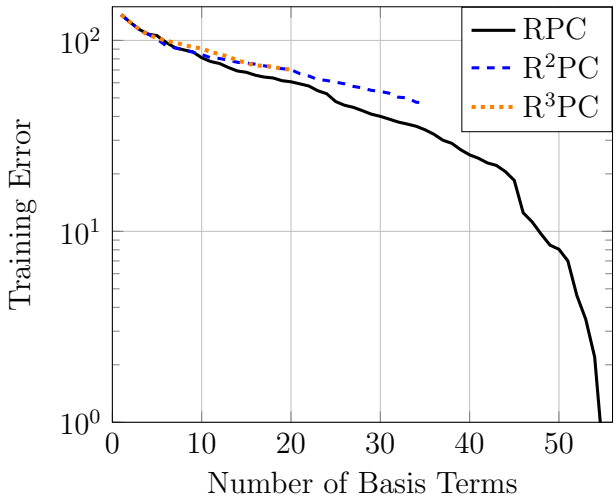
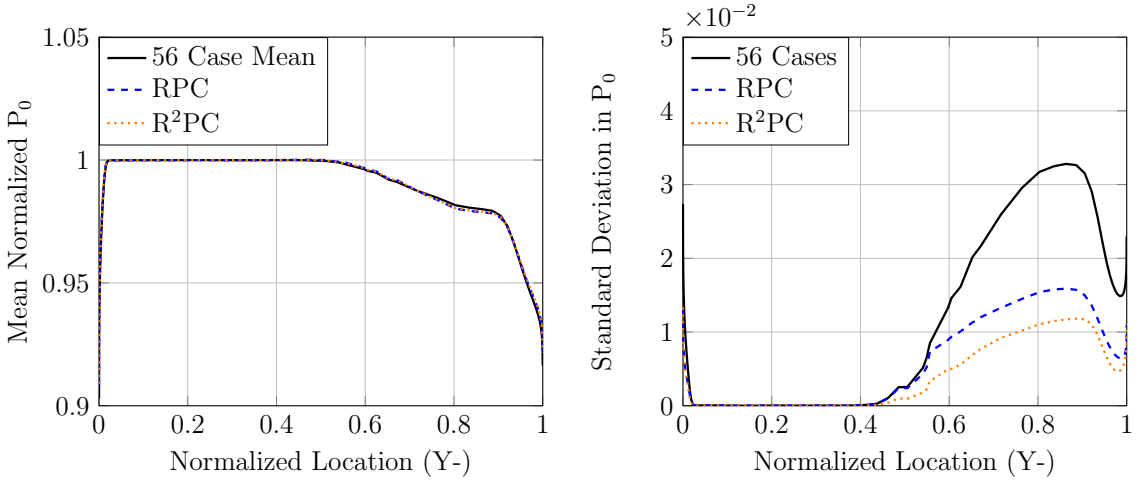


Figure 5.15: Training error for the RPC, R^2PC, and R^3PC metamodels.

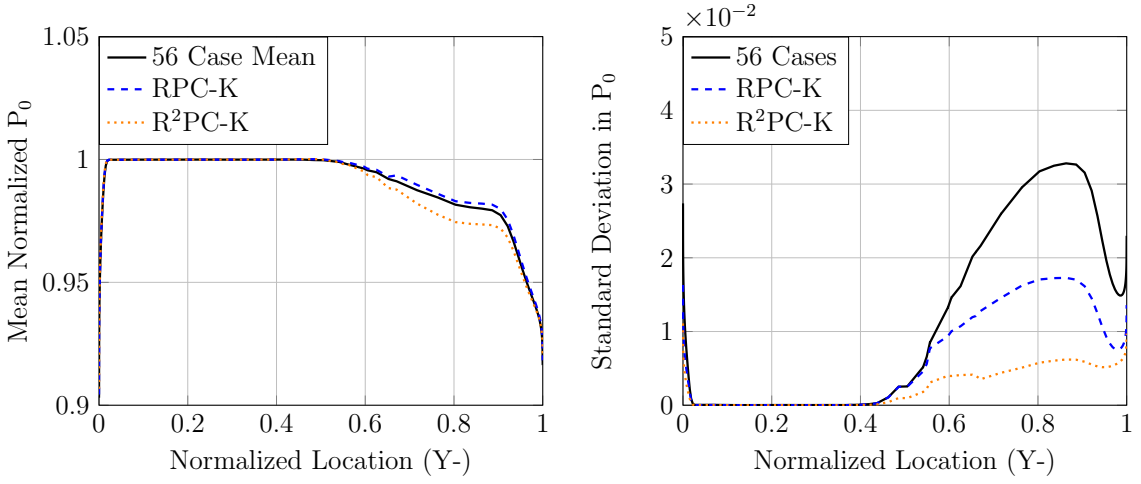
The resulting data of these metamodels can then be plotted against the data from the 56 full evaluations. The total pressure is plotted along the normalized Y- profile where zero is the hub wall and one is the nacelle wall. The mean and standard deviations are plotted to provide an estimate of the problem mean and realistic deviations. Critically, because only 56 cases were run due to computational cost, this data may differ from the true values. The results of the RPC and R^2PC metamodels are shown in Fig. 5.16 while the results of the RPC-K and RPC-K metamodels are

shown in Fig. 5.17. The R^{3}PC and R^{3}PC-K metamodels are not compared due to their inferior error measures.



(a) Mean P_0 for RPC and R^{2}PC. **(b)** Standard deviation in P_0 for RPC and R^{2}PC.

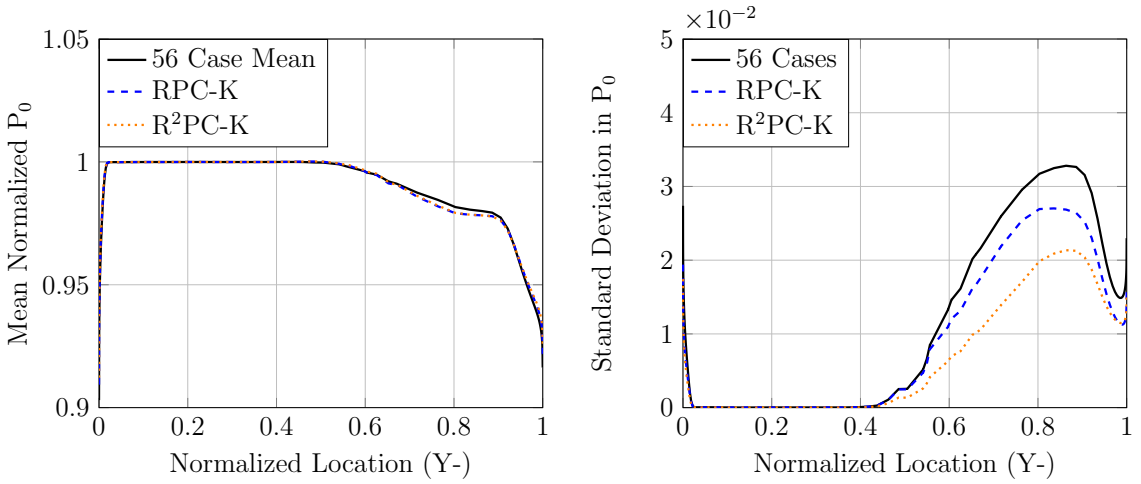
Figure 5.16: Mean and standard deviation of P_0 on the Y- profile of the engine nacelle problem for PC metamodels.



(a) Mean P_0 for RPC-K and R^{2}PC-K. **(b)** Standard deviation in P_0 for RPC-K and R^{2}PC-K.

Figure 5.17: Mean and standard deviation of P_0 on the Y- profile of the engine nacelle problem for PC-K metamodels.

It can be seen that the mean of the data is well approximated by all of the models; however, the standard deviations of the data are greatly under-predicted. At best, the RPC and RPC-K metamodels predict about half of the standard deviation seen in the data. This difference is very significant and means that the variance of the metamodels chosen by the LOOCV error is too small. However, knowing that the standard deviation values given by the polynomial trend of the RPC-K metamodels may not match its Monte Carlo results based on the airfoil results, a Monte Carlo study is conducted on both the RPC-K and R^2PC-K metamodels. These results are shown in Fig. 5.18. Again, the standard deviation values of the Monte Carlo study of the RPC-K metamodel are superior to those of the RPC metamodel and the statistical moments given by the polynomial coefficients of the RPC-K metamodel. The average mean squared errors of the RPC, RPC-K, and Monte Carlo study of the RPC-K metamodel compared to the 56 cases are compiled in Table 5.10.



(a) Mean P_0 for RPC-K and R^2PC-K.

(b) Standard deviation in P_0 for RPC-K and R^2PC-K.

Figure 5.18: Mean and standard deviation of P_0 on the Y- profile of the engine nacelle problem calculated from a Monte Carlo study of the PC-K metamodels.

It can be seen that the error difference of the mean values for all of the metamodels is very small and is superior for the models with more terms in the polynomial expansion. This makes sense as these models should have a lower percentage of bias error based on the bias-variance

Table 5.10: LOOCV error for the engine nacelle metamodels.

Model	RPC	R^2PC	RPC-K	R^2PC-K	RPC-K MC	R^2PC-K MC
Mean MSE	5.79e-6	2.16e-5	1.58e-6	7.37e-5	9.25e-6e	1.89e-5
σ MSE	2.39e-1	2.64e-1	1.80e-1	3.80e-1	6.70e-2	1.40e-1

decomposition. All of the metamodels have more error in the standard deviation estimates, though the metamodels fit to the larger number of parameters have lower errors. This is likely due to the fact that again these models have more terms in the basis which tends to increase the variance in the resulting models. Additionally, the PC-K metamodels have lower error estimates when using Monte Carlo data with the Monte Carlo results of the RPC-K metamodel showing far superior results than any other metamodel. This again shows the superiority of the PC-K type metamodels over standard PC expansions when fitting these sparse problems due to the additional correlation data.

Chapter 6

Conclusions and Future Work

The RPC-K has been developed and applied to UQ in the CFD simulations of aerodynamics problems. Conclusions and directions for future work are summarized.

6.1 Conclusion

Our work shows that the RPC-K method can be effectively used to help conduct UQ for CFD analysis of aerodynamics problems as an accurate metamodel. This metamodel can be used to efficiently generate data in the presence of uncertainty or estimate the response of off-design points. This thesis research has also extended the earlier research into how PC-K methods can be adapted for very expensive non-linear fluid dynamics problems through reducing parameter spaces and efficiently using LOOCV error along with LARS to select sparse polynomial expansions. This process is not straightforward because it is not robust enough to be fully automated for strongly non-linear fluid dynamics problems. The best solution to combat this issue is still a good sampling method and a sufficient number of simulation evaluations. Several tools are also used to help combat this. A model selection algorithm, such as the hierarchical LARS algorithm, is effective in providing good sparse PC expansions. These expansions can capture a large portion of the problem variability and produce physically realistic models based on the heredity of terms included. Estimating the generalization error of the model using LOOCV is very efficient, although it may lead to systematic under-fitting of a model if the problem is sparsely sampled in a high-dimensional space. This can be circumnavigated by manually inspecting the problem, but the additional time cost and loss of automatization negates many of the benefits of using a metamodel. Fundamentally, this issue is due to using mean squared error measures to construct a model which have a bias-variance decomposition. This problem can be alleviated by reducing the model to fewer parameters which can more robustly fit the data with only a minimal loss in accuracy. The resulting RPC-K metamodels are superior to their higher-dimensional forms because of this robustness. Additionally, the

RPC-K metamodel can be superior to RPC methods because a Monte Carlo analysis of the metamodel produces more accurate data based on the additional correlation information and resulting variance.

For an example of the computational savings engineers can expect, in an example study of an aircraft wing configuration with ten input parameters fit using third order polynomials, a CFD practitioner will need to run a minimum of 572 CFD simulations. While far less demanding than the cost of a quasi-Monte Carlo study, this is still a significant computational cost. Instead, a sensitivity study can be run from an initial sample of 62 CFD simulations. If this study finds that only half of the parameters contribute more than 2% of the total sensitivity (realistic based on our studies), the CFD practitioner can fit the same polynomial metamodel with only 112 CFD simulations. Thus, for 20% of the total computational cost, the metamodel provides results accurate to within roughly 5% of the full study. Further, if the problem shows significant non-linear effects, the RPC-K metamodel will fit the data better than low-order PC expansions can. This extends the usefulness of PC metamodels for sparsely sampled, non-linear flow problems due to the RPC-K metamodel's ability to fit non-linear data more robustly. Ultimately, the RPC-K metamodel will help facilitate practical engineering designs by providing CFD practitioners with statistical data from computational designs efficiently.

6.2 Future Work

Additional research on PC-K metamodels should focus on making the method more robust. This is especially true for the RPC-K metamodel and all applications to sparsely sampled problems. Currently, the most fragile part of the metamodel is fitting the PC trend to the data by estimating the coefficients. As explained in the engine nacelle results, more ideal polynomial bases may exist for sparsely sampled problems; however, it is unlikely there is a global solution. Because of this, PC expansions are assumed to be the best form in general. Linear least squares regression works well for problems with few outliers, generally linear behavior, and sufficient over-sampling of the problem. The computer experiments conducted using CFD often do not match these idealized

cases. One possible method could be to use non-linear basis functions and non-linear regression methods to more accurately model the non-linear effects of the model. Unfortunately, this type of methodology will no longer use PC expansions. Another possible solution is to transform the problem parameters using principal component analysis, partial least squares regression, active subspaces, or a similar variable projection method. When used, these methods could make it easier for the PC expansion to fit the model to more "ideal" parameters which are a linear combination of the natural problem parameters. This can be an improvement in case the problem parameters are accidentally correlated, as the new parameters are not. Also, the variance of the principal components is maximized and so the problem can likely be more accurately estimated from fewer parameters. These methods, though, still address non-linear effects with linear models.

The RPC-K metamodel can also be improved if robust regression methods are used over ordinary least squares. These methods are specifically designed to be more robust to outliers. Additionally, the output responses can also be inspected for outliers and non-linear effects. If they exist in the data set, the metamodel can be fit to only the features which it is capable of fitting based on the sparse sampling of the problem. While the resulting trend function may not fit all of the problem physics, it does more robustly represent the important physics of the problem. Unfortunately, this manual inspection of the data does not help automate the metamodeling process. Additionally, any derivative or uncertainty obtained from an intrusive method could also be used by a RPC-K metamodel to improve its ability to fit these high-dimensional, sparsely-sampled problems. As improvements are made, additional complex validation cases will help the technology mature. Fundamentally, all of these methods have the potential to help improve the RPC-K metamodel for CFD practitioners and engineers in general.

Bibliography

[1] Christopher Rumsey. 2DN44: 2D NACA 4412 airfoil trailing edge separation. Online, March 2018.

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

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

[4] T. W. Simpson, J. D. Peplinski, P. N. Koch, and J. K. Allen. Metamodels for computer-based engineering design: Survey and recommendations. *Engineering with Computers*, 17:129–150, 2001.

[5] G. G. Wang and S. Shan. Review of metamodeling techniques in support of engineering design optimization. *ASME Transactions, Journal of Mechanical Design*, 129(4):370–380, April 2007.

[6] B. Sudret. Meta-models for structural reliability and uncertainty quantification. In *Fifth Asian-Pacific Symposium on Structural Reliability and its Applications*, 2012.

[7] Felipe A. C. Viana, Timothy W. Simpson, Vladimir Balabanov, and Vasilli Toropov. Metamodeling in multidisciplinary design optimization: How far have we really come? *AIAA Journal*, 52(4):670–690, April 2014.

[8] A. K. Prasad and S. Roy. Global sensitivity based dimension reduction for fast variability analysis of nonlinear circuits. *24th IEEE Conference on Electrical Performance of Electronic Packages*, pages 97–99, October 2015. in Proc.

[9] L. L. Gratiet, S. Marelli, and B. Sudret. Metamodel-based sensitivity analysis: Polynomial chaos expansions and gaussian processes. Technical report, ETHZurich, 2016.

[10] R. Schöbi. *Surrogate models for uncertianty quantification in the context of imprecise probability modeling*. PhD thesis, ETH Zurich, 2017.

[11] Dimitri Mavriplis. High performance computational engineering: Putting the 'E' back in 'CSE'. In *Proceedings of the 21st International Conference on Parallel Computational Fluid Dynamics*, 2009.

[12] Jeffrey Slotnick, Abdollah Khodadoust, Juan Alonso, David Darmofal, William Gropp, Elizabeth Lurie, and Dimitri Mavriplis. CFD vision 2030 study: A path to revolutionary computational aerosciences. Technical Report NASA/CR-2014-218178, NASA, 2014.

[13] S. M. Guzik, X. Gao, and C. Olschanowsky. A high-performance finite-volume algorithm for solving partial differential equations governing compressible viscous flows on structured grids. *Comput. Math Appl.*, 72:2098–2118, 2016.

[14] H. N. Najm. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. *Annual Review of Fluid Mechanics*, 41:35–52, 2009.

[15] John Faragher. Probabilistic methods for the qunatification of uncertainty and error in computational fluid dynamics simulations. Technical Report DSTO-TR-1633, Defence Science and Technology Organisation, Department of Defense, Australian Government, 2004.

[16] Bruno Sudret and Armen Der Kiureghian. Stochastic finite element methods and reliability: A state-of-the-art report. Technical Report UCB/SEMM-2000/08, University of California, Berkeley, 2000.

[17] Sharon C. Glotzer, Sangtae Kim, Peter T. Cummings, Abhijit Deshmukh, Martin Head-Gordon, George Karniadakis, Linda Petzold, Celeste Sagui, and Masanobu Shinozuka. International assesment of research and development in simulation-based engineering and science. Technical report, World Technology Evaluation Center, 2009.

[18] J. Tinsley Oden, Ted Belytschko, Jacob Fish, Thomas J. R. Hughes, Chris Johnson, David Keyes, Alan Laub, Linda Petzold, David Srolovitz, and Sidney Yip. Simulation-based engineering science. Technical report, National Science Foundation, May 2006.

[19] Joaquim R. R. A. Martins and Andrew B. Lambe. Multidisciplinary design optimization: A survey of architectures. *AIAA Journal*, 51(9):2049–2075, 2013.

[20] Michael Meinel and Gunnar Ólafur Einarsson. The flowsimulator framework for massively parallel cfd applications. In *PARA 2010*, 2010.

[21] Yu-Chi Ho and David L. Pepyne. Simple explanation of the no free lunch theorem of optimization. In *Proceedings of the 40th IEEE Conference of Decision and Control*, Orlando, Florida, 2001. IEEE.

[22] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13:455–492, 1998.

[23] A. J. Booker, J. E. Dennis Jr., P. D. Frank, D. B. Serafini, V. Torczon, and M. W. Trosset. A rigorous framework for optimization of expensive functions by surrogates. *Structural and Multidisciplinary Optimization*, 17(1):1–13, 1999.

[24] L. Bajer and Martin HoleÅĹa. Rbf-based surrogate model for evolutionary optimization, 2011.

[25] R. Schöbi, P. Kersaudy, B. Sudret, and J. Wiart. Combining polynomial chaos expansions and kriging. Technical Report RSUQ-2014-001, ETH Zurich, 2014.

[26] J. Barche. Experimental data base for computer program assesment. Technical Report AGARD-AR-138, Advisory Group for Aerospace Research and Development, 1979.

[27] R. Schöbi and B. Sudret. PC-kriging: A new metamodelling method combining polynomial chaos expansions and kriging. In *Proceedings of the 2nd International Symposium on uncertainty Quantification and Stochastic Modeling*, Rouen, France, June 2014.

[28] Roland Schöbi, Bruno Sudret, and Joe Wiart. Polynomial-chaos-based kriging. *International Journal of Uncertainty Quantification*, 5, 02 2015.

[29] P. Kersaudy, B. Sudret, N. Varsier, O Picon, and J. Wiart. A new surrogate modeling technique combining kriging and polynomial chaos expansions âĂŞ application to uncertainty analysis in computational dosimetry. *Journal of Computational Physics*, 286:103 – 117, 2015.

[30] Emmanuel Pagnacco, Eduardo Souza de Cursi, and Rubens Sampaio. Investigations into coupling kriging and polynomial chaos methods. In *XVII International Symposium on Dynamic Problems of Mechanics*, 2015.

[31] Roland Schöbi, B. Sudret, and S. Marelli. Rare event estimation using polynomial-chaos kriging. *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng.*, 3(2), 2017.

[32] Pramudita Satria Palar and Koji Shimoyama. Polynomial-chaos-kriging-assisted efficient global optimization. In *IEEE Symposium Series on Computational Intelligence*, Honolulu, HI, 2017.

[33] S. Dubreuil, N. Bartoli, C. Gogu, T. Lefebvre, and J. Mas Colomer. Extreme value oriented random field discretization based on an hybrid polynomial chaos expansion - kriging approach. *Comput. Methods Appl. Mech. Engrg.*, 332:540–571, 2018.

[34] Luca Margheri and Pierre Sagaut. A hybrid anchored-anova - pod/kriging method for uncertainty quantification in unsteady high-fidelity cfd simulations. *Journal of Computational Physics*, 324:137–173, 2016.

[35] Andrea Francesco Cortesi. *Predictive numerical simulations for rebuilding freestream conditions in atmospheric entry flows*. PhD thesis, L'Université De Bordeaux, 2018.

[36] Donald Coles and Alan J. Wadcock. Flying-hot-wire study of flow past an naca 4412 airfoil at maximum lift. *AIAA Journal*, 17(4):321–329, April 1979.

[37] Alan J. Wadcock. Structure of the turbulent separated flow around a stalled airfoil. Technical Report NASA-CR-152263, Beam Engineering, Inc., Sunnyvale, CA, February 1979.

[38] N. Spotts. 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.

[39] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, Oct. 1938.

[40] D. Xiu and G. E. Karniadakis. The wiener-askey polynomial chaos for stochastic differential equations. *SIAM Journal Sci. Comput.*, 24(2):619–644, July 2002.

[41] Dongbin Xiu and George Em Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal for Computational Physics*, 187:137–167, 2003.

[42] Géraud Blatman. *Adaptive sparse polynomial chaos expansions for uncertainty propogation and sensitivity analysis*. PhD thesis, Université Blaise Pascal - Clermont II, 2009.

[43] M. S. Eldred. Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design. In *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dyanmics, and Materials Conference*, Palm Springs, Florida, 2009. AIAA.

[44] G. Blatman and B. Sudret. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Prob. Engineering Mechanics*, 25(2):183–197, April 2010.

[45] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, 2nd edition, 2017.

[46] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. *The Annuals of Statistics*, 32(2):407–451, April 2004.

[47] M. Yuan, V. R. Joseph, and Y. Lin. An efficient variable selection approach for analyzing designed experiments. *Technometrics*, 49(4):430–439, Nov 2007.

[48] T. Hesterberg, N. H. Choi, L. Meier, and C. Fraley. Least angle and l_1 penalized regression: A review. *Statistics Surveys*, 2:61–93, 2008.

[49] G. Blatman and B. Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *J. Comput. Physics*, 230(6):2345–2367, March 2011.

[50] A. K. Prasad and S. Roy. Accurate reduced dimensional polynomial chaos for efficient uncertainty quantification of microwave/rf networks. *IEEE Transactions on Microwave Theory and Techniques*, 65(10):3697–3708, Oct 2017.

[51] X. Gao, Y. Wang, N. Xie, N. Spotts, S. Roy, and A. Prasad. Fast uncertainty quantification in engine nacelle inlet design using a reduced dimensional polynomial chaos approach. AIAA 2016-0298, 54th AIAA Aerospace Sciences Meeting, 2016.

[52] Oliver G. Ernst, Antje Mugler, Hans-J org Starkloff, and Elisabeth Ullmann. On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46:317–339, 2012.

[53] Anthony O'Hagan. Polynomial chaos: A tutorial and critique from a statistician's perspective. *submitted to SIAM/ASA Journal of Uncertainty Quantification*, 2013.

[54] Georges Matheron. Principles of geostatistics. *Economic Geology*, 58(8):1246–1266, 1963.

[55] Richard Webster and Margaret A. Oliver. *Geostatistics for Environmental Scientists*. John Wiley and Sons Ltd., West Sussex, England, 2007.

[56] O. Dubrule. Cross validation of kriging in a unique neighborhood. *Mathematical Geology*, 15(6):687–699, 1983.

[57] M Armstrong. Problems with universal kriging. *Mathematical Geology*, 16(1):101–108, 1984.

[58] Mark S. Handcock and Michael L. Stein. A bayesian analysis of kriging. *Technometrics*, 35(4):403–410, Nov. 1993.

[59] G. Bohling. Kriging. Technical report, University of Kansas, http://www.people.ku.edu/ gbohling/cpe940, 2005.

[60] J. D. Martin and T. W. Simpson. Use of kriging models to approximate deterministic computer models. *AIAA Journal*, 43(4):853–863, April 2005.

[61] S. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[62] Koji Shimoyama, Soshi Kawai, and Juan J. Alonso. Dynamic adaptive sampling based on kriging surrogate models for efficient uncertainty quantification. In *54th AIAA/ASME/ASCE/AHS/ACS Structures, Structural Dynamics, and Materials Conference*. AIAA, 2013.

[63] Ying Xiong, Wei Chen, Daniel Apley, and Xuru Ding. A non-stationary covariance-based kriging method for metamodelling in engineering design. *Int. J. Numer. Meth. Engng.*, 71:733–756, 2007.

[64] Matthew Plumlee and Daniel W. Apley. Lifted brownian kriging models. *Technometrics*, 59(2):165–177, 2107.

[65] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4(4):409–435, 1989.

[66] Noel Cressie. *Statistics for Spatial Data*. Wiley Classics Library, 2015.

[67] F. Bachoc. Cross validation and maximum likelihood estimations of hyper-parameters of guassian processes with model misspecification. *Computational Statistics and Data Analysis*, 66:55–69, March 2013.

[68] MATLAB. *UserGuide R2018a*. The MathWorks Inc., 2018.

[69] S. Kawai and K. Shimoyama. Kriging-model-based uncertainty quantification in computational fluid dynamics. In *32nd AIAA Applied Aerodynamics Conference*, Atlanta, GA, June 2014. AIAA.

[70] Mohamed Amine Bouhlel, Nathalie Bartoli, Abdelkader Otsmane, and Joseph Morlier. Improving kriging surrogates of high-dimensional design models by partial least squares dimension reduction. *Struct. Multidisc. Optim.*, 2016.

[71] Paul G. Constantine, Eric Dow, and Qiqi Wang. Active subspace methods in theory and practice: Applications to kriging surfaces. *SIAM Journal on Scientific Computing*, 36(4):A1500–A1524, 2014.

[72] Pramudita Satria Palar and Koji Shimoyama. Exploiting active subspaces in global optimization: How complex is your problem? *CoRR*, abs/1707.02533, 2017.

[73] Pramudita Satria Palar and Koji Shimoyama. On the accuracy of kriging model in active subspaces. In *AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Kissimmee, Florida, 2018. AIAA.

[74] I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling & Computational Experiment*, 1(4), 1993.

[75] A. Saltelli, S. Tarantola, and F. Campolongo. Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15(4):377–395, 2000.

[76] M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment*, 55(1-3):271–280, Feb 2001.

[77] Bertrand Iooss and Paul Lemaître. *Uncertainty Management in Simulation-Optimization of Complex Systems*, chapter A review on global sensitivity analysis methods. Springer, 2014.

[78] Justin Weinmeister, Nelson Xie, Xinfeng Gao, Aditi Krishna Prasad, and Sourajeet Roy. Combining a reduced polynomial chaos expansion approach with universal kriging for un-

certainty quantification. In *8th Theoretical Fluid Mechanics Conference*, Denver, CO, June 2017.

[79] Justin Weinmeister, Nelson Xie, Xinfeng Gao, Aditi Krishna Prasad, and Sourajeet Roy. Analysis of a polynomial chaos-kriging metamodel for uncertainty quantification in aerospace applications. In *2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Kissimmee, FL, January 2018.

[80] H. Rabitz and O. F. Alis. General formulation of high-dimensional model representations. *Journal of Math. Chem*, 50(2-3):197–233, 1999.

[81] G. Blatman and B. Sudret. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering and System Safety*, 95:1216–1229, June 2010.

[82] G. Blatman and B. Sudret. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. *Comptes Rendus MÃl'canique*, 336(6):518–523, June 2008.

[83] Art B. Owen. Better estimation of small sobol' sensitivity indices. *ACM TOMACS*, 23(2), 2012.

[84] Nam Hee Choi, William Li, and Ji Zhu. Variable selection with the strong heredity constraint and its oravle property. *Journal of the American Statistical Association*, 105(489):354–364, March 2010.

[85] M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

[86] Ruichen Jin, Wei Chen, and Agus Sudjianto. An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, 134:268–287, 2005.

[87] F. A. C. Viana, Venter Gerhard, and Balabanov Vladimir. An algorithm for fast optimal latin hypercube design of experiments. *International Journal for Numerical Methods in Engineering*, 82(2):135–156, 2009.

[88] Michael Sasena, Matthew Parkinson, Pierre Goovaerts, Panos Papalambros, and Matthew Reed. Adaptive experimental design applied to an ergonomic testing procedure. In *ASME 2002 International Design Engineering Technical Conferences and Computers*, volume 2. ASME, 2002.

[89] Andrea Garbo and Brian J. German. Adaptive sampling with adaptive surrogate model selection for computer experiment applications. In *18th AIAA/ISSMO Multidisciplinary Analysis and Optimiation Conference*. AIAA, 2017.

[90] N. Fajraoui, S. Marelli, and B. Sudret. On optimal experimental designs for sparse polynomial chaos expansions. Technical Report RSUQ-2017-001, ETH Zurich, 2017.

[91] Annette M. Molinaro, Richard Simon, and Ruth M. Pfeiffer. Prediction error estimation: a comparison of resampling methods. *Bioinformatics*, 21(15):3301–3307, 2005.

[92] Huu Minh Nguyen, Ivo Couckuyt, Luc Knockaert, Tom Dhaene, Dirk Gorissen, and Yvan Saeys. An alternative approach to avoid overfitting for surrogate models. In *Proceedings of the Winter Simulation Conference*, pages 2765–2776. Winter Simulation Conference, 2011.

[93] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer, 2017.

[94] D. Xiu. Fast numerical methods for stochastic computations: A review. *Commun. Comput. Phys.*, 5(2-4):242–272, Feb. 2009.

[95] http://www.metacomptech.com, 2015.

[96] Michael J. Moran, Howard N. Shapiro, Daisie D. Boettner, and Margaret B. Bailey. *Fundamentals of Engineering Thermodynamics*. John Wiley & Sons, 7th edition, 2011.

[97] Sutherland's law. Web, October 2008. https://www.cfd-online.com/Wiki/Sutherland's law.

[98] A. Satelli, S. Tarantola, and F. Campolongo. Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15(4):377–395, 2000.