An overview of uncertainty quantification techniques with application to oceanic and oil-spill simulations

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2015JC011366

Special Section:
Physical Processes Responsible for Material Transport in the Gulf of Mexico for Oil Spill Applications

An overview of uncertainty quantification techniques with application to oceanic and oil-spill simulations

Mohamed Iskandarani[1], Shitao Wang[1], Ashwanth Srinivasan[2], W. Carlisle Thacker, Justin Winokur[3], and Omar M. Knio[4,5]

[1]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, USA, [2]Tendral LLC, Miami, Florida, USA, [3]Sandia National Laboratories, Albuquerque, New Mexico, USA, [4]Department of Mechanical Engineering and Material Science, Duke University, Durham, North Carolina, USA, [5]Division of Computer, Electrical and Mathematical Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Key Points:
- Overview of surrogate construction with polynomial chaos and Gaussian process methods
- Uncertainty quantification with surrogates is more efficient than Monte Carlo methods
- Noisy model outputs require filtering

Correspondence to:
M. Iskandarani,
miskandarani@rsmas.miami.edu

Citation:
Iskandarani, M., S. Wang, A. Srinivasan, W. Carlisle Thacker, J. Winokur, and O. M. Knio (2016), An overview of uncertainty quantification techniques with application to oceanic and oil-spill simulations, *J. Geophys. Res. Oceans*, *121*, 2789–2808, doi:10.1002/2015JC011366.

Received 1 OCT 2015
Accepted 31 MAR 2016
Accepted article online 4 APR 2016
Published online 22 APR 2016

Abstract We give an overview of four different ensemble-based techniques for uncertainty quantification and illustrate their application in the context of oil plume simulations. These techniques share the common paradigm of constructing a model proxy that efficiently captures the functional dependence of the model output on uncertain model inputs. This proxy is then used to explore the space of uncertain inputs using a large number of samples, so that reliable estimates of the model's output statistics can be calculated. Three of these techniques use polynomial chaos (PC) expansions to construct the model proxy, but they differ in their approach to determining the expansions' coefficients; the fourth technique uses Gaussian Process Regression (GPR). An integral plume model for simulating the Deepwater Horizon oil-gas blowout provides examples for illustrating the different techniques. A Monte Carlo ensemble of 50,000 model simulations is used for gauging the performance of the different proxies. The examples illustrate how regression-based techniques can outperform projection-based techniques when the model output is noisy. They also demonstrate that robust uncertainty analysis can be performed at a fraction of the cost of the Monte Carlo calculation.

1. Introduction

The rapid increase in computational power and the increased sophistication of numerical models have made simulations an essential tool in engineering and scientific disciplines. Reliable simulations demand good fidelity between the real system and its virtual counterpart, and modelers attempt to limit the disparities between the two. The sources of these disparities can be roughly categorized into: modeling errors (misspecifying the mathematical model describing the real system), numerical errors generated by the discretization of the continuous mathematical equations, and data errors caused by uncertainties in the model's input data, such as its initial and boundary conditions, forcing, model parameters, etc. The focus here is on data errors.

Uncertainty quantification (UQ) is concerned with estimating the impact of the uncertain input data on the model's outputs. It aims to enhance the model's usefulness by presenting an output in a probabilistic framework that allows users to estimate robustly and efficiently its most likely value as well as the confidence and uncertainties associated with that value. This propagation of uncertainty through the model allows additional uncertainty analyses such as determining the principal contributors to the output uncertainties (sensitivity analysis), using observational data to refine *input* uncertainties and thereby to reduce output uncertainties (inverse propagation analysis and data assimilation), and guiding the acquisition of additional observational data (observing system simulation experiments).

The present article provides an overview of several inexpensive techniques for propagating uncertainty in inputs to outputs and compares their performance within the context of a simple integral plume model of a deep sea oil spill. Such a comparison is not feasible for more complex models due to the high cost of each run, but this model's relatively low cost permitted the use of traditional Monte Carlo simulations to obtain output statistics for comparison with those resulting from the inexpensive techniques. Applications of these different techniques to oil-fate modeling [*Gonçalves et al.*, 2016], plume modeling (S. Wang et al.,

Propagation of uncertainty and sensitivity analysis in an integral oil-gas plume model, submitted to Journal of Geophysical Research: *Oceans*, 2016, Wang 2016 hereafter), and to circulation forecasting in the Gulf of Mexico (M. Iskandarani et al., Quantifying Uncertainty in Gulf of Mexico Forecasts Stemming from Uncertain Initial Conditions, submitted to Journal of Geophysical Research: Oceans, 2016, Iskandarani 2016 hereafter) will be presented in separate articles in the special JGR issue on "Physical Processes Responsible for Material Transport in the Gulf of Mexico for Oil Spill Applications." The present article consolidates the methodology sections of these papers. It also provides the opportunity to present to oceanographers the rapidly evolving field of uncertainty quantification.

The propagation of uncertainty can be viewed as the transformation of an input probability density function (PDF) into an output probability density function. For example, the ensemble of input data specifying the Monte Carlo simulations is drawn from the input PDF and the ensemble of outputs these runs produce can be regarded as being drawn from the output PDF. The main drawback of this direct approach is the large number of simulations required, which prohibits its application when sufficient computational resources are unavailable or when the uncertainty analysis must be performed in a relatively short period.

Alternatives to Monte Carlo sampling, which are the focus of the present article, rely on what could be described as "indirect" sampling. An inexpensive proxy for the model is constructed using a relatively small ensemble size, and this proxy is then used to generate a much larger ensemble. In effect, most of the costly simulations are replaced with cheap proxy counterparts. The reliability and efficiency of this approach hinges on producing an efficient and accurate proxy for the forward model. Polynomial chaos expansions and Gaussian Process Regressions are two categories of such techniques (the proxy is often referred to as surrogate in the first case and an emulator in the second; here we will use the three terms interchangeably). Three of the four proxy construction techniques presented here use spectral expansions made up of orthogonal polynomials to represent the dependence of the model output on the uncertain inputs; they differ in their approach to calculate the series coefficients (Galerkin projection, least squares, and basis pursuit denoising). The fourth technique relies on Gaussian Process Regression (GPR) and requires the specification of a covariance function and its hyperparameters. The four techniques can be regarded as different ways to reconstruct the output's response surface [*Thacker et al*., 2015] using a relatively small ensemble. All four can be characterized as "nonintrusive" (requiring no modification of model code) and as ensemble-based [*Le Maître et al*., 2002].

The four techniques could also be grouped according to how they handle noise, which could be generated inadvertently by iterative procedures in the model or intrinsically by stochastic model parameterizations and forcing. PC Galerkin projection does not handle noise explicitly, and thus can suffer from aliasing effects when sampling noisy model output. The other techniques can extract the model signal by filtering the noise via either a least squares [*Blatman*, 2009] or a Bayesian approach [*Sargsyan et al*., 2009].

The layout of this paper is as follows. Section 2 discusses the characterization of the input uncertainties in terms of their PDFs and establishes some notation for later use. Section 3 briefly describes the different PC-based approaches, and section 4 introduces the Gaussian process approach. Section 5 introduces the integral plume model and its UQ analysis using the different approaches presented here. Finally, section 6 concludes with a discussion.

2. Characterizing the Input Uncertainty

Let $u(\boldsymbol{x},\boldsymbol{\theta})$ refer to the model output requiring uncertainty analysis; u depends on the vector of deterministic-independent variables $\boldsymbol{x}$ (such as space and time) and on the vector of uncertain input data $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_d,\ldots,\theta_D)^\top$, where D is the number of uncertain inputs (and the dimension of the uncertainty space). Computational models require the specification of a value for each θ_d. The "correct" value may not be known however, and modelers must then pick a value from within an acceptable uncertainty range: $\theta_d^{\min} \le \theta_d \le \theta_d^{\max}$ (the uncertainty interval can be infinite or semi-infinite); most often the value chosen coincide with a mean or central value. The quantification of the uncertainty in θ_d requires that an input PDF also be assigned for each parameter.

A first step in UQ analysis consists of identifying and characterizing the input uncertainties by their ranges and PDFs. Modelers can easily recognize the uncertain inputs in a given model, but specifying the PDF can

be quite difficult in practice. This is largely due to a lack of knowledge and to the scarcity of observational data. The problem is particularly severe in the case of oceanic flows where dynamical scales are small compared to the basin size, and where extreme conditions make measurements very difficult and costly. One course of action is to use available data to identify the ranges of the uncertain inputs and, in the absence of additional information, to consider the parameter to be uniformly distributed. If additional information is available, the user can use it to sharpen his/her estimate of the input PDF. The applications discussed in *Thacker et al*. [2012], *Sraj et al*. [2013a, 2014], and *Gonçalves et al*. [2016] and in Wang 2016 and Iskandarani 2016 illustrate how the observational data are used to identify the uncertainty ranges of the input parameters. See also *Webster and Sokolov* [2000] for an interesting discussion on how expert opinion can be used to inform the specification of input PDFs.

An additional complication arises when the uncertain input is a field and not a scalar quantity. For example, initial conditions uncertainties can conceivably involve treating every prognostic model variable at each computational grid cell as an independent random variable whose input PDF needs to be specified; this is clearly impractical. Some mechanism is needed to limit the dimension of the uncertain space while capturing most of the uncertainty associated with that field. One solution is to decompose the field into Empirical Orthogonal Functions (EOFs) and to characterize the field's uncertainties as uncertainties in the amplitudes of a few EOFs. Iskandarani 2016 and *Gonçalves et al*. [2016] rely on these EOFs to perturb fields. For the remainder of the paper, we assume that the user has capitalized on the available a priori information to specify the uncertain parameters and their input PDFs.

It is convenient to introduce (dimensionless) standard random variables that are related to the original parameters by a simple map, e.g.,

$$\xi_d = \begin{cases} \dfrac{\theta_d - \theta_d^{\min}}{\theta_d^{\max} - \theta_d^{\min}} & \text{when the uncertainty range is finite} \\ \dfrac{\theta_d - \bar{\theta}_d}{\sigma_{\theta_d}} & \text{when the uncertainty range is infinite} \end{cases}, \tag{1}$$

where $\bar{\theta}_d$ and σ_{θ_d} are the mean and standard deviation of θ_d. The vector $\boldsymbol{\xi}=(\xi_1,\ldots,\xi_D)^\top$ will refer to the uncertain input variables in the remainder of this article, and we will omit referencing the deterministic variables $\boldsymbol{x}$ for clarity of presentation.

3. Polynomial Chaos Expansions

Polynomial chaos (PC) [*Wiener*, 1938; *Ghanem and Spanos*, 1991; *Xiu and Karniadakis*, 2002; *Najm*, 2009; *Le Maître and Knio*, 2010] is a spectral-type expansion where the model dependence on the uncertain input variables is represented as a polynomial expansion of the form:

$$u(\boldsymbol{\xi}) \approx u_P(\boldsymbol{\xi}) = \sum_{k=0}^{P} \hat{u}_k \psi_k(\boldsymbol{\xi}), \tag{2}$$

where $u(\boldsymbol{\xi})$ is the model output (presumed to have finite variance) and u_P is its truncated series representation, $\hat{u}_k$ are the PC coefficients (both of which in general depend on $\boldsymbol{x}$), ξ is the vector of independent uncertain input variables with joint PDF $p(\boldsymbol{\xi})$, and $\psi_k(\boldsymbol{\xi})$ are orthogonal multivariate polynomial basis functions in the $\boldsymbol{\xi}$-space. Appendix A provides some details on the multidimensional construction and truncation of the basis functions. For reference, we also define the discrepancy between the model response and its truncated representation, which defines the error associated with using the surrogate rather than the original model:

$$\epsilon(\boldsymbol{\xi}) = u(\boldsymbol{\xi}) - u_P(\boldsymbol{\xi}). \tag{3}$$

The basis functions $\psi_k(\boldsymbol{\xi})$ are chosen to form an orthogonal basis in the weighted inner product defined by

$$\langle f, g\rangle = \int f(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, p(\boldsymbol{\xi})\ \mathrm{d}\boldsymbol{\xi}, \ \text{so that } \langle \psi_\mathrm{j}, \psi_\mathrm{k}\rangle = \delta_{\mathrm{j,k}} \langle \psi_\mathrm{k}, \psi_\mathrm{k}\rangle = \delta_{\mathrm{j,k}} ||\psi_\mathrm{k}||_2^2, \tag{4}$$

where $f(\boldsymbol{\xi})$ and $g(\boldsymbol{\xi})$ are any two square integrable functions of $\boldsymbol{\xi}$ (meaning that their norms as defined in equation (4), $||f||_2$ and $||g||_2$, are finite), and $\delta_{j,k}$ is equal to 1 when $j=k$ and 0 otherwise. Table 5 lists

examples of orthogonal bases for some commonly used weight functions. Since the weight $p(\xi)$ appearing in the inner product is the input PDF, the inner product coincides with an expectation operation on the quantity fg. This simplifies the calculation of statistical moments of u; for example, the mean reduces to the first coefficient $\hat{u}_0$ whereas the variance is simply a weighted sum of the square of the remaining coefficients (see Appendix B). The truncation level P controls the accuracy of the series representation and, along with the dimension of the uncertain space D, the computational cost required to calculate the $(P+1)$ coefficients. Once the $\hat{u}_k$ are available, the series can be used as a surrogate for the model and sampled much more efficiently. The bulk of the computational burden is then associated with the determination of these coefficients.

Several nonintrusive (ensemble-based) approaches are available to determine these coefficients, namely Galerkin projection [*Ghanem and Spanos*, 1991; *Le Maître et al*., 2002; *Le Maître and Knio*, 2010], least squares [*Blatman*, 2009], and compressed sensing [*Doostan and Owhadi*, 2011]. The main philosophy of the different approaches revolves around minimizing different norms of the truncation error ϵ.

3.1. Galerkin Projection

The Galerkin projection determines the coefficients by minimizing the error norm $||\epsilon||_2 = \sqrt{\langle \epsilon, \epsilon \rangle}$, where the angle brackets refer to the inner product defined in equation (4). Thanks to the orthogonality property of the basis set, this minimization reduces to a simple projection of u on the basis set with the coefficients given by:

$$\hat{u}_k = \frac{\langle u, \psi_k \rangle}{\langle \psi_k, \psi_k \rangle} \,. \tag{5}$$

The remaining task is the evaluation of the integral to compute the inner product in the numerator. This can be done by numerical quadrature:

$$\langle u, \psi_k \rangle = \int u(\xi)\psi_k(\xi)p(\xi)\ \mathrm{d}\xi \approx \sum_{i=1}^{S} u(\xi_i)\psi_k(\xi_i)\,\omega_i \,, \tag{6}$$

where ξ_i and ω_i are the quadrature's points and weights and where S is the number of quadrature points needed for an accurate approximation. The Galerkin approach is often referred to as "pseudospectral" when numerical quadrature is used to compute the series coefficients. The main computational expense is the evaluation of the model output $u(\xi_i)$ for inputs specified by the quadrature points, since each quadrature point requires a run of the numerical model. A number of quadrature rules are available such as tensorized Gauss quadrature, Sparse Smolyak, and their adaptive variants (see Appendices A and C for a brief discussion on the construction of multidimensional basis functions and quadrature).

In this study, we use the pseudospectral projection based on Gauss-Kronrod-Patterson 1-D rule for the Galerkin projection; its details can be found in *Gerstner and Griebel* [1998, 2003], *Huan* [2010], *Constantine et al*. [2012a], *Conrad and Marzouk* [2013], and *Winokur et al*. [2013]. The projection approach is most suitable when the function $u(\xi)$ varies smoothly with respect to ξ as a low-order expansion is then sufficient to approximate the model output accurately. The projection's computational cost can however become problematic when high-order expansions are necessary, for example, when localized sharp transitions occur within the interval of interest. Moreover, when $u(\xi)$ is polluted by small-scale noise, other approaches are required; sections 3.2, 3.3, and 4 discuss regression-based techniques that can filter this noise effectively.

3.2. Least Squares Regression

The least squares approach oversamples the response surface to filter the noise and thus requires more samples than model coefficients, $S > (P+1)$. The series coefficients are determined by minimizing the sum of the squared discrepancy between surrogate and model response at the S sample points:

$$||\epsilon||_{\mathrm{LS}}^2 = \sum_{i=1}^{S} \left(u(\xi_i) - u_P(\xi_i)\right)^2. \tag{7}$$

Note that the S sampling points are not necessarily quadrature points. This leads to the least squares solution [*Blatman*, 2009] for the series coefficients:

$$\hat{\boldsymbol{u}} = \left(\boldsymbol{\Psi}\boldsymbol{\Psi}^{\top}\right)^{-1}\boldsymbol{\Psi}\boldsymbol{u}, \tag{8}$$

where $\boldsymbol{\Psi}$ is a $(P+1)\times S$ matrix whose entries are given by $\Psi_{k,i}=\psi_k(\xi_i)$; $\hat{\boldsymbol{u}}^{\top}=(\hat{u}_0,\ldots,\hat{u}_P)$ is the vector of PC coefficients that minimizes the sum of the squared discrepancy, and $\boldsymbol{u}^{\top}=(u(\xi_1),\ldots,u(\xi_S))$ is the $S\times 1$ vector of samples of the response surface. Again, the computation of $\boldsymbol{u}$ constitutes the bulk of the cost; the least squares approach incurs the additional cost of solving the matrix system (8). The choice of sampling points is a delicate matter since an ill-conditioned matrix $\boldsymbol{\Psi}\boldsymbol{\Psi}^{\top}$ may lead to an unstable calculation. A number of recent articles [*Hampton and Doostan*, 2015; *Zhou et al*., 2014; *Shin and Xiu*, 2016; P. Seshadri et al., Optimal quadrature subsampling for least squares polynomial approximations, arXiv:1601.05470, 2016] have proposed optimal and stable sampling strategies for least squares polynomial approximations in high-dimensional spaces.

If the number of samples is the same as the number of coefficients, $S=(P+1)$, the matrix $\boldsymbol{\Psi}$ is square and, provided it is nonsingular, admits a solution of the form $\hat{\boldsymbol{u}}=(\boldsymbol{\Psi}^{\top})^{-1}\boldsymbol{u}$ [*Tatang et al*., 1997]. Then, rather than smoothing the response samples, as one might want when the response is noisy, the series reproduces them exactly, $u(\xi_i)=u_P(\xi_i)$ for $1\le i\le S$, so that the discrepancy in the surrogate, equation (3), is exactly zero at the sampling points and $||\epsilon||_{\mathrm{LS}}$ vanishes. In this case, the choice of points to sample is still a delicate matter, as it can cause $\boldsymbol{\Psi}^{\top}$ to be ill conditioned. Alternatively, the stochastic collocation method [*Klimke*, 2006a,b; *Ganapathysubramanian and Zabaras*, 2007; *Le Maître and Knio*, 2010; *Narayan and Zhou*, 2015] also enforces zero discrepancy at the sampling points but avoids the matrix inversion by linearly recombining the orthogonal polynomials to produce the so-called nodal basis, where each basis functions is equal to 1 at only one sample point and 0 at all the others (the coefficient of that basis function becomes the value of the response at the sampling point); again, numerical stability is achieved by careful selection of the sampling points. Approaches to determine the coefficients that force the discrepancy to be zero at the sampling points are, however, ineffective when model noise is present, irrespective of the nature of the basis functions and the sampling strategy.

3.3. Basis Pursuit Denoising

The Basis Pursuit Denoising (BPDN) approach [*Doostan and Owhadi*, 2011] seeks to build a sparse representation of the response function (2), i.e., one that represents the simulated response to within a given tolerance while retaining as few coefficients in the series as possible (the sparsity requirement). To this end, the BDPN algorithm minimizes the sum of the magnitude of the series coefficients while requiring the discrepancy between the surrogate and model to be small:

$$\min\left(\sum_{k=1}^{P}|\hat{u}_k|\right)\ \text{subject to}\ ||\epsilon||_{\mathrm{LS}}=\left[\sum_{i=1}^{S}\left(u(\xi_i)-u_P(\xi_i)\right)^2\right]^{\frac{1}{2}}\le\sigma. \tag{9}$$

The use of $||\epsilon||_{\mathrm{LS}}$ as constraint allows this technique to choose the fewest terms while guaranteeing that root-mean-square differences between surrogate and samples is no greater than the specified tolerance σ. The inequality constraint on the magnitude of the discrepancy allows BDPN to handle noisy model outputs, and the chosen tolerance σ should reflect the noise's amplitude. The BDPN algorithm works best when the response function is sparse, i.e., when a small fractions of the terms dominate the truncated PC expansion. In that case, the number of samples needed to reconstruct the function can be much smaller than the number of terms retained in the series: $S\ll(P+1)$. The choice of the noise level σ is important for BDPN. If σ is too small, the surrogate may overfit the data, while if σ is too large, it may not be accurate. A cross-validation approach proposed in *Doostan and Owhadi* [2011] and *Peng et al*. [2014] can be used to select the optimal σ. The BPDN software used here is the SPGL1 [*van den Berg and Friedlander*, 2007, 2008] Matlab solver.

4. Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) [*Rasmussen and Williams*, 2006] relies on an intrinsically statistical approach to approximate the model output's dependence on its uncertain input data. Constructing a GPR surrogate is not viewed as the truncation of a spectral expansion, as was the case for the polynomial chaos surrogates, instead, reproducing-kernel Hilbert spaces [*Weinert*, 1982], provide its function-theoretic

foundations. In practice, GPR involves the choice of a mean function, which specifies the default response when there are no simulations that can provide guidance, and a covariance function, which specifies how closely the surrogate for one set of inputs resembles another as the two sets of inputs become separated. This is exactly the same mathematics used commonly for data assimilation [*Daley*, 1993; *Kalnay*, 2003]: a forecast, which is to be corrected by observations, provides the mean (background) function, an error-covariance matrix is the covariance function evaluated for points on a computational grid, and the observations provide information for updating the forecast. Similarly, the sample outputs from model runs provide information for updating the mean function. And just as is the case for data assimilation, the statistics are best thought of as Bayesian. The mean function that is specified is the prior mean of possible surrogates and the updated mean based on samples provided by model runs is the surrogate that is constructed. We refer the reader to the appendix in *Thacker et al.* [2015] for more details on Gaussian Process Regression.

Because of its statistical underpinnings, GPR regression is ideal for providing a proxy for models with noise. However, its approach to incorporating noise is different from those of polynomial chaos projection or basis-pursuit denoising. For the least squares approach, more terms in the spectral expansion can be retained until the proxy agrees with the model within a tolerance that allows for the magnitude of the noise, while the compressed-sensing approach specifies the tolerance as an inequality constraint that is used in controlling which terms in the expansion should be retained. GPR, on the other hand, handles noise through the specification of a Gaussian probability density function relating sampled values to their noise-free counterparts. Consequently, the maximum-likelihood GPR proxy is obtained by minimizing a sum of two terms, one guaranteeing it is close to the samples, but not so close that it reflects the noise, and the other guaranteeing it is not far from the prior mean, so the estimate would not be absurd for any choice of input.

In practice, Gaussian Process Regression is quite straight-forward, as there is a simple equation relating the proxy response $u_a = u(\xi_a)$ to any input ξ_a desired:

$$u_a = m_a + \boldsymbol{K}_{as}(\boldsymbol{K}_{ss} + \boldsymbol{R}_{ss})^{-1}(\boldsymbol{u}_s - \boldsymbol{m}_s)\,, \tag{10}$$

where m_a is the prior mean function evaluated at the analysis point ξ_a and where the vectors $\boldsymbol{u}_s$ and $\boldsymbol{m}_s$ contain the sampled model output and the values of the prior mean function at the corresponding points, respectively. The covariance matrices $\boldsymbol{K}_{as}$ and $\boldsymbol{K}_{ss}$ are obtained by sampling the covariance function at pairs of points, the former containing the covariance of the response at all sampled points with the response at the analysis point and the later, for all pairs of sampled points. Finally, the covariance matrix $\boldsymbol{R}_{ss}$ characterizes the model noise manifesting in the samples. Similarly, there is a simple expression for the posterior covariance matrix characterizing the error of the proxy:

$$\boldsymbol{Q}_{aa} = \boldsymbol{K}_{aa} - \boldsymbol{K}_{as}(\boldsymbol{K}_{ss} + \boldsymbol{R}_{ss})^{-1}\boldsymbol{K}_{as}^{\top}\,, \tag{11}$$

where the elements of $\boldsymbol{Q}_{aa}$ are covariances of the approximate response at pairs of analysis points and where those of $\boldsymbol{K}_{aa}$, $\boldsymbol{K}_{ss}$, and $\boldsymbol{K}_{as}$ are given by the covariance function evaluated using the pairs of analysis or sampled points.

Before equations (10) and (11) can be used, it is necessary to specify the Gaussian process' mean and covariance functions and the covariance matrix for the model's noise. Unless there is information about the expected response, the mean function is typically taken to be an average of the samples. A common choice for covariance function is the squared exponential:

$$k(\xi_i, \xi_j) = \sigma_u^2 \exp\left[-\frac{(\xi_i - \xi_j)^{\top}\boldsymbol{L}(\xi_i - \xi_j)}{2}\right], \tag{12}$$

where σ_u^2 sets the magnitude and $\boldsymbol{L}$ is a positive definite matrix that determines the separation required between inputs to produce a significant change in the models response. In the simplest case, $\boldsymbol{L}$ is diagonal with $L_{dd} = l_d^{-2}$, where l_d is the correlation length for the response in the direction corresponding to the dth input. A simple choice for the noise covariance matrix is $\boldsymbol{R}_{ss} = \sigma^2\boldsymbol{I}$ where $\boldsymbol{I}$ is an identity matrix; this is equivalent to assuming that the noise of each sample is independent and identically distributed with zero mean and variance σ^2. The choice of the hyperparameters (σ, σ_u, and $\{\boldsymbol{L}\}$) affects the performance of GPR; *Rasmussen* [2006] discusses various techniques to find their optimal values given the data.

Table 1. Uncertain Input Variables and Their Initial Distributions[a]

Parameter	Parameter Range	Units
Entrainment coefficient	(0.06, 0.116)	
Entrainment ratio	(0.4, 0.6)	
Gas-to-oil ratio	(39.64, 84.95)	m^3/barrel
95th percentile of the droplet size	(1, 10)	mm
Droplet distribution spreading ratio	(1.5, 4)	

[a]The Ranges of the entrainment coefficient and entrainment ratio are from *Bhaumik* [2005], the Gas-to-oil ratio range is from *Reddy et al.* [2012] and *Valentine et al.* [2010], the ranges of the 95th percentile of the droplet size is from *Johansen et al.* [2001], and the Droplet distribution spreading ratio range from *Lefebvre* [1988].

5. Test Problem

The performance of the different techniques will be illustrated here using the TAMOC integral plume model [*Socolofsky et al.*, 2008]. The test problem consists of simulating an oil-gas plume resulting from a Deep Water Horizon blowout scenario. TAMOC simulates the plume within a few hundred meters of the wellhead where the dominant dynamical processes are buoyancy, turbulent mixing, buoyancy peeling, and gas dissolution [*Crounse et al.*, 2007]. TAMOC relies on a combination of theoretical assumptions and empirical data to parametrize these processes; as such the model output depends on quite a large number of uncertain inputs. Here we focus on a set of five uncertain parameters that significantly influence the buoyancy of the plume. The parameters and their uncertainty ranges are listed in Table 1; details on how the ranges were identified and a description of TAMOC can be found in Wang 2016. The uncertain parameters will be assumed to be uniformly distributed within the ranges indicated and we use the simple linear map in equation (1) to calculate the corresponding standardized input variables ξ_d. Furthermore, we restrict our attention to two model outputs, namely the trap height (i.e., the depth of the deepest intrusion layer) and the peel height which designates the level at which the plume stops ascending. Analyses of additional quantities of interest are presented in Wang 2016.

5.1. Proxy Model Configurations

Four different surrogates for the TAMOC plume model are considered, three that use polynomial chaos approaches and one that uses a Gaussian Process approach. The PC-based surrogates are based on a uniform probability density for inputs within their prescribed minima and maxima, and use Legendre polynomials as they are orthogonal when weighted uniformly. The PC-surrogates, however, differ in their truncation to a finite basis and in the way their coefficients are determined. On the other hand, the GPR-based surrogate requires the specification of suitable mean and covariance functions, the noise variance, and the associated hyperparameters. The present subsection summarizes the setups used for the surrogates' construction.

The multivariate PC basis consists of tensor products of univariate Legendre polynomials. For both trap height and peal height, in all cases except one, the maximal degree for each uncertain input was set to 5, the exception being a single PC regression experiment using ninth degree univariate polynomial factors. However, the particular combination of univariate factors is different for the three different methods (see Appendix A and Table 2 for more details). A total-order truncation (aka triangular truncation) discards terms where the sum of the degrees of the univariate Legendre polynomial factors contributed by the various inputs exceeds 5 and thus retains 252 multivariate polynomial terms, each with a coefficient to be determined. With this truncation, least squares regression using twice as many samples as coefficients to filter model noise requires an ensemble of 504 TAMOC runs. With a library of these 252 multivariate basis functions to choose from, basis pursuit denoising, on the other hand, requires only an ensemble of 100 simulations and finds 32 nonzero coefficients for the trap height and 174 for the peel height. The multidimensional quadrature (sparse Smolyak based on 1-D Gauss-Kronrod-Patterson quadrature) used with the projection approach maximizes the number of mixed degree terms retained in the series and leads to a truncation requiring the determination of 738 coefficients; their calculation required an ensemble of 903 realizations. The configurations of these different PC bases and their supporting ensembles are summarized in Table 3 along with the list of numerical experiments.

Table 2. Ensemble Size Needed to Calculate the Coefficients in a Projection-Based Five-Dimensional PC Series Truncated Isotropically at a Maximum Polynomial Degree p, Using Different Quadrature Rules

Polynomial Degree (p)	1	2	3	4	5
Basis size $(P+1)$	6	21	56	126	252
Tensorized Gauss $(p+1)^5$	6	36	216	1296	7776
Sparse Smolyak	11	51	151	391	903

The GPR was configured with a constant but unknown background $\tilde{u}$, and with a squared exponential covariance

Table 3. Table of Experiments, Proxy, and Sampling Configurations, and L2 Approximation Errors[a]

Experiment	Coefficient	Ensemble Type/Size	Basis Size or Kernel	Max. Poly. Degree	Multidim. Truncation S	Trap Error	Peel Error
PC-P	Galerkin	Quadrature/903	738	5	Pseudospectral	0.0509	0.0267
PC-R5	Regression	Random/504	252	5	$\left(\sum_{d=1}^{5} m_d\right) \leq 5$	0.0177	0.0097
PC-BPDN	Sparse	Random/100	252	5	Sparse	0.0195	0.0152
PC-R9	Regression	Random/4004	2002	9	$\left(\sum_{d=1}^{5} m_d\right) \leq 9$	0.0247	0.0139
GPR	Regression	Random/100	Squ. Exp			0.0159	0.0095
MC		Random/50,000					
REF-1D		200×5					

[a]The pseudospectral projection scheme calculates a larger number of coefficients than a traditional total-order truncation which has only 252 coefficients. The additional terms retained are the high-order mixed terms disregarded by the traditional total-order truncation. The numbers of nonzero PC coefficients determined using BPDN are 32 for trap height and 174 for peel height. Only the Galerkin projection required a quadrature-based ensemble, all the other techniques used randomly generated samples.

function of the form shown in equation (12). The matrix $\boldsymbol{L}$ is diagonal with entries $L_{dd}=l_d^{-2}$, where l_d determines the range of influence of each sample along the dth direction. The list of hyperparameters consisted of the eight parameters $(\sigma, \sigma_u, l_{1\leq i\leq 5}, \tilde{u})$ where $\tilde{u}$ refers to the mean background value. A maximum-likelihood method was used to identify the optimal hyperparameters and relied on conjugate gradient iterations; 20 random restarts were used for each hyperparameter to avoid local extrema. The posterior GPR hyperparameters are shown in the Table 4. The σ hyperparameter in GPR represents the noise in the model which is relatively small compared with the signal variance σ_u, and the characteristic length scales represent the decorrelation length scales in different parameter dimensions.

The relative efficiency of TAMOC allowed us to perform an independent uncertainty analysis which relied on a 50,000 Monte Carlo ensemble designed using Latin Hypercube Sampling. The MC analysis, with its direct sampling of the model response, will allow us to gauge the merits of the different surrogates by measuring their approximation errors and their usefulness in representing the output's statistics. Additional samples were calculated to visualize the nature of the model response to variations in the uncertain inputs along the five unidimensional axes passing through the center of the uncertain space.

5.2. Response Curves Analysis

Figure 1 compares the different proxy response curves of trap and peel heights to unidimensional variations in the uncertain inputs with those obtained from the unidimensional validation samples. The validation samples reveal that the trap and peel heights calculations exhibit quite a bit of small-scale noise which is attributable to the way the plume model terminates an iterative procedure and/or to TAMOC's sensitivity to the measured stratification profile. The noise is of small amplitude (less than 2% of the signal's amplitude) and the errors incurred are negligible from a practical point of view. The representation of this small-scale noise in the PC surrogate series would require a very long series and a costly ensemble for no practical benefit. Except for the noise, the curves exhibit very little curvature (in fact the trap height seems to vary linearly with the entrainment coefficient and ratio but exhibits some nonlinear dependence on the 95th percentile droplet size), so low-degree polynomials should be able to capture this behavior. The response curves of the different proxies, whether PC-based or GPR, capture the large-scale trends but not the noise. The largest errors seem to be associated with PC-projection, which is incapable of filtering the noise; its response curves exhibits larger oscillations than the other approaches. The presence of noise in the response curves makes

Table 4. GPR Hyperparameters for Trap Height and Peel Height[a]

Hyperparameter	Mean	l_1	l_2	l_3	l_4	l_5	σ_u	σ
Trap height (m)	327.22	5.13	7.43	11.75	1.95	7.77	162.09	3.38
Peel height (m)	558.43	2.56	6.00	7.30	1.72	6.23	222.55	3.17

[a]$l_1 - l_5$ represents the characteristic length scales in each dimension, σ_u is the signal standard deviation, and σ is the noise standard deviation.

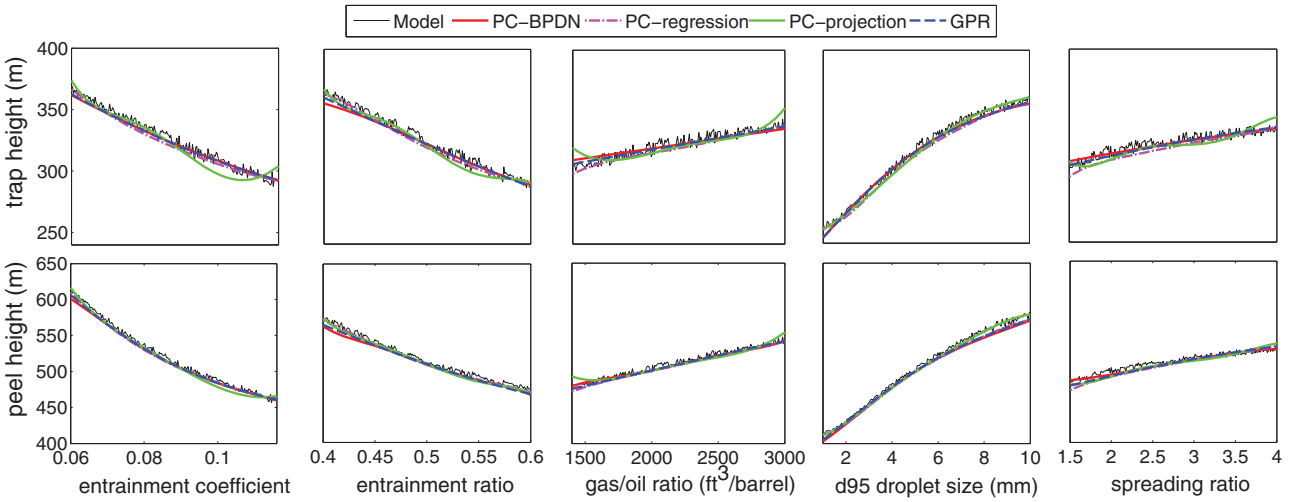

Figure 1. Trap and peel height samples from the univariate validation sets (thin noisy black lines) and the proxy models' response curves (colored thick lines) along the five central axes of the space of uncertain inputs. The 1-D proxy response curves were calculated from the 5-D proxies by varying only one parameter at a time while fixing the others at their central values.

this problem very challenging for projection-based techniques as they perform best when the dependence of the model output on the model input is smooth [*Sargsyan et al.*, 2009; *Canuto et al.*, 2006].

The coefficients of the different PC approaches are shown in Figure 2. All methods return coefficients of similar magnitude for the low-degree polynomial terms (left end of the *x* axis); the major differences are in the magnitudes of the coefficients of the high-degree polynomials (right end of the *x* axis). The PC-P series (black) retains the most polynomial terms but not enough to resolve the small-scale noise; the ensuing aliasing leads to high mode coefficients with substantial magnitude. The PC-R5 proxy (blue) exhibits a spectrum similar to PC-P for the high-order modes, except its spectrum is one third as long. Its ratio of number of samples to coefficients is 2 whereas the pseudospectral projection's is only 1.15; this higher ratio is probably responsible for the regression's better overall performance. This suggests that regression using a larger ratio of samples to number of coefficients would perform even better. The PC-BPDN series for the peel height (right) is less sparse than the trap height series (left) as the former retains more coefficients than the latter; most of these extra coefficients, however, have small magnitudes. Noisy model outputs thus represent a challenge to the construction of PC surrogates, and long PC series in this context are not necessarily better, as they risk overfitting to model noise.

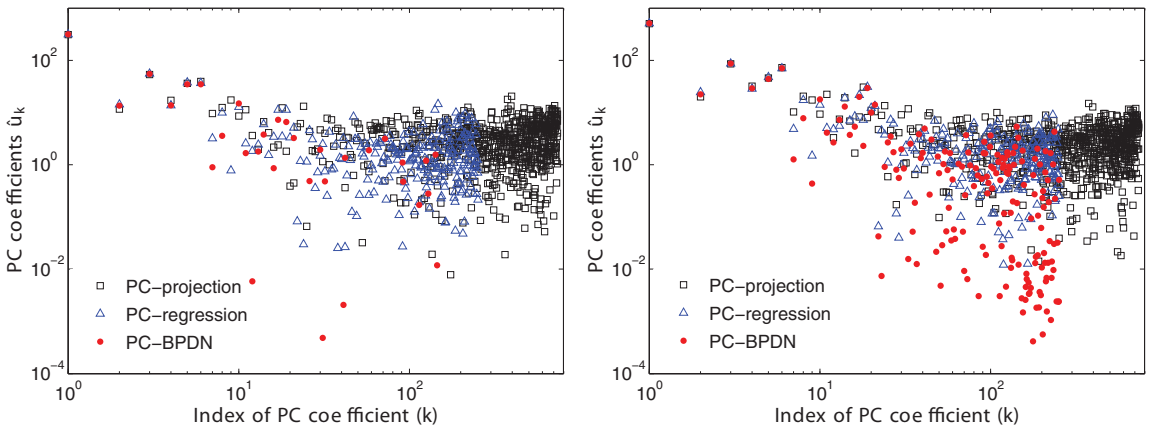

Figure 2. Distribution of coefficients for surrogates constructed using three polynomial-chaos methods. Recall that the number of coefficients and the polynomials that they multiply are different for the different surrogates; (left) trap height PC coefficients; (right) peel height PC coefficients.

5.3. Approximation Errors

The existence of model noise precludes a traditional analysis of the approximation error via a convergence study. We thus turn to analyzing the statistical distribution of the approximation errors in the five-dimensional uncertainty space. To this end, we define the relative local error as:

$$e(\xi_i^{\mathrm{MC}}) = \frac{|u_{\mathrm{proxy}}(\xi_i^{\mathrm{MC}}) - u_{\mathrm{MC}}(\xi_i^{\mathrm{MC}})|}{\max(u_{\mathrm{MC}}) - \min(u_{\mathrm{MC}})},\ i = 1, 2, \ldots, 5 \times 10^5, \tag{13}$$

where ξ_i^{MC} refers to the locations of the MC samples in the uncertain space.

Figure 3 show scatterplots of the relative local errors for trap and peel heights. Red lines indicating 95th percentile (Err95) levels have been added, so that it is possible to judge the magnitudes of the 5% of largest errors. The 1-D validation sets and the GPR posterior for the noise hyperparameter suggest a relative noise amplitudes of about 0.01–0.02, which provides a basis for judging whether relative local errors are large. There is no need for the local errors to be much smaller than the noise, but few samples should be much larger either. PC-P shows the largest errors: its Err95 is the largest (0.11 for trap height and 0.06 for peel height) with the local error reaching as much as 0.40 for the trap height and 0.30 for the peel height at some isolated sampling points. Quite a few PC-P surrogate samples have errors ranging between 0.15 and 0.30 for the trap height and 0.10 and 0.25 for the peel height. The regression methods, whether PC or GPR based, perform substantially better: their Err95 level stands at 0.03 and 0.02 for the trap and peel heights, respectively, and their maximum trap and peel heights error levels are 0.20 and 0.08, respectively, for GPR and PC-BDPN. PC-R5 maximum errors are somewhat higher reaching 0.30 at five isolated points for the trap height and 0.12 for the peel height. The best performers for the present problem are thus the PC-BDPN and GPR as they produce the lowest error levels for the least amount of computations requiring only 100 samples each. Both PC-BPDN and GPR account for model error and for a sparse representation while PC-regression only accounts for model error but does not guard against aliasing errors.

We also present a measure for the global error $||e||$

$$||e|| = \frac{\left(\sum_i \left(u_{\mathrm{MC}}(\xi_i^{\mathrm{MC}}) - u_{\mathrm{proxy}}(\xi_i^{\mathrm{MC}})\right)^2\right)^{\frac{1}{2}}}{\left(\sum_i \left(u_{\mathrm{MC}}(\xi_i^{\mathrm{MC}})\right)^2\right)^{\frac{1}{2}}}. \tag{14}$$

The global errors for trap height and peel height for the different proxy methods are shown in Table 3. Again, PC-P exhibits the largest error while other methods perform similarly; one exception is that the PC-BPDN peel height global error is larger than that from PC-R5 and GPR.

5.4. Comparison of Proxy-PDF to MC-PDF

The PDFs of the model outputs can be obtained by using the proxies to calculate a large number of model realizations without any additional model runs. Using Latin Hypercube Sampling, 10^5 new input points were generated and the trap and peel heights were then estimated using the different proxies. Figure 5 presents PDFs for the resulting peel (left) and trap (middle) heights. For comparison, corresponding PDFs obtained directly from the TAMOC Monte Carlo ensemble are also shown (Figure 4 confirms that the 50×10^3 MC samples delivers a "converged" estimate for the PDFs). The MC-based PDF for the trap height indicates a narrow peak near 275 m with a substantial shoulder between 300 and 360 m, while the peel height MC PDF has a broader peak between 450 and 500 m and shows some skewness toward higher values but no shoulder. The mode and spread of the peel height PDFs using proxies agree well with the MC reference solution. The proxy-based PDFs for the trap height, however, differ from the reference solution in the central portion of the distribution. Both PC-BDPN and GPR, with 100 realizations only, miss the mode and shoulder of the MC-PDF; they exhibit a Gaussian-like shape with a mode at 325 m. PC-P shows a broad peak where the shoulder is expected to be. Only the PC-R approach shows a hint of a peak near 275 m. All methods, however, seem to capture the tails well.

The disparities between the MC PDF for trap height and its proxy counterparts were further investigated to determine whether the discrepancies are caused by sampling errors, approximation errors or a combination of both. An analysis of the MC samples to identify the origins of the shoulder, i.e., to identify which uncertain input values lead to a trap height in the shoulder range, was inconclusive; no pattern was identifiable.

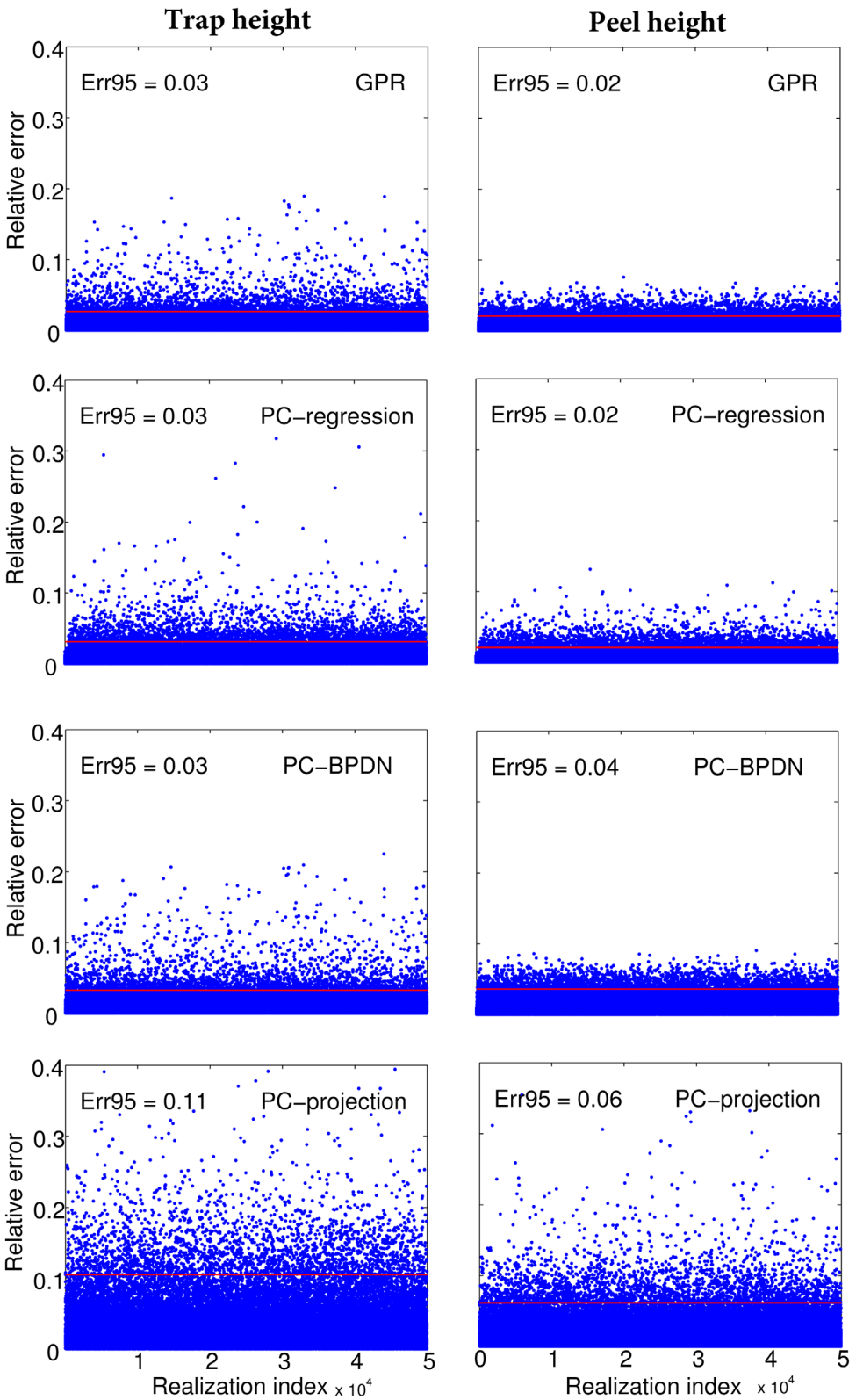

Figure 3. Blue points indicate the relative errors of the proxies for (left) trap height and (right) peel height. From top to bottom: proxies constructed using Gaussian-process regression, polynomial chaos regression, polynomial chaos basis-pursuit denoising, and polynomial chaos projection. Horizontal red lines indicate the 95-percentile error.

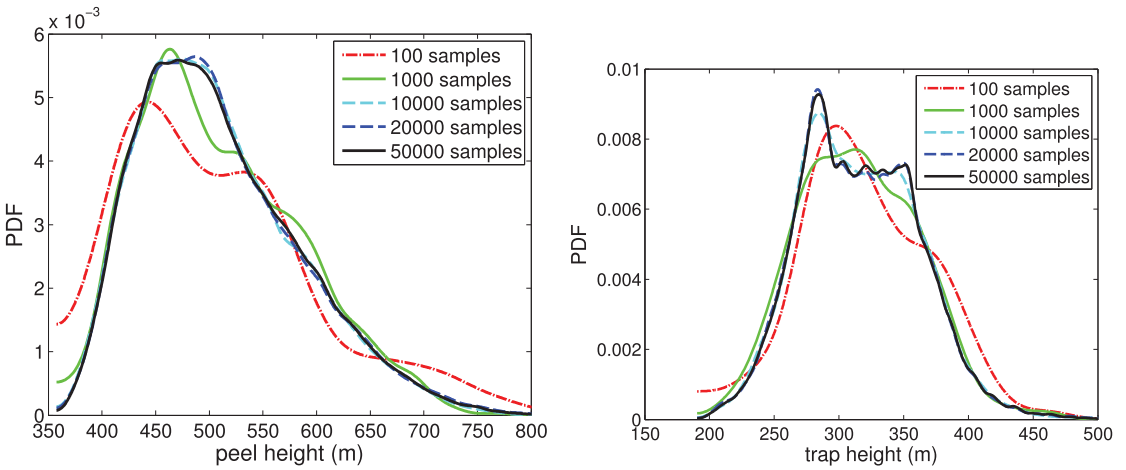

Figure 4. Probability density function of the (left) peel height and (right) trap height as estimated by Monte Carlo sampling using different sample sizes. The pdfs of the two largest samples are in agreement.

To investigate whether a PC series of higher polynomial degree might capture the shoulder, a new ninth degree PC-R proxy is constructed using a basis set of 2002 coefficients and 4004 realizations; the ensuing PDF is compared to that of the fifth degree PC-BPDN (100 realizations) and PC-R5 (504 realizations) in Figure 5, right. The ninth degree proxy PDF shows better agreement with the reference MC-PDF including the appearance of a shoulder, and a better defined narrow mode along with a better estimate of its probability. Additionally, we use the Kullback-Leibler [*Kullback*, 1959] divergence (also known as relative entropy):

$$D_{KL}(p||q) = \sum_i p(i) \ln \frac{p(i)}{q(i)}$$

to quantify the distance between two discretely defined probability distributions p and q. D_{KL} is nonnegative, is zero if p and q are identical, and increases as p and q grow apart. The D_{KL} divergences of the PC-BPDN, PC-R5, and PC-R9 PDFs from the Monte Carlo PDF are 0.1084, 0.0276, and 0.0125, respectively; the ninth degree basis is thus the closest to the MC-PDF. It appears that an improvement in the PDF estimate can be obtained using a higher degree PC basis, but this improvement occurs at the expense of a very large increase in the number of realizations. The following section attempts to describe the impact of misrepresenting the PDF on the sensitivity analysis.

5.5. Analysis of Variance

Analysis of variance, also known as sensitivity analysis in the UQ community, is an attempt at identifying the largest contributors to forecast uncertainty by estimating the variance injected by the different variables to the total variance; it is an important application of the forward uncertainty propagation. Here the impacts of proxy approximation errors on the sensitivity analysis are investigated. Two measures of sensitivity are

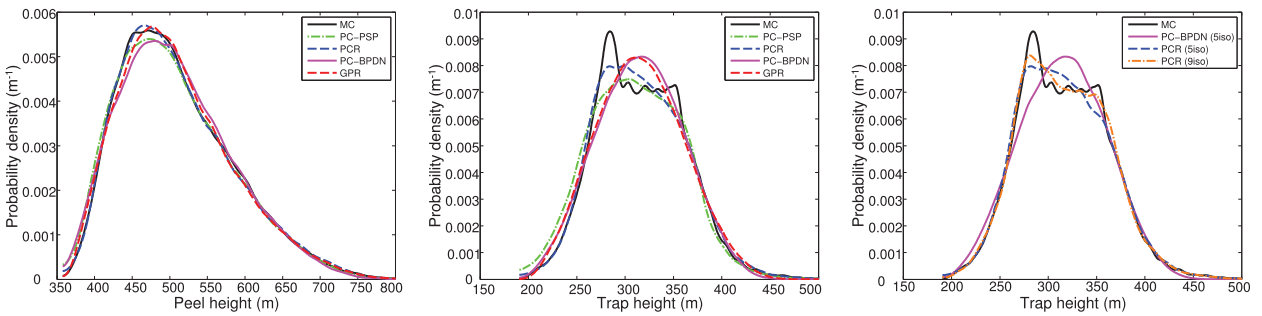

Figure 5. Comparisons of (left) peel height and (middle) trap height PDFs obtained from different proxies with the reference Monte Carlo PDF. (right) Comparison of the trap height PDF obtained using a new ninth degree polynomial basis (4004 realizations) with those obtained from a fifth degree basis using either regression or basis-pursuit denoising.

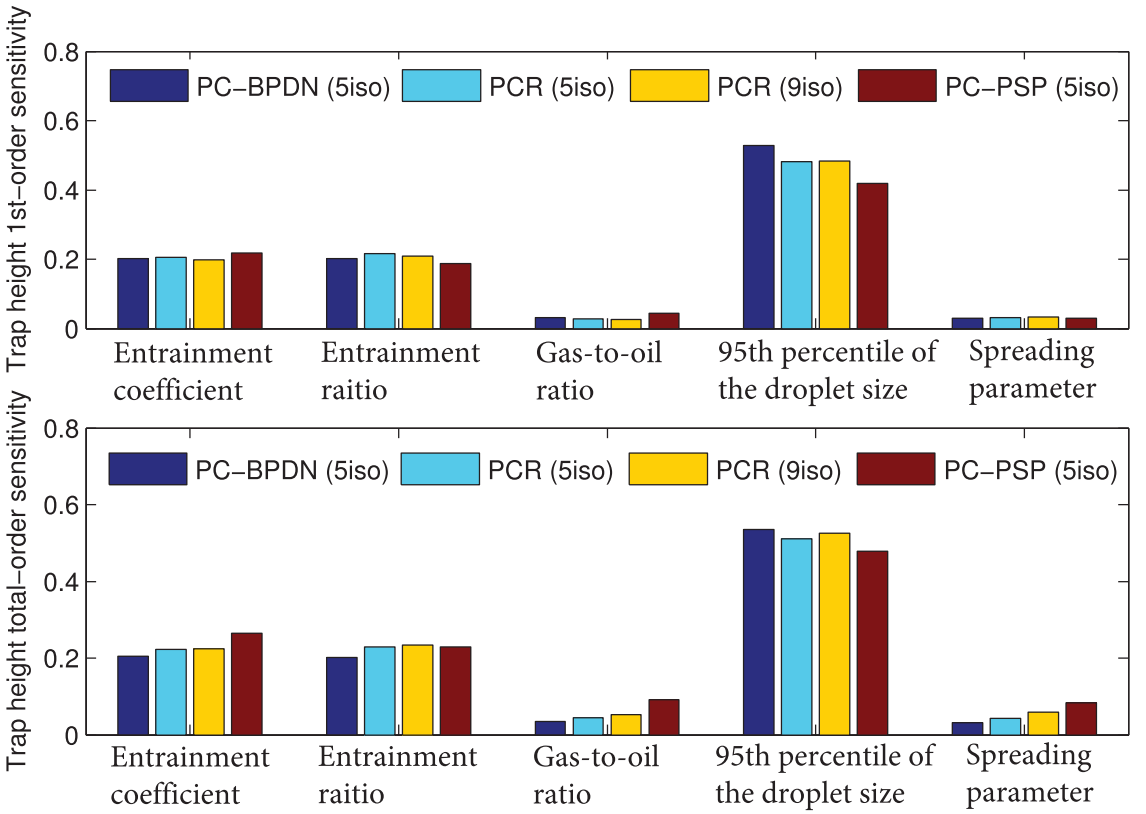

Figure 6. First-order and total-order sensitivity of trap height to the five uncertain inputs. Colors indicate the different polynomial chaos surrogates, bar heights indicate sensitivity, and group labels indicate the input variable.

commonly used. The first-order sensitivity index is the ratio of the variance contributed by the polynomial terms involving only a single specified input to the variance accounted for by all the terms; it measures the fraction of variance attributed to that input alone. The total-order sensitivity index is the relative variance contributed by all terms involving that specified input; it attempts to account for the variance due to the interaction of that input variable acting in concert with all the others. The orthogonality of the polynomials makes the computations of these ratios extremely easy [*Crestaux et al*., 2009; *Le Maître and Knio*, 2010; *Alexanderian et al*., 2012], whereas their estimation directly from model realizations is considerably more difficult [*Homma and Saltelli*, 1996a; *Saltelli*, 2002]. Details of the estimation of sensitivity indices can be found in Appendix D. First-order and total-order sensitivity of trap height and peel height to the five uncertain inputs calculated in the previous three experiments are shown in Figures 6 and 7, respectively. The fact that the sensitivity indices for the different surrogates agree with each other suggest that reliable sensitivity indices can be obtained using PC surrogates.

6. Discussion

All surrogate construction methods discussed here are ensemble based and are nonintrusive; that is, no code modification is necessary, and all that is needed is the ability to run the forward model at specified values of the uncertain parameters. The optimal approach to build a model surrogate is problem dependent, and insight about the response surface, $u(\xi)$, is clearly very useful in deciding which surrogate type is most suitable for the problem at hand. For model outputs that depends smoothly on the uncertain input data, the Galerkin projections are quite practical and straightforward to implement as exemplified in several oceanic applications [*Thacker et al*., 2012; *Alexanderian et al*., 2012]; moreover, the adaptive construction used in Winokur et al. [2013] and *Sraj et al*. [2013] was extremely useful in steering the sampling toward regions in the uncertainty space that contributed most toward the output uncertainties. The test problem presented here, on the other hand, features small-scale noise and methods that can filter it will fare better on this class

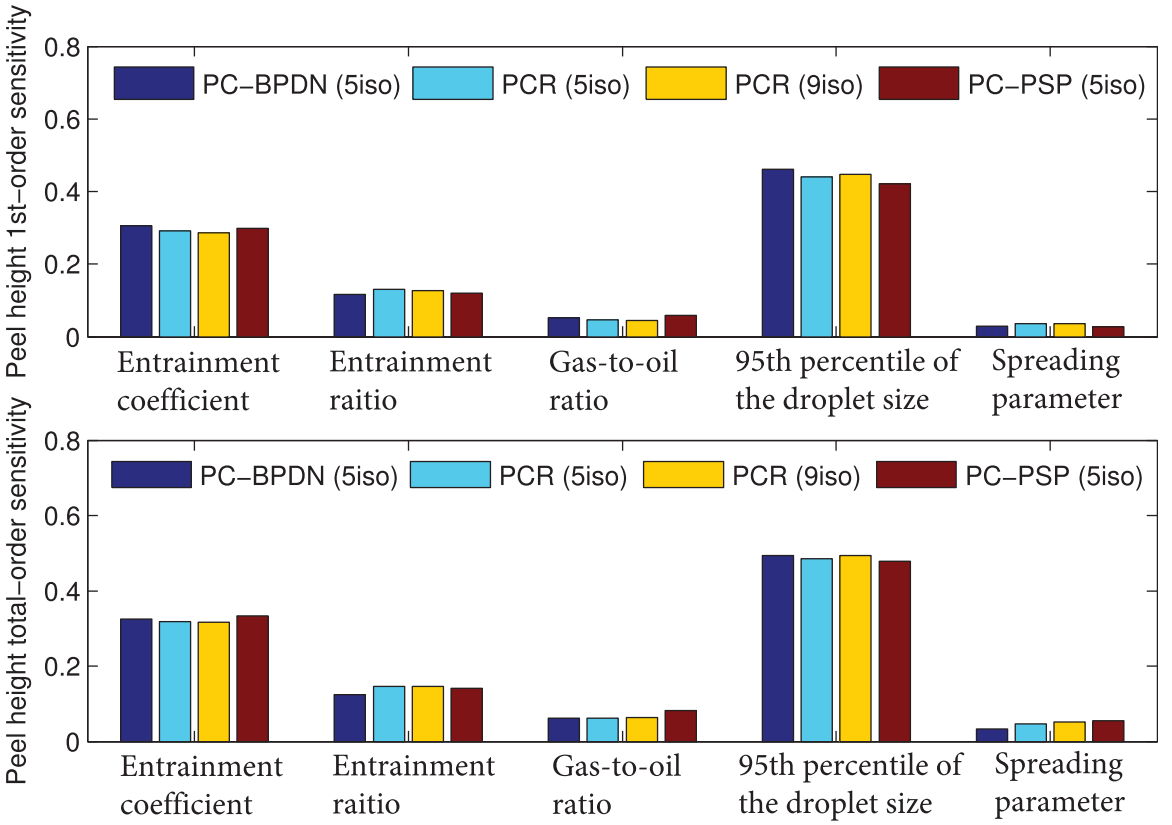

Figure 7. First-order and total-order sensitivity of peel height to the five uncertain inputs. Colors indicate the different polynomial chaos surrogates, bar heights indicate sensitivity, and group labels indicate the input variable.

of problems. The two model outputs analyzed here are the trap and peel heights. Both quantities exhibit small-scale noise according to the one-dimensional plots in Figure 1; however, all proxies tested here perform better at representing the peel height than the trap height. The discrepancies between the different proxies are most visible in the PDF of the trap height where, for low-order polynomials, the shoulder and mode of the reference trap height PDF are either missed or barely present. Increasing the maximum polynomial degree to 9 allows the regression-based PC to represent these features better at the expense of increasing the ensemble size substantially. These errors in representing the PDF however do not seem to affect the sensitivity calculation in substantial ways. A number of additional general remarks can be made.

Projection-based PC methods minimize $\langle \epsilon, \epsilon \rangle$, dispense with the inversion of a linear system of equation, sample the model at quadrature points, and do not suffer from ill conditioning because of the orthogonality of the basis functions. Different quadrature rules can be used to approximate the projection integral, most notably the pseudospectral technique of *Constantine et al*. [2012b] and *Conrad and Marzouk* [2013] compute the coefficients incrementally and truncate the series adaptively by focusing the sampling along directions contributing the most to the variance (see also Winokur et al. [2013] for an example application). Quadrature rules, however, are not very flexible, as they require model runs at a set of prescribed points, many of which are near the edges of the uncertainty space, and the whole set may have to be discarded if a single sampling point falls outside the operating regime of the numerical model. Finally, the Galerkin projection does not account for model noise and may incur a high cost in representing inconsequential model noise if proper precautions are not taken. The projection technique exhibited the largest approximation errors among the different techniques presented here.

Least-squares-based PC methods minimize the algebraic residual norm $||\epsilon||_{LS}$, and allow more freedom in the selection of sampling points since the latter do not have to be quadrature points. The sampling points, however, must be chosen carefully so as to avoid severe ill conditioning. Least squares approaches are more useful in cases where the model includes algorithmic errors and/or stochastic processes. In these

cases, a projection approach would be suboptimal as the higher-order coefficients would grow to accommodate noise. In the present work, the regression approaches proved to be more useful than the projection one because of the noisy model output.

PC-BPDN selects the sparsest basis set that can represent the response within a specified tolerance. This tolerance can be obtained from the model samples using cross validation, a process that, though computationally expensive, can guard against the dangers of overfitting the surrogate to model noise. One concern with BPDN is in specifying the basis set on which the sparsity will be enforced. The PC-BPDN provided the best overall approach as it delivered the smallest errors for the least amount of work. Furthermore, its orthogonal bases allowed a straightforward calculation of the sensitivity indices.

The GPR approach also accounts for the presence of model noise but uses Bayesian principles to construct the model surrogate. Its sampling strategy is quite flexible, as it is not constrained by quadrature requirements. Most importantly, GPR delivers estimates of the mean and variance (uncertainty) of the response $u(\boldsymbol{\xi})$ at all points. From a practical point of view, GPR does not concern itself with the specification and truncation of a basis set, rather it requires the specification of mean and covariance functions and their hyperparameters. Although a variety of covariance functions are available to model quite general surfaces the squared exponential covariance functions have proven extremely useful in a large number of applications, even when the hyperparameters are not optimal. Furthermore, relatively sophisticated covariance functions can be constructed by combining simpler ones in order to model quite complicated response surfaces; for example, a Matérn class could have been added to represent the characteristics of the small-scale noise [*Rasmussen and Williams*, 2006].

In conclusion, PCE and GPR methods provide an efficient way for propagating the input uncertainties, making it possible to compute the full probabilistic descriptions of model outputs. Verification using a large Monte Carlo ensemble of simulations confirms their usefulness in delivering accurate surrogates and sensitivity analyses in spite of discrepancies in estimating the trap height PDF. The examples in this paper illustrate a great gain in computational efficiency using indirect sampling methods compared with direct Monte Carlo sampling. The UQ procedures discussed here have been used to quantify uncertainties in a number of oceanic applications [*Thacker et al*., 2012; *Alexanderian et al*., 2012; *Sraj et al*., 2013, 2014; *Chen et al*., 2015]. The articles of *Gonçalves et al*. [2016] and Wang 2016 focus on their application in oil plume and oil-fate modeling while Iskandarani 2016 focus on quantifying uncertainties in oceanic forecasts in the Gulf of Mexico due to initial condition uncertainties.

Appendix A: Multidimensional Basis Functions

Multidimensional-polynomial basis functions $\psi_k(\boldsymbol{\xi})$ are products of orthogonal polynomials P_{m_d} of the individual uncertain inputs ξ_d:

$$\psi_k(\boldsymbol{\xi}) = P_{m_1}(\xi_1) P_{m_2}(\xi_2) P_{m_3}(\xi_3) P_{m_4}(\xi_4) P_{m_5}(\xi_5)\,. \tag{A1}$$

Note that the index k of the multidimensional basis $\psi_k(\boldsymbol{\xi})$ can be represented collectively by the degrees of each polynomial factor: $(m_1, m_2, m_3, m_4, m_5)$; in practice, these five numbers are mapped to the sequential index k ordered so that the degrees of the terms in (A1) increase monotonically and information relating the individual degrees of the factors to this index allows both to be accessed as needed. The multivariate degree of $\psi_k(\boldsymbol{\xi})$ is simply the sum of the degrees of its univariate factors: $\sum_{d=1}^{5} m_d$.

While truncation of the multidimensional expansion to eliminate terms with high degree strongly restricts the univariate polynomial factors, the number of possible combinations increases rapidly with increasing univariate degrees. For example, if all univariate factors have degrees no greater than M, total-order (triangular) truncation retains $P+1=(M+5)!/M!5!$ terms, while maximum-order truncation retains $(M+1)^5$. The nature of the truncation is influenced strongly by the particular method used for constructing the polynomial chaos surrogate. Adaptive pseudospectral construction of *Conrad and Marzouk* [2013], for example, allows the basis

Table 5. Some Common Input PDFs and Their Associated Orthogonal Polynomial Bases

$p(\xi)$	Orthogonal Polynomial $\psi_n(\xi)$
Gaussian	Hermite
Gamma	Laguerre
Beta	Jacobi
Uniform	Legendre

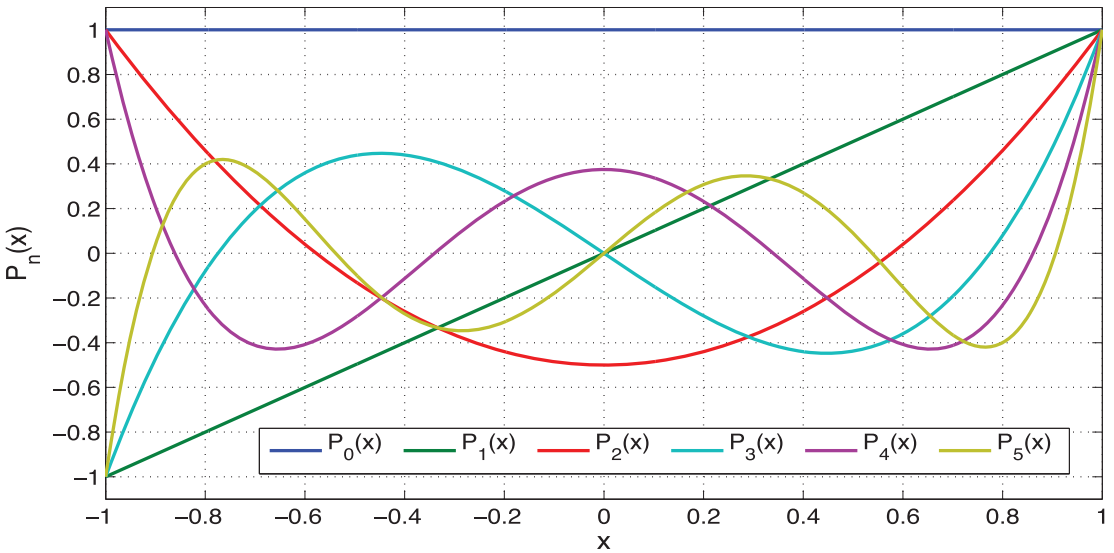

Figure 8. Legendre polynomials of degree 0–5.

polynomials for different input variables to have different maximal degrees, which are selected automatically according to their contributions to the variance; see e.g., Winokur et al. [2013] for an example application. And basis pursuit denoising generally discards some terms with lower degree than others that are retained.

For independent random input variables, the joint probability density function is the product of the univariate densities p_d for the individual inputs ξ_d: $p(\boldsymbol{\xi}) = \prod_{d=1}^{5} p_d(\xi_d)$. Hence the multidimensional basis functions inherit their orthogonality properties. The orthogonal univariate polynomials associated with given distributions are listed in Table 5. Legendre polynomials, which are basis functions for a uniform distribution $p(\xi) = \frac{1}{2}$ with $|\xi| \leq 1$, are shown in Figure 8.

Appendix B: Statistical Properties of PC Series

The statistics of the uncertain output can be calculated straightforwardly with polynomial series that are orthogonal with respect to their pdf. The expectation or the mean of the model output is

$$\mathbb{E}[u] = \int u(\boldsymbol{\xi})\, p(\boldsymbol{\xi})\ \mathrm{d}\boldsymbol{\xi} = \langle u, \psi_0 \rangle = \langle u_P, \psi_0 \rangle + \langle \epsilon_P, \psi_0 \rangle = \hat{u}_0, \tag{B1}$$

since $\psi_0(\boldsymbol{\xi})$ is just a constant for all orthogonal polynomials that can be scaled such that its norm $\langle \psi_m, \psi_m \rangle = 1$, and the orthogonality property has been invoked to arrive to the last equality in equation (B1). The mean is then simply the first coefficients $\hat{u}_0$. Likewise, the variance of u can also be derived:

$$\mathbb{E}[(u - \mathbb{E}[u])^2] = \sum_{m,n=1}^{\infty} \hat{u}_m \hat{u}_n \int \psi_m \psi_n p(\boldsymbol{\xi})\ \mathrm{d}\boldsymbol{\xi} = \sum_{m,n=1}^{\infty} \hat{u}_m \hat{u}_n \langle \psi_m, \psi_n \rangle, \tag{B2}$$

$$= \sum_{n=1}^{P} \hat{u}_n^2 ||\psi_n||^2 + \underbrace{\sum_{m=P+1}^{\infty} \hat{u}_n^2 ||\psi_n||^2}_{\text{truncation errors}}. \tag{B3}$$

The mean of the model output is unaffected by truncation when the inner product is evaluated exactly; it does however include aliasing errors when quadrature rules are used to approximate the inner products. The truncated series underestimates the variance but the error remains small if the amplitude of the higher-order terms is small. Monitoring the variance is hence a good indicator of whether the series includes enough terms or not.

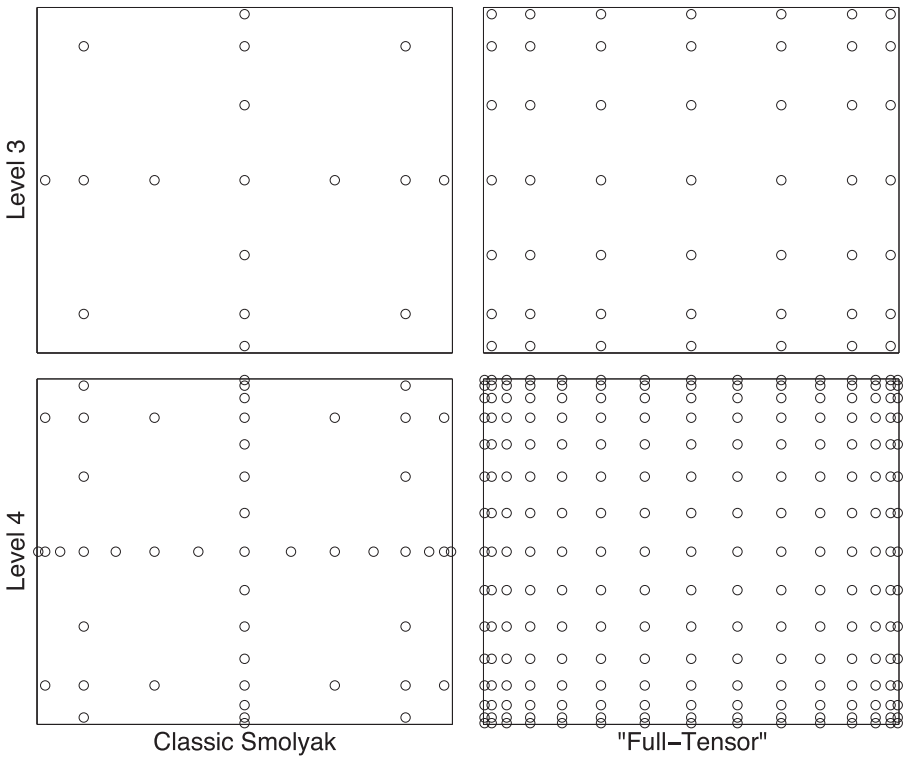

Figure 9. Comparison of quadrature nodes between classic Smolyak sparse construction and "Full-Tensor" construction. Quadrature is built with Gauss-Kronrod-Patterson 1-D rule.

Appendix C: Numerical Quadrature

Quadratures are numerical approximations to integrals of the form:

$$I = \int_a^b f(\xi) p(\xi) d\xi \approx Q_l^1 f = \sum_{i=1}^{S} f(\xi_i)\omega_i$$

where S is the number of quadrature points, and ω_i and ξ_i are the quadrature weights and quadrature roots associated with the weight function $p(\xi)$; Q stands for the quadrature operator. The traditional Gauss quadrature is well known for maximizing the accuracy of the integration per sample evaluation of f: if the quadrature roots and weights are chosen to be those of the orthogonal polynomials associated with the weight function $p(\xi)$, then an S-point quadrature would integrate exactly polynomials of degree $2S-1$, conversely a polynomial of degree $2M$ would be integrated exactly using a Gauss rule using $S > M+\frac{1}{2}$ points. Multidimensional quadrature can be done using tensor products of 1-D rules.

$$(Q_{l_1}^1 \otimes \ldots \otimes Q_{l_d}^1) f := \sum_{i_1=1}^{S_{l_1}^1} \ldots \sum_{i_d=1}^{S_{l_d}^1} \omega_{l_1,i_1} \ldots \omega_{l_d,i_d} f(\xi_{l_1,i_1}, \ldots, \xi_{l_d,i_d})$$

where Q_{l_i} is the 1-D quadrature operator using l_i samples in the ith variable, S_{l_i} is the number of quadrature points ω_{l_d,i_d} is the quadrature weights associated with different dimensions. There are a number of disadvantages to Gauss quadrature which include: the roots cluster near the edge of the interval to maximize the order, this is equivalent to sampling the edges of the uncertainty interval; the roots are not nested so the users cannot reuse the old quadrature samples if the number of sample points is increased after a first calculation is performed; and finally the number of quadrature points increases exponentially with the number of uncertain variables.

Fortunately, sparse nested quadrature rules can be constructed so that adaptive integration with reuse of results can be performed, and so that the curse of dimensionality can be mitigated [*Gerstner and Griebel*, 1998].

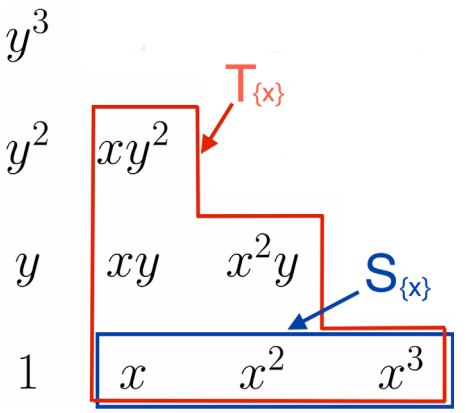

Figure 10. The basis functions contributing to S_i are within the blue zone whereas those contributing to the total sensitivity are within the red zone. The different bases are represented by the highest degree in x and y.

The starting point is 1-D nested quadrature at level l: $Q_l f = \sum_{i=1}^{S_l} f(\xi_{l,i})\omega_{l,i}$ such as Gauss-Kronrod-Patterson rules, where $S_l = 2^l - 1$ is the number of quadrature points, and the quadrature points of level l include all the quadrature points in level $l-1$. Smolyak construction capitalizes on the generally weak effect of high-order mixed terms, such as $\xi_1^M \xi_2^M$, to reduce the number of function evaluations needed. The Smolyak quadrature construction is a telescoping sum of different level of quadrature rules. Figure 9 is a comparison of full tensor product tensorization and classic Smolyak sparse grid construction.

Appendix D: Sensitivity Indices

The analysis of variance is based on sensitivity indices as defined in *Homma and Saltelli* [1996b], but the orthogonal nature of PC series with respect to the input PDF makes this analysis extremely simple. First notice that the total variance in a PC expansion is given by the sum of positive quantities (equation (B3)) with each coefficient in the series contributing a single term in the variance estimate. Thus, it would be possible to attribute the variance contributed by a combination of uncertain parameters by forming different partial sums of these same terms. The process is made simple by the mapping that keeps track of the degree of the polynomial along each dimension

$$k \leftrightarrow (m_1^k, m_2^k, \ldots, m_d^k, \ldots, m_D^k), \tag{D1}$$

where $k = 0, \ldots, P$ and m_d^k refers to the 1-D polynomial degree in the dth variable in the kth term of the expansion.

The variance contributed by the dth variable alone, for example, would correspond to a partial sum where $m_d^k > 0$ and $m_j^k = 0$ for $j \neq d$. If the set $\mathcal{K}_d$ designates the collection of multi-indices that satisfies this relationship then this contribution is:

$$V_d[u] = \sum_{k \in \mathcal{K}_d} \hat{u}_k^2 \langle \psi_k, \psi_k \rangle. \tag{D2}$$

The first-order sensitivity index $S_{\{d\}}$ which measures the fraction of the variance due ξ_d alone is then

$$S_d[u] = \frac{V_d[u]}{\sum_{k=1}^{P} \hat{u}_k^2 \langle \psi_k, \psi_k \rangle} \quad \text{with} \quad \sum_{d=1}^{D} S_d[u] \leq 1. \tag{D3}$$

The variance contributed by the dth variable in conjunction with other variables involves a different partial sum. Let $\mathcal{T}_d$ designates the set of basis where $m_d^k > 0$, then the variance caused by the ξ_d is

$$V_{\{1,\ldots,D\}}[u] = \sum_{k \in \mathcal{T}_d} \hat{u}_k^2 \langle \psi_k, \psi_k \rangle, \tag{D4}$$

and the total sensitivity index is the ratio:

$$T_d[u] = \frac{V_{\{1,\ldots,D\}}[u]}{\sum_{k=1}^{P} \hat{u}_k^2 \langle \psi_k, \psi_k \rangle} \quad \text{with} \quad \sum_{d=1}^{D} T_d[u] \geq 1. \tag{D5}$$

Figure 10 shows a sketch of the basis functions whose coefficients are involved in computing the first-order and total sensitivity indices for a two-dimensional case.

Acknowledgments
This research was made possible in part by a grant from BP/The Gulf of Mexico Research Initiative to the CARTHE and DEEP-C Consortia and by the Office of Naval Research, award N00014-101-0498. J. Winokur and O. M. Knio were also supported in part by the U.S. Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research, under award DE-SC0008789. This research was conducted in collaboration with and using the resources of the University of Miami Center for Computational Science. The source code for the model used in this study, TAMOC, is freely available at https://github.com/socolofs/tamoc. The data and input files necessary to reproduce the experiments are available from the authors upon request (swang@rsmas.miami.edu). The data are archived at https://github.com/Shitao/A-comparison-of-uncertainty-quantification-techniques-using-integral-plume-model.

References

Alexanderian, A., J. Winokur, I. Sraj, A. Srinivasan, M. Iskandarani, W. C. Thacker, and O. M. Knio (2012), Global sensitivity analysis in an ocean general circulation model: A sparse spectral projection approach, *Comput. Geosci.*, *16*(3), 757–778.

Bhaumik, T. (2005), Numerical modeling of multiphase plumes: A comparative study between two-fluid and mixed-fluid integral models, PhD thesis, Texas A&M Univ., College Station.

Blatman, G. (2009), Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis, Universite Blaise Pascal - Clermont-Ferrand II, PhD thesis, Clermont-Ferrand France.

Canuto, C. G., M. Y. Hussaini, A. Quarteroni, and T. A. Zang (2006), *Spectral Methods: Fundamentals in Single Domains, Scientific Computation*, 581 pp., Springer, Berlin, Heidelberg.

Chen, G., B. P. Kirtman, and M. Iskandarani (2015), An efficient perturbed parameter scheme in the Lorenz system for quantifying model uncertainty, *Q. J. R. Meteorol. Soc.*, *141*, 2552–2562, doi:10.1002/qj.2541.

Conrad, P. R., and Y. M. Marzouk (2013), Adaptive Smolyak pseudospectral approximations, *SIAM J. Sci. Comput.*, *35*(6), A2643–A2670.

Constantine, P. G., M. S. Eldred, and E. T. Phipps (2012a), Sparse pseudospectral approximation method, *Comput. Methods Appl. Mech. Eng.*, *229*, 1–12.

Constantine, P. G., M. S. Eldred, and E. T. Phipps (2012b), Sparse pseudospectral approximation method, *Comput. Methods Appl. Mech. Eng.*, *28*(24), 11,751–11,762, doi:10.1016/j.cma.2012.03.019.

Crestaux, T., O. P. Le Ma itre, and J.-M. Martinez (2009), Polynomial chaos expansion for sensitivity analysis, *Reliab. Eng. Syst. Safety*, *94*(7), 1161–1172.

Crounse, B. C., E. J. Wannamaker, and E. E. Adams (2007), Integral model of a multiphase plume in quiescent stratification, *J. Hydraul. Eng.*, *133*(1), 70–76.

Daley, R. (1993), *Atmospheric Data Analysis*, 457 pp., Cambridge Univ. Press, Cambridge, U. K.

Doostan, A., and H. Owhadi (2011), A non-adapted sparse approximation of pdes with stochastic inputs, *J. Comput. Phys.*, *230*(8), 3015–3034.

Ganapathysubramanian, B., and N. Zabaras (2007), Sparse grid collocation schemes for stochastic natural convection problems, *J. Comput. Phys.*, *225*(1), 652–685, doi:10.1016/j.jcp.2006.12.014.

Gerstner, T., and M. Griebel (1998), Numerical integration using sparse grids, *Numer. Algorithms*, *18*(3-4), 209–232.

Gerstner, T., and M. Griebel (2003), Dimension–adaptive tensor–product quadrature, *Computing*, *71*(1), 65–87.

Ghanem, R. G., and P. D. Spanos (1991), *Stochastic Finite Elements: A Spectral Approach*, vol. 41, Springer, N. Y.

Gonçalves, R., M. Iskandarani, A. Srinivasan, C. Thacker, E. Chassignet, and O. M. Knio (2016), A framework to quantify uncertainty in simulations of oil transport in the ocean, *J. Geophys. Res. Oceans*, doi:10.1002/2015JC011311, in press.

Hampton, J., and A. Doostan (2015), Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression, *Comput. Methods Appl. Mech. Eng.*, *290*, 73–97, doi:10.1016/j.cma.2015.02.006.

Homma, T., and A. Saltelli (1996a), Importance measures in global sensitivity analysis of nonlinear models, *Reliab. Eng. Syst. Safety*, *52*(1), 1–17, doi:10.1016/0951-8320(96)00002-6.

Homma, T., and A. Saltelli (1996b), Importance measures in global sensitivity analysis of nonlinear models, *Reliab. Eng. Syst. Safety*, *52*(1), 1–17.

Huan, X. (2010), Accelerated Bayesian experimental design for chemical kinetic models, Master's thesis, Mass. Inst. of Technol., Cambridge.

Johansen, O. I., H. Rye, and A. Melbye (2001), Deep spill JIP-Experimental discharges of gas and oil at Helland Hansen–June 2000, *SINTEF Rep. 5TF66 F01082*, pp. 1–159, SINTEF Applied Chemistry, Trondheim, Norway.

Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation, and Predictability*, 341 pp., Cambridge Univ. Press, N. Y.

Klimke, A. (2006a), Uncertainty modeling using fuzzy arithmetic and sparse grids, PhD thesis, Univ. Stuttgart, Stuttgart, Germany.

Klimke, A. (2006b), Sparse Grid Interpolation Toolbox—User's guide, *Tech. Rep. IANS Doc. 2006/001*, Univ. Stuttgart, Stuttgart, Germany.

Kullback, S. (1959), *Information Theory and Statistics*, John Wiley, N. Y.

Lefebvre, A. (1988), *Atomization and Sprays*, 434 pp., CRC Press.

Le Maître, O. P., and O. Knio (2010), *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Springer, N. Y.

Le Maître, O. P., M. T. Reagan, H. N. Najm, R. G. Ghanem, and O. M. Knio (2002), A stochastic projection method for fluid flow: II. Random process, *J. Comput. Phys.*, *181*(1), 9–44, doi:10.1006/jcph.2002.7104.

Najm, H. N. (2009), Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics, *Annu. Rev. Fluid Mech.*, *41*, 35–52, doi:10.1146/annurev.fluid.010908.165248.

Narayan, A., and T. Zhou (2015), Stochastic collocation on unstructured multivariate meshes, *Commun. Comput. Phys.*, *18*(1), 1–36, doi:10.4208/cicp.020215.070515a.

Peng, J., J. Hampton, and A. Doostan (2014), A weighted 1-minimization approach for sparse polynomial chaos expansions, *J. Comput. Phys.*, *267*, 92–111.

Rasmussen, C. E., and C. K. I. Williams (2006), *Gaussian Processes for Machine Learning*, vol. 14, MIT Press, Cambridge, doi:10.1142/S0129065704001899.

Reddy, C. M., et al. (2012), Composition and fate of gas and oil released to the water column during the Deepwater Horizon oil spill, *Proc. Natl. Acad. Sci. U. S. A.*, *109*(50), 20,229–20,234.

Saltelli, A. (2002), Making best use of model evaluations to compute sensitivity indices, *Comput. Phys. Commun.*, *145*(2), 280–297, doi:10.1016/S0010-4655(02)00280-1.

Sargsyan, K., B. Debusschere, H. N. Najm, and Y. Marzouk (2009), Bayesian inference of spectral expansions for predictability assessment in stochastic reaction networks, *J. Comput. Theor. Nanosci.*, *6*(10), 2283–2297, doi:10.1166/jctn.2009.1285.

Shin, Y., and D. Xiu (2016), Nonadaptive quasi-optimal points selection for least squares linear regression, *SIAM J. Sci. Comput.*, *38*(1), A385–A411, doi:10.1137/15M1015868.

Socolofsky, S. A., T. Bhaumik, and D.-G. Seol (2008), Double-plume integral models for near-field mixing in multiphase plumes, *J. Hydraul. Eng.*, *134*(6), 772–783.

Sraj, I., M. Iskandarani, A. Srinivasan, W. C. Thacker, J. Winokur, A. Alexanderian, C.-Y. Lee, S. S. Chen, and O. M. Knio (2013), Bayesian inference of drag parameters using AXBT data from Typhoon Fanapi, *Mon. Weather Rev.*, *141*(7), 2347–2367, doi:10.1175/MWR-D-12-00228.1.

Sraj, I., M. Iskandarani, W. C. Thacker, A. Srinivasan, and O. M. Knio (2014), Drag parameter estimation using gradients and hessian from a polynomial chaos model surrogate, *Mon. Weather Rev.*, *142*, 933–941, doi:10.1175/MWR-D-13-00087.1.

Sraj, I., K. T. Mandli, O. M. Knio, C. N. Dawson, and I. Hoteit (2014), Uncertainty quantification and inference of manning friction coefficients using DART buoy data during the Tōhoku tsunami, *Ocean Modell.*, *83*, 82–97, doi:10.1016/j.ocemod.2014.09.001.

Tatang, M. A., W. Pan, R. G. Prinn, and G. J. McRae (1997), An efficient method for parametric uncertainty analysis of numerical geophysical models, *J. Geophys. Res.*, *102*(D18), 21,925–21,932, doi:10.1029/97JD01654.

Thacker, W. C., A. Srinivasan, M. Iskandarani, O. M. Knio, and M. L. Hénaff (2012), Propagating boundary uncertainties using polynomial expansions, *Ocean Modell.*, *43*, 52–63.

Thacker, W. C., M. Iskandarani, R. C. Gonalves, A. Srinivasan, and O. M. Knio (2015), Pragmatic aspects of uncertainty propagation: A conceptual review, *Ocean Modell.*, *95*, 25–36, doi:10.1016/j.ocemod.2015.09.001.

Valentine, D. L., et al. (2010), Propane respiration jump-starts microbial response to a deep oil spill, *Science*, *330*(6001), 208–211.

van den Berg, E., and M. P. Friedlander (2007), *SPGL1: A Solver For Large-Scale Sparse Reconstruction*, British Columbia, Canada. [Available at http://www.cs.ubc.ca/labs/scl/spgl1.]

van den Berg, E., and M. P. Friedlander (2008), Probing the pareto frontier for basis pursuit solutions, *SIAM J. Sci. Comput.*, *31*(2), 890–912, doi:10.1137/080714488.

Webster, M. D., and A. P. Sokolov (2000), A methodology for quantifying uncertainty in climate projections, *Clim. Change*, *46*(4), 417–446, doi:10.1023/A:1005685317358.

Weinert, H. L. (Ed.) (1982), *Reproducing Kernel Hilbert Spaces—Applications in Statistical Signal Processing, Benchmark Pap. Electr. Eng. Comput. Sci.*, vol. 25, 654 pp., Hutchinson Ross, Stroudsburg, Pa.

Wiener, N. (1938), The homogeneous chaos, *Am. J. Math.*, *60*(4), 897–936.

Winokur, J., P. Conrad, I. Sraj, O. Knio, A. Srinivasan, W. C. Thacker, Y. Marzouk, and M. Iskandarani (2013), A priori testing of sparse adaptive polynomial chaos expansions using an ocean general circulation model database, *Comput. Geosci.*, *17*(6), 899–911.

Xiu, D., and G. E. Karniadakis (2002), The Wiener–Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, *24*(2), 619–644.

Zhou, T., A. Narayan, and Z. Xu (2014), Multivariate discrete least-squares approximations with a new type of collocation grid, *SIAM J. Sci. Comput.*, *36*(5), A2401–A2422, doi:10.1137/130950434.