Surrogate Modeling for Uncertainty Assessment with Application to Aviation Environmental System Models

D. Allaire,* K. Willcox†

Department of Aeronautics and Astronautics

Massachusetts Institute of Technology, Cambridge, MA, 02139

Numerical simulation models to support decision-making and policy-making processes are often complex, involving many disciplines, many inputs, and long computation times. Inputs to such models are inherently uncertain, leading to uncertainty in model outputs. Characterizing, propagating, and analyzing this uncertainty is critical both to model development and to the effective application of model results in a decision-making setting; however, the many thousands of model evaluations required to sample the uncertainty space (e.g. via Monte Carlo sampling) present an intractable computational burden. In this paper we present a novel surrogate modeling methodology designed specifically for propagating uncertainty from model inputs to model outputs and for performing a global sensitivity analysis—which characterizes the contributions of uncertainties in model inputs to output variance—while maintaining the quantitative rigor of the analysis by providing confidence intervals on surrogate predictions. The approach is developed for a general class of models and is demonstrated on an aircraft emissions prediction model that is being developed and applied to support aviation environmental policy-making. The results demonstrate how the confidence intervals on surrogate predictions can be used to balance the tradeoff between computation time and uncertainty in the estimation of the statistical outputs of interest.

Nomenclature

AEDT	Aviation Environmental Design Tool
AEM	Aircraft Emissions Module
$\mathbb{B}$	Borel σ-field
CO	Carbon Monoxide
CO_2	Carbon Dioxide

---

*Postdoctoral Associate, Member AIAA, dallaire@mit.edu
†Associate Professor of Aeronautics and Astronautics, Associate Fellow AIAA, kwillcox@mit.edu

American Institute of Aeronautics and Astronautics

D	Output variance
D_i	Single factor partial variance for factor i
$\mathbb{E}[y_k]$	Expected value of NO_x emissions from operation k, gm
f	Any Borel-measurable function
$\mathcal{G}$	Group of inputs of interest
g	Emissions index of NO_x, gm NO_x/kg fuel
$\mathcal{N}(\alpha, \beta)$	Normal distribution with mean, α, and variance β
n	Dimension of the input space
N	Number of model evaluations in a Monte Carlo simulation
n_o	Number of operations in the AEM surrogate model
N_o	Total number of AEM operations
NO_x	Oxides of Nitrogen
N_s	Number of flight segments in an operation
$\mathcal{O}$	Subset of n_o operations in surrogate model
$\mathbb{P}$	Probability measure
$\mathcal{Q}$	Subset of N_q random variables in generic surrogate model
q	Fuel burn input, kg
r	Temperature input, K
$REINO_x$	Reference Emissions Index of Oxides of Nitrogen, gm NO_x/kg fuel
s	Pressure input, N/m^2
S_i	Main effect sensitivity index for factor i
SO_x	Oxides of Sulfur
t	Relative humidity input
u	Fuel flow input, kg/s
v	$REINO_x$ input, gm NO_x/kg fuel
$\mathbf{x}$	Generic vector of random variables
x_i	Random variable that defines input factor i
x_i^m	mth sample from the random variable x_i
y_{kl}	NO_x emissions produced by flight segment l of operation k, gm
Y	Generic output of interest
y_k	NO_x emissions produced by operation k, gm
y_{tot}	Total NO_x emissions, gm
z_k	Generic kth constituent part of a model

$\mu_{\mathbb{E}}$	Expected value of the distribution of operation-level NO_x emissions expected values, gm
$\mu_{\mathbb{E}[z_k]}$	Expected value of the distribution of expected values of the z_k
μ_{σ^2}	Expected value of the distribution of operation-level NO_x emissions variances, gm^2
$\mu_{\mathrm{var}(z_k)}$	Expected value of the distribution of variances of the z_k
$\sigma_{\mathbb{E}}^2$	Variance of the distribution of operation-level NO_x emissions expected values, gm^2
$\sigma_{\mathbb{E}[z_k]}^2$	Variance of the distribution of expected values of z_k
$\sigma_{\sigma^2}^2$	Variance of the distribution of operation-level NO_x emissions variances, gm^4
$\sigma_{\mathrm{var}(z_k)}^2$	Variance of the distribution of variances of the z_k
$\sigma_{y_k}^2$	Variance of NO_x emissions from operation k, gm^2
τ_i	Total effect sensitivity index for factor i

I.   Introduction

Numerical simulation models to support decision-making and policy-making processes, while becoming increasingly widespread, typically have uncertainty associated with their inputs, leading to uncertainty in model outputs. Effective application of model results to decision-making and in support of model development require proper characterization, propagation, and analysis of that uncertainty. The process of propagating the uncertainty from inputs to outputs, for example via Monte Carlo simulation, could require many thousands of model evaluations, thus presenting an intractable computational burden. Here we present a novel surrogate modeling methodology based on invoking the Central Limit Theorem, which is designed specifically for propagating uncertainty from model inputs to model outputs and for performing a global sensitivity analysis—which characterizes the contributions of uncertainties in model inputs to output variance—while maintaining the quantitative rigor of the analysis by providing confidence intervals on surrogate predictions.

Our approach is developed for a general class of models where the application of uncertainty propagation and global sensitivity analysis on a full model is computationally impractical, and demonstrated on the specific case of the Aircraft Emissions Model (AEM) of the Aviation Environmental Design Tool (AEDT). The computational models of AEDT are being developed and applied to support aviation environmental policy-making by providing the capability to characterize and quantify interdependencies among aviation-related noise and emissions, impacts on health and welfare, and industry and consumer costs, under different policy, technology, operational, and market scenarios. A key priority is to inform the analyses conducted by these tools with associated uncertainty from the inputs and assumptions used in the analysis process. The scale and complexity of these analyses are immense; for example, a single simulation of a one-year analysis

involves over thirty million flight operations with 350 aircraft types and thousands of inputs, analyzed with computationally intensive models spanning airline economics, environmental economics, aircraft operations, aircraft performance and emissions, noise, local air quality, and global climate. Thus, the propagation and analysis of uncertainty in such models with a method such as Monte Carlo simulation, which in some cases can take several thousand model evaluations, is computationally impractical.

Surrogate models that provide substantial computational speedups are therefore crucial to the process of uncertainty analysis in tools such as AEDT. However, quantifying the impacts on the analyses due to exercising a surrogate in place of the full model is essential to producing defensible claims in the context of decision-making. While surrogate modeling methodologies have been successfully applied in many settings, a key challenge here is the derivation of surrogate models—and the associated confidence in uncertainty and sensitivity analyses conducted with the surrogate models—for large-scale complex system models with high-dimensional input spaces.

This paper proposes a systematic method to reduce the complexity and computational cost of a general class of large-scale models in such a way that input uncertainty may still be quantified and analyzed. The method is applied to the AEM, which is designed to estimate global emissions from aviation. Section II presents background on the uncertainty analysis methods employed, and describes the structure of the general class of problems considered. The methodology, described in Section III, focuses on the creation of a hierarchical surrogate model for the general model class, by selecting a small subset of inputs to represent the large-scale complex system. These representative inputs form a surrogate model with which an inexpensive computation can be performed in place of the originally expensive computation. In Section IV the method is applied to create surrogates for the AEM. We demonstrate how these surrogates can be used for both uncertainty and sensitivity analysis with rigorous confidence intervals on surrogate predictions. Limitations and additional sources of error are discussed in Section V, and conclusions are drawn in Section VI.

II. Background

In the context of numerical simulation tools, *uncertainty analysis* encompasses the process of characterizing and analyzing the effects of uncertainty in model inputs, with a focus on quantitative assessment of the effects on model outputs and thus, on the conclusions drawn from simulation results. *Sensitivity analysis* studies how variability in model outputs can be apportioned to sources of uncertainty in model factors.[1] To carry out uncertainty and sensitivity analyses for large-scale numerical models requires first an understanding of the purpose of the analyses and the way in which quantitative results will be employed for decision-making. Second, knowledge of the character/structure of the underlying model is important to determine appropriate analysis methods and for an appreciation of the associated computational complexity,

which may mandate the use of surrogate models. Background on each of these areas is given in the following subsections.

II.A. Uncertainty Analysis and Sensitivity Analysis

A detailed overview of both deterministic and statistical methods for uncertainty and sensitivity analysis of large-scale systems is presented in Ref. 1. Here, we consider two general applications of uncertainty and sensitivity analysis. The first is to support decision-making, for which uncertainty analysis should provide the ability to compare various scenarios (e.g. different policies, different input assumptions, etc.) in terms of output means, output variances and other distributional information that may be used to help make a decision. The second application is to help further model development. In this second case, the primary goal is a sensitivity analysis that apportions model output variability to model factors[2–4] to help determine where future research and development efforts should focus.

The computation of model output means, output variances and other distributional information in support of uncertainty analysis for decision-making can be carried out with Monte Carlo simulation. We consider a general model $f(\mathbf{x})$, where $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ is the vector of n inputs to the model. If the model inputs are viewed as random variables with some associated probability distribution, then the mean value of the model output can be computed from a Monte Carlo simulation as

$$\frac{1}{N} \sum_{m=1}^{N} f(\mathbf{x}^m) \to \mathbb{E}[f(\mathbf{x})] \text{ as } N \to \infty \ , \tag{1}$$

where N is the number of model evaluations in the Monte Carlo simulation and $\mathbf{x}^m = [x_1^m, x_2^m, \ldots, x_n^m]^T$ denotes the mth sample realization of the random vector $\mathbf{x}$. Convergence of the sample mean in (1) to the expected value of $f(\mathbf{x})$ is guaranteed by the law of large numbers and the convergence rate is $1/\sqrt{N}$, as given by the Central Limit Theorem.[5] Output variances and other distributional quantities can similarly be computed using Monte Carlo simulation results. The process of computing such quantities requires a large number of model evaluations, which for computationally intensive models is in many cases impractical; hence the need for developing surrogate models for this type of uncertainty analysis.

For model development purposes, global sensitivity analysis is a rigorous method for quantitatively apportioning output variance.[2] The goal of a global sensitivity analysis is shown notionally in Figure 1, where the pie represents the variance in a model output, which is then broken out according to factor contributions. The results of a global sensitivity analysis permit a ranking of model factors that can be used in different development settings such as *factor prioritization* for future research, where the goal is to determine which factors, once fixed will cause the largest reduction in variance, and *factor fixing*, for which the goal is to identify noninfluential factors that may be fixed without substantially affecting model outputs.[6]
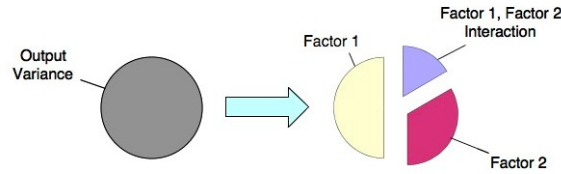
American Institute of Aeronautics and Astronautics

Figure 1. Apportioning Output Variance

The process of apportioning output variance across model factors in a global sensitivity analysis can be carried out rigorously by both a Fourier Amplitude Sensitivity Test (FAST) method, and the Sobol' method.[2,3,7,8] The FAST method is based on Fourier transforms, while the Sobol' method utilizes Monte Carlo simulation. Owing to its ease of implementation, the Sobol' method is employed in this work. It should be noted here that other techniques, such as the method of elementary effects due to Morris,[9] can approximate the results of the FAST and Sobol' methods, and in the case of the Morris method, provide a good proxy to the total sensitivity index discussed below.[6] However, in this work we have focused on using rigorous statistical methods on surrogate models rather than on using approximate statistical methods on surrogate models or full models.

The Sobol' method for computing global sensitivity indices was proposed by Russian mathematician I.M. Sobol'. The method is well-developed and in wide use in the sensitivity analysis field, particularly by the Joint Research Centre of the European Commission.[2,6,10] The method is discussed here in detail because the surrogate modeling methods developed in Section III will make use of the formulation. The derivation follows the work of Homma and Saltelli.[3]

The Sobol' method is based on the ANOVA High-Dimensional Model Representation (ANOVA-HDMR). A high-dimensional model representation of a function, $f(\mathbf{x})$, can be written as

$$f(\mathbf{x}) = f_0 + \sum_i f_i(x_i) + \sum_{i<j} f_{ij}(x_i, x_j) + \ldots + f_{12\ldots n}(x_1, x_2, \ldots, x_n), \tag{2}$$

where f_0 is a constant, $f_i(x_i)$ is a function of only x_i, $f_{ij}(x_i, x_j)$ is a function of only x_i and x_j, etc. Without any constraints, the representation of $f(\mathbf{x})$ given by (2) is not unique, however, it can be made unique by enforcing the constraints

$$\int_0^1 f_{i_1,\ldots,i_s}(x_{i_1}, \ldots, x_{i_s})\, dx_\omega = 0, \text{ for } \omega = i_1, \ldots, i_s, \quad s = 1, \ldots, n, \tag{3}$$

where the function $f(\mathbf{x})$, and hence all its components, has been assumed to be integrable. For simplicity of presentation, the inputs to the function in (3) have been defined on the interval [0,1], but this assumption is not essential to the method. For each s, the indices $i_1, \ldots, i_s$ in (3) are all sets of s integers such that

American Institute of Aeronautics and Astronautics

$1 \leq i_1 < \ldots < i_s \leq n$. Thus, for $s = 1$, the constraint (3) applies to all terms f_i in (2), while for $s = 2$, the constraint (3) applies to all terms f_{ij} with $i < j$ as in (2), etc. The application of the constraint (3) makes the HDMR a unique representation of the function $f(\mathbf{x})$, referred to as an ANOVA-HDMR. Integration of $f(\mathbf{x})$ over all inputs results in $\int f(\mathbf{x})d\mathbf{x} = f_0$, which assuming each input x_i is a uniform random variable on [0,1], is the mean value of the function $f(\mathbf{x})$.

The constraint given by (3) also forces the different components of $f(\mathbf{x})$ within the ANOVA-HDMR to be orthogonal. That is, if $(i_1, \ldots, i_s) \neq (j_1, \ldots, j_l)$, then

$$\int f_{i_1,\ldots,i_s}(x_{i_1}, \ldots, x_{i_s}) \times f_{j_1,\ldots,j_l}(x_{j_1}, \ldots, x_{j_l}) \; d\mathbf{x} = 0, \tag{4}$$

since at least one index is not repeated.

Assuming now that $f(\mathbf{x})$ is square integrable, and therefore all components within the ANOVA-HDMR are as well, the variance of $f(\mathbf{x})$ is written as

$$D = \int f(\mathbf{x})^2 \; dx - f_0^2, \tag{5}$$

and partial variances are defined as

$$D_{i_1 \ldots i_s} = \int f_{i_1,\ldots,i_s}(x_{i_1}, \ldots, x_{i_s})^2 \; dx_{i_1} \; \ldots \; dx_{i_s}. \tag{6}$$

Given the ANOVA-HDMR for some $f(\mathbf{x})$, we square and then integrate both sides of (2), and employ the orthogonality constraint to arrive at

$$\int f(\mathbf{x})^2 \; dx = f_0^2 + \sum_i D_i + \sum_{i<j} D_{ij} + \ldots + D_{12\ldots n}, \tag{7}$$

which implies

$$D = \sum_i D_i + \sum_{i<j} D_{ij} + \ldots + D_{12\ldots n}. \tag{8}$$

This is precisely the notion shown in Figure 1.

Global sensitivity indices are defined as

$$S_{i_1,\ldots,i_s} = \frac{D_{i_1 \ldots i_s}}{D}, \quad s = 1, \ldots, n. \tag{9}$$

The sum of all global sensitivities of this form for a given function is unity. Global sensitivity indices with only one subscript, (e.g. S_i), are called main effect sensitivities, and those with multiple subscripts, (e.g. $S_{i,j}$,

$S_{i,j,k}$, etc.), are called interaction effect sensitivities. The sum of a factor's main effect global sensitivity and all interaction effect sensitivities that involve that factor gives the total effect sensitivity index, τ, which is defined for input factor i as

$$\tau_i = S_i + S_{i,i^c}, \tag{10}$$

where S_i is the main effect sensitivity to factor i, and S_{i,i^c} is the sum of the sensitivity indices of all interaction effects that include factor i. Since the sum of all unique sensitivity indices is unity, we have that

$$\tau_i = S_i + S_{i,i^c} = 1 - S_{i^c}, \tag{11}$$

where S_{i^c} is the sum of the sensitivity indices for all main effects and interactions effects that do not involve factor i. Since interaction effects will be counted for each factor involved in them, $\sum_i \tau_i \geq 1$.

The total effect sensitivity indices in (11) can be computed via Monte Carlo simulation as follows,[3] where hat quantities denote estimates of the corresponding true quantities. Here it should be noted that the computation of the partial variances with Monte Carlo simulation proceeds directly with the function $f(\mathbf{x})$ and does not require explicit knowledge of the functions on the right-hand side of (2). The estimate of the mean f_0 is computed as

$$\hat{f}_0 = \frac{1}{N} \sum_{m=1}^{N} f(\mathbf{x}^m), \tag{12}$$

while the estimate of the variance D is

$$\hat{D} = \frac{1}{N} \sum_{m=1}^{N} f(\mathbf{x}^m)^2 - \hat{f}_0^2. \tag{13}$$

The single-factor partial variance is then computed for factor i by resampling all factors except factor i:

$$\hat{D}_i = \frac{1}{N} \sum_{m=1}^{N} f([x_1^m, \ldots, x_i^m, \ldots, x_n^m]^T) f([\tilde{x}_1^m, \ldots, x_i^m, \ldots, \tilde{x}_n^m]^T) - \hat{f}_0^2, \quad i = 1, \ldots, n, \tag{14}$$

where $\tilde{x}_j^m$ denotes a different sample of factor x_j. The estimate of the variance due to all factors except factor i (which includes the sum of all single-factor and interaction effect partial variances that do not include factor i) is denoted as $\hat{D}_{i^c}$, and is computed by

$$\hat{D}_{i^c} = \frac{1}{N} \sum_{m=1}^{N} f([x_1^m, \ldots, x_i^m, \ldots, x_n^m]^T) f([x_1^m, \ldots, \tilde{x}_i^m, \ldots, x_n^m]^T) - \hat{f}_0^2, \tag{15}$$

where now just factor i is resampled. Finally, computing $\hat{S}_{i^c} = \hat{D}_{i^c}/\hat{D}$ and applying (11), we obtain the desired total effect sensitivity index.

The main effect sensitivity indices, S_i, may be used for factor prioritization by ranking inputs according to their main effect indices, which give the percentage of how much output variability can be expected to be eliminated by fixing a particular input somewhere on its domain. The total effect sensitivity indices, τ_i, may be used for factor fixing, since a low total effect index reveals a given input has a small main effect and also does not take part in substantial interactions among other inputs. For n inputs, the calculation of sensitivity indices requires $(2n + 1)$ Monte Carlo simulations (each with N model evaluations) if both the main effect and total effect indices are desired. Thus, like uncertainty analysis for decision-making, sensitivity analysis of large-scale models for development purposes will in most situations require surrogate models.

II.B.   Surrogate Modeling

Surrogate models can be categorized into three different classes: data-fit models, reduced-order models, and hierarchical models.[11] Data-fit models are generated using interpolation or regression of simulation data from the input/output relationships in the high-fidelity model.[11, 12] The primary challenge in adopting this surrogate modeling strategy for large-scale complex system models is the "curse of dimensionality" when the number of inputs to a model is large and design of experiment techniques must be applied with care in order to balance the computational cost of the required simulations with coverage of the input space. Reduced-order models are typically constructed for systems described by partial differential equations or large sets of ordinary differential equations.[13] Derivation of reduced-order models relies on the knowledge of the governing equations and are thus not suitable to systems for which the governing equations are unknown or empirically based. Hierarchical surrogate models, also known as variable fidelity models, employ simplified mathematical models such as coarser grids in finite element models[14] and models with simplified physics.[15–17]

The application of a particular surrogate modeling strategy depends both on what computational tasks are to be performed, and on the underlying structure of the model. In some cases, nothing will be known about a given model, and strategies that perform better in black-box situations, such as data-fit methods should be used. In other cases, everything will be known about the governing equations of a given model, and reduced-order models can be derived using projection-based approaches. In this paper, we consider models of a general form where the outputs of interest are computed as a sum of many constituent parts. As described in the next subsection, this class of models has applications in engineering and logistics systems. For such models, existing surrogate modeling methods cannot be applied. While the structure of the model is known, the input space is too large to use a projection-based model reduction approach or a data-fit method.

American Institute of Aeronautics and Astronautics

II.C.   Generalized Model Form

Consider a probability triple $(\mathbb{R}, \mathbb{B}, \mathbb{P})$, where $\mathbb{R}$ is the sample space, here the set of real numbers, $\mathbb{B}$ is the σ-field, here the Borel σ-field, and $\mathbb{P}$ is a probability measure. Also, consider a class of $\mathbb{B}$-measurable functions, f. The class of models we consider in this work are of the form

$$Y = \sum_{k=1}^{N} z_k = \sum_{k=1}^{N} f^k(\mathbf{x}), \tag{16}$$

where Y is an output of interest computed via a sum over N constituent parts, which are represented by the z_k that may themselves be functions of random inputs, $z_k = f^k(\mathbf{x})$, where $\mathbf{x} = [x_1, x_2, ..., x_n]^T$ is a vector of n independent random inputs, and $f^k \in f$. Because the f^k are all $\mathbb{B}$-measurable, the z_k and Y are random variables.

Though the form of (16) is simple, it is relevant to a broad range of applications. For example, outputs of the form (16) are common in systems whose performance is computed over many parts—e.g., emissions or fuel burn summed over mission segments for a transportation system, system failure time estimated as a sum of individual component lifetimes, many aspects of wireless communication, including current fluctuations in tunnel junctions, diversity schemes, and cochannel interference, profit summed over products or store locations, or votes summed over polling locations. In these examples, and in general, it is important to note that the assumption of independence is an aspect of the modeling. Our purpose here is to create surrogates for models of the form given by (16) so that we may carry out both uncertainty and sensitivity analyses in a reasonable amount of time while maintaining quantitative rigor. If independence is assumed when the full model is exercised, it should be assumed when the surrogate model is used in place of the full model, since the surrogate is attempting to produce estimates of results that would be obtained with the full model. The same sentiment holds when considering the assumption of randomness.

Simpson et al. present a survey of surrogate modeling techniques, along with recommendations for selecting a modeling approach.[18] They find that response surface modeling is appropriate for applications with less than ten input factors, while interpolation-based methods such as Kriging may be suitable for up to 50 input factors. For our problems of interest, the number of inputs is typically in the thousands or even millions. Of existing surrogate modeling methods, only neural networks are even remotely possible for such a large number of inputs. However, as discussed in Ref. 18, neural networks are computationally expensive to create and are best suited for deterministic problems. More recent advances in surrogate modeling methods, such as the pseudo response surface methodology,[19] address some of the computational challenges associated with high-dimensional input spaces by requiring the surrogate to be accurate only in some regions of the design space (e.g., near the Pareto front). However, even with these advances, surrogate modeling for systems with

thousands or millions of inputs remains out of reach. In the next section we present a hierarchical surrogate modeling approach that addresses this challenge for models of the form given by (16).

III. Surrogate Modeling Methodology with Quantified Confidence Intervals

Here we propose a hierarchical surrogate modeling approach targeted at uncertainty analysis and sensitivity analysis applications for the class of models described in the previous section. However, if the uncertainty associated with the use of a surrogate model in place of a full model is not properly quantified, the usefulness of the various uncertainty analyses will be limited. Thus, one of the key objectives of this work is to quantify the effects of using a surrogate model to perform uncertainty and sensitivity analyses. The methods used to achieve this goal for the general class of models given by (16) are discussed in the following subsections.

III.A.   Hierarchical Surrogate Modeling Approach

For the class of models given by (16), a natural representative for building a hierarchical surrogate is a single random variable, z_k. Our surrogate modeling approach is thus to approximate the output of interest, Y, using a subset of the z_k. For the case of a general model of the form of (16), if the z_k are such that

$$\max_{c \leq N} \frac{\mathrm{var}(z_c)}{\Sigma_{k=1}^{N} \mathrm{var}(z_k)} \to 0 \quad \text{as} \quad N \to \infty, \tag{17}$$

where $c \in \{1, 2, \ldots, N\}$, and N is the number of constituent parts used in the summation given in (16), then according to the Central Limit theorem,

$$Y \to^d \mathcal{N} \left(\sum_{k=1}^{N} \mathbb{E}[z_k], \sum_{k=1}^{N} \mathrm{var}(z_k) \right) \quad \text{as} \quad N \to \infty, \tag{18}$$

where the convergence is in distribution, $\mathcal{N}(\alpha, \beta)$ is a normal distribution with mean α and variance β, and the constraint given by (17) is referred to as the Lyapunov condition.[20] Though (18) states that the convergence to a normal distribution occurs as $N \to \infty$, it is common in statistical practice to assume Y may be appropriately modeled with a normal distribution when $N \geq 30$.[21] The constraint given by (17) can be met for example, by any set of z_k such that $\mathrm{var}(z_k) < \infty$ for $k = 1, 2, \ldots, N$ and $\sum_{k=1}^{N} \mathrm{var}(z_k)$ is unbounded as $N \to \infty$. Although here independence was used to invoke the Central Limit theorem, there are other methods for invoking the Central Limit theorem, such as m-dependence,[22] for situations where the independence condition is not met. The application of the Central Limit theorem to the general model given in (16) for both uncertainty and sensitivity analyses is discussed in the following section.

III.B. Surrogate Modeling for Decision-Making Uncertainty Analysis

As given by (18), the output of interest, Y, is normally distributed. To estimate the distribution of Y with a surrogate model, only estimates of $\sum_{k=1}^{N} \mathbb{E}[z_k]$ and $\sum_{k=1}^{N} \mathrm{var}(z_k)$ are required. We may estimate these quantities by noting that if we were to compute $\mathbb{E}[z_k]$ for every constituent part z_k, we could view the resulting set of expected values as representing a set of N samples drawn from some distribution. Thus, the expected value of some z_k can be considered as a sample from a random variable, and can be estimated using a subset of random variables chosen from the full set. We denote by $\mathcal{Q}$ the subset of N_q random variables chosen randomly from the full set of N random variables. Then using the law of large numbers, $\sum_{k=1}^{N} \mathbb{E}[z_k]$ is estimated as $N \frac{1}{N_q} \sum_{k \in \mathcal{Q}} \mathbb{E}[z_k]$, since

$$N \frac{1}{N_q} \sum_{k \in \mathcal{Q}} \mathbb{E}[z_k] \to N \mathbb{E}[\mathbb{E}[z_k]] = N \frac{1}{N} \sum_{k=1}^{N} \mathbb{E}[z_k] = \sum_{k=1}^{N} \mathbb{E}[z_k] \text{ as } N_q \to N. \tag{19}$$

For the sum of the variances of the z_k in (18), a similar method is followed to derive an analogous expression for the variance estimate of Y. Thus, the surrogate model estimate of the output distribution using the subset $\mathcal{Q}$ of N_q random variables to represent the full N random variables is given by

$$\hat{Y} \sim \mathcal{N} \left(\frac{N}{N_q} \sum_{k \in \mathcal{Q}} \mathbb{E}[z_k], \frac{N}{N_q} \sum_{k \in \mathcal{Q}} \mathrm{var}(z_k) \right), \tag{20}$$

where $\hat{Y}$ is a random variable that is an estimate of the random variable Y.

In (20), the terms $\frac{1}{N_q} \sum_{k \in \mathcal{Q}} \mathbb{E}[z_k]$ and $\frac{1}{N_q} \sum_{k \in \mathcal{Q}} \mathrm{var}(z_k)$, are sample means of the distributions of expected values of the z_k and of the variances of the z_k, respectively. According to the Central Limit theorem, these sample means have the following normal distributions:

$$\frac{1}{N_q} \sum_{k \in \mathcal{Q}} \mathbb{E}[z_k] \sim \mathcal{N}(\mu_{\mathbb{E}[z_k]}, \frac{N - N_q}{N - 1} \sigma_{\mathbb{E}[z_k]}^2 / N_q),$$

$$\frac{1}{N_q} \sum_{k \in \mathcal{Q}} \mathrm{var}(z_k) \sim \mathcal{N}(\mu_{\mathrm{var}(z_k)}, \frac{N - N_q}{N - 1} \sigma_{\mathrm{var}(z_k)}^2 / N_q), \tag{21}$$

where $\mu_{\mathbb{E}[z_k]}$ is the expected value of the distribution of expected values of the z_k, $\sigma_{\mathbb{E}[z_k]}^2$ is the variance of the distribution of expected values of z_k, $\mu_{\mathrm{var}(z_k)}$ is the expected value of the distribution of variances of the z_k, and $\sigma_{\mathrm{var}(z_k)}^2$ is the variance of the distribution of variances of the z_k. The $\frac{N - N_q}{N - 1}$ terms are finite population correction factors that must be applied since N is finite and the sampling of N_q random variables from N total random variables is done without replacement.[23]

As noted in Section II, a key outcome of an uncertainty analysis intended to support decision-making is the ability to compare such quantities as output means and variances. These quantities cannot be computed

exactly using a surrogate model; however, confidence intervals for these quantities can be rigorously computed since, as shown in the analysis above, the parameters are normally distributed. The confidence intervals for the mean and variance of the output Y can be constructed from

$$\frac{N}{N_q} \sum_{k \in \mathcal{Q}} \mathbb{E}[z_k] - Z_{\alpha/2} \sqrt{\frac{N - N_q}{N - 1} \frac{N2}{N_q} \sigma^2_{\mathbb{E}[z_k]}} < \mathbb{E}[Y] < \frac{N}{N_q} \sum_{k \in \mathcal{Q}} \mathbb{E}[z_k] + Z_{\alpha/2} \sqrt{\frac{N - N_q}{N - 1} \frac{N2}{N_q} \sigma^2_{\mathbb{E}[z_k]}} \tag{22}$$

and

$$\frac{N}{N_q} \sum_{k \in \mathcal{Q}} \mathrm{var}(z_k) - Z_{\alpha/2} \sqrt{\frac{N - N_q}{N - 1} \frac{N2}{N_q} \sigma^2_{\mathrm{var}(z_k)}} < \mathrm{var}(Y) < \frac{N}{N_q} \sum_{k \in \mathcal{Q}} \mathrm{var}(z_k) + Z_{\alpha/2} \sqrt{\frac{N - N_q}{N - 1} \frac{N2}{N_q} \sigma^2_{\mathrm{var}(z_k)}}, \tag{23}$$

where $Z_{\alpha/2}$ is the value of the inverse cumulative distribution function of a standard normal random variable evaluated at $(1 - \alpha/2)$, where α sets the level of confidence.[24] A typical value of $Z_{\alpha/2}$ is 1.96, which corresponds to a 95% confidence interval. In practice, constructing these confidence intervals requires estimating the variance of the distribution of the expected values of the z_k, $\sigma^2_{\mathbb{E}[z_k]}$, and the variance of the distribution of the variances of the z_k, $\sigma^2_{\mathrm{var}(z_k)}$. We estimate these parameters using the sample variance for each, which are calculated from

$$\hat{\sigma}^2_{\mathbb{E}[z_k]} \quad = \quad \frac{1}{N_q - 1} \sum_{k=1}^{N_q} (\mathbb{E}[z_k] - \overline{\mathbb{E}[z_k]})^2 \tag{24}$$

$$\hat{\sigma}^2_{\mathrm{var}(z_k)} \quad = \quad \frac{1}{N_q - 1} \sum_{k=1}^{N_q} (\mathrm{var}(z_k) - \overline{\mathrm{var}(z_k)})^2, \tag{25}$$

where $\hat{\sigma}^2_{\mathbb{E}[z_k]}$ and $\hat{\sigma}^2_{\mathrm{var}(z_k)}$ are the sample variances of $\sigma^2_{\mathbb{E}[z_k]}$ and $\sigma^2_{\mathrm{var}(z_k)}$ respectively, and $\overline{\mathbb{E}[z_k]}$ and $\overline{\mathrm{var}(z_k)^2}$ are the sample means of the distributions of the expected values and the variances of the z_k respectively. The estimates, $\hat{\sigma}^2_{\mathbb{E}[z_k]}$ and $\hat{\sigma}^2_{\mathrm{var}(z_k)}$, are then used in (22) and (23). These intervals also require the estimation of expected values of the random variables z_k, $\mathbb{E}[z_k]$, and variances, $\mathrm{var}(z_k)$. These parameters, as will be discussed in Section IV, are estimated from a Monte Carlo simulation with a large number of model evaluations and thus, uncertainty associated with these estimates is neglected.

As can be seen from (22), as N_q approaches N, the confidence interval around $\mathbb{E}[Y]$ narrows, eventually becoming a single point when $N_q = N$. Thus, there is a tradeoff between how many random variables are considered in the surrogate model and the tightness of the confidence intervals for the mean and variance of the output Y.

III.C. Surrogate Modeling for Model Development Sensitivity Analysis

The ANOVA-HDMR for a random variable z_k given by $f^k(\mathbf{x})$ may be written as

$$z_k = f_0^k + \sum_i f_i^k(x_i) + \sum_{i<j} f_{ij}^k(x_i, x_j) + \ldots + f_{12\ldots n}^k(x_1, x_2, \ldots, x_n), \tag{26}$$

which can be squared and integrated, as was done to arrive at (7), giving

$$\text{var}(z_k) := \sum_i D_i^k + \sum_{i<j} D_{ij}^k + \ldots + D_{12\ldots n}^k, \tag{27}$$

where D_i^k is the partial variance of z_k due to input x_i and the rest of the terms represent partial variances of z_k due to various levels of interactions between the components of $\mathbf{x}$. Summing over all the z_k gives

$$Y = \sum_{k=1}^N f_0^k + \sum_{k=1}^N \sum_i f_i^k(x_i) + \sum_{k=1}^N \sum_{i<j} f_{ij}^k(x_i, x_j) + \ldots + \sum_{k=1}^N f_{12\ldots n}^k(x_1, x_2, \ldots, x_n). \tag{28}$$

At this point we may proceed as we did in Section II.A and estimate total and main effect sensitivity indices for each x_i. However, the models of the form given by (16) may contain millions of inputs, rendering computation of sensitivity indices for each input impractical and of limited use. For such models, it is typical that these inputs comprise just a few distinct physical quantities. For example, for the emissions model analyzed in the next section, the inputs consist of six physical quantities (fuel burn, temperature, pressure, relative humidity, fuel flow, and reference emission index) which are defined for each flight segment of each operation in a set of aircraft flights. A sensitivity analysis could consider the sensitivities of each of these independently sampled inputs individually, resulting in millions of sensitivity indices. From a practical standpoint, we are more interested in determining the sensitivity of model outputs to groups of inputs. For example, for the nitrous oxide (NO_x) emissions for an aircraft operation, we might wish to compute the contribution to variance of all fuelburn inputs for that operation, where each input is sampled independently across flight segments in the operation. Alternatively, for the total NO_x emissions summed over a set of operations, we might wish to compute the contribution to variance of all fuelburn inputs for those operations, where again each input is sampled independently across all flight segments. In this section, we present the extension of the global sensitivity analysis methodology to handle such cases for the class of general models given by (16). Further, application of the Central Limit Theorem, which is permissible given the additive nature of the models and ANOVA-HDMR, enables the calculation of confidence intervals around sensitivity indices computed using the surrogate models in place of the full models.

Consider a case where we have two physical quantities (type 1 and type 2, e.g. pressure and temperature) within $\mathbf{x}$ that we wish to group. We define the set of physical quantities of type 1 as $\mathcal{G}_1 = \{x_i | x_i \text{ are of type 1}\}$

and the set of physical quantities of type 2 as $\mathcal{G}_2 = \{x_i | x_i \text{ are of type 2}\}$. An ANOVA-HDMR can then be written for Y in terms of the groups $\mathcal{G}_1$ and $\mathcal{G}_2$ and their interaction rather than the individual components of $\mathbf{x}$. The variance of Y can then be written as

$$\text{var}(Y) := D = D_{\mathcal{G}_1} + D_{\mathcal{G}_2} + D_{\mathcal{G}_1 \mathcal{G}_2}, \tag{29}$$

where $D_{\mathcal{G}_1}$ is the variance of Y due to the elements of $\mathcal{G}_1$ and their interactions, $D_{\mathcal{G}_2}$ is the variance of Y due to the elements of $\mathcal{G}_2$ and their interactions, and $D_{\mathcal{G}_1 \mathcal{G}_2}$ is the variance of Y due to the interactions involving elements from both $\mathcal{G}_1$ and $\mathcal{G}_2$. A global sensitivity analysis may then be carried out to compute total effect and main effect sensitivity indices for the groups. For $\mathcal{G}_1$, the total effect sensitivity index can be computed as

$$\tau_{\mathcal{G}_1} = 1 - \frac{D_{\mathcal{G}_1^c}}{D}, \tag{30}$$

where $D_{\mathcal{G}_1^c}$ is the sum of the variances due to all main effect terms that do not involve elements of $\mathcal{G}_1$, which here is just $D_{\mathcal{G}_2}$. The main effect sensitivity index can be computed as

$$S_{\mathcal{G}_1} = \frac{D_{\mathcal{G}_1}}{D}. \tag{31}$$

Similar expressions may be written for $\mathcal{G}_2$.

For the class of models with which we are concerned, which consist of a summation of N constituent parts as in (16), it is typical that particular physical inputs be defined on each part. Thus, the groups in this case will each have N elements. As will be shown in Section IV, computation of the sensitivity indices given by (30) and (31) can proceed by performing global sensitivity analyses on the constituent parts of the model. This enables the creation of hierarchical surrogate models for global sensitivity analysis in the form of subsets of constituent parts much like the surrogate models constructed in Section III.B. Confidence intervals may be derived for the sensitivity indices computed using these surrogates, the formulation of which is given in the following section.

IV. AEM Application and Results

In this section, our approach is applied to the example of estimating emissions of aircraft flights. We first describe the model and then present surrogate models for decision-making uncertainty analysis and model development sensitivity analysis. This is followed by results from using the surrogate models in place of the full model for a typical emissions analysis.

IV.A. Aircraft Emissions Model

The AEM is used to calculate emissions inventories of such pollutants as CO_2, CO, NO_x, SO_x, and many others. The calculation is done on an operation-by-operation basis, and the emissions computed for each operation in a given scenario are then aggregated to produce an emissions inventory. An operation is in turn simulated on a flight segment-by-segment basis as shown in Figure 2, where emissions are calculated for each segment of the operation and then aggregated to produce the total emissions of the operation.
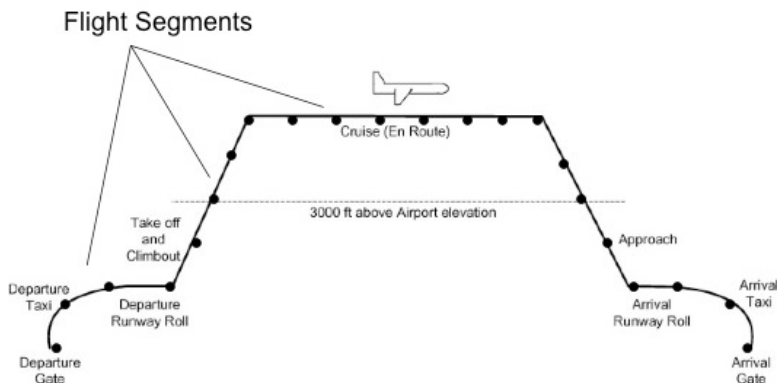
Figure 2. AEM segmentation of an aircraft operation. Adapted from Ref. 25.

The AEM inputs considered in our analyses are shown in Table 1, where each input is defined for each segment of each operation. We consider the emissions resulting from a total of N_o operations, each consisting of N_s flight segments. Thus, the total number of inputs is given by $n = 6N_oN_s$. Table 1 also shows the probability density functions that are defined for each input on a segment-by-segment basis. These density functions were arrived at through previous studies and expert opinions of the Partnership for AiR Transportation Noise and Emissions Reduction (PARTNER).[26] The samples from the density functions are applied as multipliers to default values of the various inputs that are specific to aircraft type, engine type, and geographic location. For triangular distributions, the defining values are the minimum, mode, and maximum values. For uniform distributions, the defining values are the minimum and the maximum.

Table 1. AEM inputs and their probability density functions. All inputs are applied as multipliers to nominal input values.

Input Variable	Input Quantity	Distribution Type	Defining Values
$x_1 := q$	Fuel Burn	Uniform	$[0.95, 1.05]$
$x_2 := r$	Temperature	Triangular	$[0.89, 1.00, 1.11]$
$x_3 := s$	Pressure	Triangular	$[0.97, 1.00, 1.03]$
$x_4 := t$	Relative Humidity	Triangular	$[0.82, 1.00, 1.17]$
$x_5 := u$	Fuel Flow	Uniform	$[0.95, 1.05]$
$x_6 := v$	$REINO_x$	Triangular	$[0.76, 1.00, 1.24]$

The outputs of the AEM (global emissions of NO_x, CO, CO_2, etc.) are all computed in a similar manner,

American Institute of Aeronautics and Astronautics

thus the modeling methodology is developed here only for the NO_x output. Other outputs are treated in an analogous way. The NO_x produced for operation k, y_k, is calculated as

$$y_k = \sum_{l=1}^{N_s} q_{kl} g_{kl}(r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl}),$$ (32)

where q_{kl} is the fuelburn on segment l of operation k, and $g_{kl}(r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl})$ is the emissions index of NO_x (EINO$_x$) on segment l of operation k, which is calculated using Boeing Method 2 (Ref. 27), with the inputs defined in Table 1, specifically the temperature, pressure, relative humidity, fuel flow, and reference emissions index of NO_x (REINO$_x$) for the given segment. The Boeing Method 2 is an empirical method for correcting reference emission indices for flight conditions using fuel flow and atmospheric conditions. The total NO_x output for a set of N_o operations is then calculated as

$$y_{tot} = \sum_{k=1}^{N_o} y_k = \sum_{k=1}^{N_o} \sum_{l=1}^{N_s} q_{kl} g_{kl}(r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl}).$$ (33)

Since the functions within the AEM are all continuous, and each input of the AEM is a random variable, each output of the AEM is also a random variable. Thus, the total NO_x output, y_{tot}, can be thought of as a random sample from the random variable Y_{tot}, for which confidence intervals and sensitivity indices are desired.

Equation (33) reveals the structure of the AEM once it has been decomposed by operations. Given that computations are performed separately on each operation within the AEM, it has the general form of (16), where the $z_k \equiv y_k$, $f^k \equiv \sum_{l=1}^{N_s} q_{kl} g_{kl}(r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl})$, and $\mathbf{x} \equiv (q_{kl}, r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl})$. Our surrogate modeling approach thus amounts to approximating the output of interest, total NO_x emissions, using a subset of flight operations. This is illustrated in Figure 3, where it can been seen that the surrogate modeling approach is based on reducing the dimension of the input space.

Since the functions within the AEM are all continuous, and each input of the AEM is a random variable, each single operation output of the AEM is also a random variable. These random variables are independent and satisfy the Lyapunov condition;[20] therefore, the AEM is a member of the general class of models described in Section II.C and the results presented in Section III are applicable. Here, it is noted that, owing to such circumstances as aircraft operating on similar routes, certain operations in the real-world would encounter similar environmental factors, such as temperature, pressure, and humidity, thus causing some dependence in how those factors enter into emissions estimates, which would lead to some dependence in the outputs of these operations. However, the AEM does not currently include these factor dependencies, which implies that the operation-level outputs of the AEM are completely independent of each other. Thus, the Central Limit theorem may be invoked.

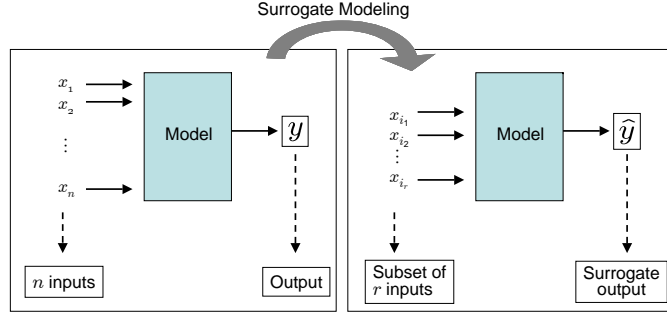American Institute of Aeronautics and Astronautics

Figure 3. The hierarchical surrogate modeling approach achieves a reduction in computational complexity through a reduction of the input space. For the AEM, this amounts to selecting a subset of r operations, denoted by the subscripts, $i_1, ..., i_r$, over which to estimate the total emissions.

IV.B. AEM Surrogate for Decision-Making Uncertainty Analysis

According to the Central Limit theorem, the output distribution of total NO_x emissions, y_{tot}, is normally distributed since

$$y_{tot} = \sum_{k=1}^{N_o} y_k \to^d \mathcal{N}(\sum_{k=1}^{N_o} \mathbb{E}[y_k], \sum_{k=1}^{N_o} \sigma_{y_k}^2) \text{ as } N_o \to \infty, \tag{34}$$

where the convergence is in distribution, and $\mathcal{N}(\alpha, \beta)$ is a normal distribution with mean α and variance β. Typical analyses involving the AEM calculate emissions inventories for representative days of operations, for which $N_o \approx 70,000$, and one year of operations, for which $N_o \approx 30,000,000$. As noted in Section II.C, to assume normality N_o should be greater than about 30, thus the number of samples is much greater than required.

To estimate the distribution of y_{tot} with a surrogate model, only estimates of $\sum_{k=1}^{N_o} \mathbb{E}[y_k]$ and $\sum_{k=1}^{N_o} \sigma_{y_k}^2$ are required. As shown in Section III, we may generate a surrogate model estimate of the the total NO_x output distribution using a subset $\mathcal{O}$ of n_o operations to represent the full N_o operations as

$$\hat{y}_{tot} \sim \mathcal{N} \left(\frac{N_o}{n_o} \sum_{k \in \mathcal{O}} \mathbb{E}[y_k], \frac{N_o}{n_o} \sum_{k \in \mathcal{O}} \sigma_{y_k}^2 \right), \tag{35}$$

In (35), the terms $\frac{1}{n_o} \sum_{k \in \mathcal{O}} \mathbb{E}[y_k]$ and $\frac{1}{n_o} \sum_{k \in \mathcal{O}} \sigma_{y_k}^2$, are sample means of the distributions of expected values of y_k and of the variances of the y_k, respectively. These sample means then have the following normal

distributions:

$$\frac{1}{n_o} \sum_{k \in \mathcal{O}} \mathbb{E}[y_k] \sim \mathcal{N}(\mu_{\mathbb{E}}, \frac{N_o - n_o}{N_o - 1} \sigma_{\mathbb{E}}^2 / n_o),$$

$$\frac{1}{n_o} \sum_{k \in \mathcal{O}} \sigma_{y_k}^2 \sim \mathcal{N}(\mu_{\sigma^2}, \frac{N_o - n_o}{N_o - 1} \sigma_{\sigma^2}^2 / n_o), \tag{36}$$

where $\mu_{\mathbb{E}}$ is the expected value of the distribution of expected values of the y_k, $\sigma_{\mathbb{E}}^2$ is the variance of the distribution of expected values of y_k, μ_{σ^2} is the expected value of the distribution of variances of the y_k, and $\sigma_{\sigma^2}^2$ is the variance of the distribution of variances of the y_k.

As given in Section III, confidence intervals for the mean and variance of total NO_x emissions can be constructed from

$$\frac{N_o}{n_o} \sum_{k \in \mathcal{O}} \mathbb{E}[y_k] - Z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\mathbb{E}}^2} < \mathbb{E}[y_{tot}] < \frac{N_o}{n_o} \sum_{k \in \mathcal{O}} \mathbb{E}[y_k] + Z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\mathbb{E}}^2} \tag{37}$$

and

$$\frac{N_o}{n_o} \sum_{k \in \mathcal{O}} \sigma_{y_k}^2 - Z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\sigma^2}^2} < \mathrm{var}(y_{tot}) < \frac{N_o}{n_o} \sum_{k \in \mathcal{O}} \sigma_{y_k}^2 + Z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\sigma^2}^2}. \tag{38}$$

As noted in Section III, construction of these confidence intervals requires estimating the variance of the distribution of the expected values of the y_k, $\sigma_{\mathbb{E}}^2$, and the variance of the distribution of the variances of the y_k, $\sigma_{\sigma^2}^2$. We estimate these parameters using the sample variance for each, which are calculated from

$$\hat{\sigma}_{\mathbb{E}}^2 = \frac{1}{n_o - 1} \sum_{k=1}^{n_o} (\mathbb{E}[y_k] - \overline{\mathbb{E}[y_k]})^2 \tag{39}$$

$$\hat{\sigma}_{\sigma^2}^2 = \frac{1}{n_o - 1} \sum_{k=1}^{n_o} (\sigma_{y_k}^2 - \overline{\sigma_{y_k}^2})^2, \tag{40}$$

where $\hat{\sigma}_{\mathbb{E}}^2$ and $\hat{\sigma}_{\sigma^2}^2$ are the sample variances of $\sigma_{\mathbb{E}}^2$ and $\sigma_{\sigma^2}^2$ respectively, and $\overline{\mathbb{E}[y_k]}$ and $\overline{\sigma_{y_k}^2}$ are the sample means of the distributions of the expected values and the variances of the y_k respectively. The estimates, $\hat{\sigma}_{\mathbb{E}}^2$ and $\hat{\sigma}_{\sigma^2}^2$, are then used in (37) and (38). As will be shown in Section IV.D, n_o is sufficiently large to neglect the uncertainty associated with these estimates. These intervals also require the estimation of operation-level expected values, $\mathbb{E}[y_k]$, and variances, $\sigma_{y_k}^2$, of NO_x emissions. These parameters, as will be discussed in Section IV.D, are estimated from a Monte Carlo simulation with a large number of model evaluations and thus, uncertainty associated with these estimates is also neglected.

As can be seen from (37), as n_o approaches N_o, the confidence interval around $\mathbb{E}[y_{tot}]$ narrows, eventually becoming a single point when $n_o = N_o$. Thus, as mentioned in Section III, there is a tradeoff between how

many operations are analyzed in the surrogate model, and the tightness of the confidence intervals for the mean and variance of the total NO_x. This tradeoff is discussed further in Sections IV.D and VI. Results from applying this method to construct confidence intervals for the mean and variance of the AEM NO_x output are presented in Section IV.D.

IV.C.   AEM Surrogate for Model Development Sensitivity Analysis

Consider the ANOVA-HDMR for the calculation of the NO_x emissions from a single operation, l:

$$
\begin{aligned}
y_k &= f_{0,k} + \sum_{l=1}^{N_s} f_{q_{kl}}(q_{kl}) + \sum_{l=1}^{N_s} f_{r_{kl}}(r_{kl}) + \sum_{l=1}^{N_s} f_{s_{kl}}(s_{kl}) + \sum_{l=1}^{N_s} f_{t_{kl}}(t_{kl}) \\
&\quad + \sum_{l=1}^{N_s} f_{u_{kl}}(u_{kl}) + \sum_{l=1}^{N_s} f_{v_{kl}}(v_{kl}) + \text{interaction terms},
\end{aligned}
\tag{41}
$$

where we use the input variable notation defined in Table 1. The term $\sum_{l=1}^{N_s} f_{q_{kl}}(q_{kl})$ is the sum of all the single-factor functions of factor q_{kl}; that is, the functions that depend only on the segment fuelburn inputs. The second summation is over those functions that depend only on the segment temperatures, r_{kl}, and so on for the other summations. Here, as in (32), N_s segments have been assumed for operation k.

Since the goal is to compute sensitivities for inputs grouped across flight segments, we define $q_k = \{q_{kl}\}_{l=1}^{N_s}$ to be the set of fuelburn segment inputs for operation k. Define r_k, s_k, t_k, u_k, and v_k similarly for the other input quantities. Each summation in (41) can then be written as

$$
f_{q_k} = \sum_{l=1}^{N_s} f_{q_{kl}}(q_{kl}),
\tag{42}
$$

with analogous expressions defining f_{r_k}, etc. Then (41) is written as

$$
y_k = f_{0,k} + f_{q_k} + f_{r_k} + f_{s_k} + f_{t_k} + f_{u_k} + f_{v_k} + \text{interaction terms}.
\tag{43}
$$

Squaring and integrating (43), as was done to arrive at (7), gives

$$
\text{var}(y_k) := D_{y_k} = D_{q_k} + D_{r_k} + D_{s_k} + D_{t_k} + D_{u_k} + D_{v_k} + \text{interaction partial variances},
\tag{44}
$$

where D_{q_k} is the partial variance due to all fuelburn inputs, and so on for the other inputs.

Similarly, the AEM output y_{tot}, which, as noted in Section II, is computed by aggregating the operational

level outputs, is written in ANOVA-HDMR form by summing over the operations in (41), which yields

$$
\begin{aligned}
y_{tot} &= \sum_{k=1}^{N_o} f_{0,k} + \sum_{k=1}^{N_o}\sum_{l=1}^{N_s} f_{q_{kl}}(q_{kl}) + \sum_{k=1}^{N_o}\sum_{l=1}^{N_s} f_{r_{kl}}(r_{kl}) + \sum_{k=1}^{N_o}\sum_{l=1}^{N_s} f_{s_{kl}}(s_{kl}) + \sum_{k=1}^{N_o}\sum_{l=1}^{N_s} f_{t_{kl}}(t_{kl}) \\
&+ \sum_{k=1}^{N_o}\sum_{l=1}^{N_s} f_{u_{kl}}(u_{kl}) + \sum_{k=1}^{N_o}\sum_{l=1}^{N_s} f_{v_{kl}}(v_{kl}) + \text{interaction terms.}
\end{aligned}
\tag{45}
$$

Now let $q = \{q_k\}_{k=1}^{N_o}$ denote the set of fuelburn inputs across all operations, and $f_q = \sum_{k=1}^{N_o}\sum_{l=1}^{N_s} f_{q_{kl}}(q_{kl})$ be the sum of all the single-factor functions of all segment fuelburn inputs, then (45) is written as

$$
y_{tot} = f_0 + f_q + f_r + f_s + f_t + f_u + f_v + \text{interaction terms,}
\tag{46}
$$

where $f_0 = \sum_{k=1}^{N_o} f_{0,k}$ is the expected value of y_{tot} and the functions f_r, f_s, f_t, f_u, and f_v are defined analogously to f_q. Squaring and integrating (46) gives

$$
\text{var}(y_{tot}) := D = D_q + D_r + D_s + D_t + D_u + D_v + \text{interaction partial variances,}
\tag{47}
$$

which may also be written as

$$
D = \sum_{k=1}^{N_o} D_{q_k} + \sum_{k=1}^{N_o} D_{r_k} + \sum_{k=1}^{N_o} D_{s_k} + \sum_{k=1}^{N_o} D_{t_k} + \sum_{k=1}^{N_o} D_{u_k} + \sum_{k=1}^{N_o} D_{v_k} + \text{interaction partial variances.}
\tag{48}
$$

The total effect sensitivity index for q, denoted τ_q, represents the relative contribution to the variance D of all fuelburn inputs over all operations and segments. As in (30), we write

$$
\tau_q = 1 - \frac{D_{q^c}}{D},
\tag{49}
$$

where D_{q^c} is the sum of the variances due to all main effect terms and interaction effect terms that do not involve fuelburn. By breaking this expression into a sum over operations and using the fact that $D = \sum_{k=1}^{N_o} D_{y_k}$, the expression (49) can be written as

$$
\tau_q = 1 - \frac{\sum_{k=1}^{N_o}(1 - \tau_{q_k})D_{y_k}}{\sum_{k=1}^{N_o} D_{y_k}},
\tag{50}
$$

where τ_{q_k} is the total effect sensitivity index for q_k, the fuelburn inputs over operation k. A similar derivation for main effect sensitivity indices leads to

$$
S_q = \frac{\sum_{k=1}^{N_o} S_{q_k}D_{y_k}}{\sum_{k=1}^{N_o} D_{y_k}}.
\tag{51}
$$

As was the case for the expected values of NO_x emissions on the operational level in Section III.A, the terms in (50) and (51) can be considered as samples from distributions. Therefore, to estimate the sensitivity indices given by (50) and (51), we apply the same process used to arrive at (20) from (18). In (51) for example, the distribution of the numerator, $\sum_{l=1}^{N_o} S_{q_l} D_{y_l}$, is equal to $N_o \mathbb{E}[S_{ql} D_{yl}]$, which may be estimated from $\frac{N_o}{n_o} \sum_{l \in \mathcal{O}} [S_{ql} D_{yl}]$. Just as in (36), this estimate is normally distributed and converges to a single value when $n_o = N_o$. To estimate confidence intervals for τ_q and S_q, we sample from distributions of the numerators and denominators to estimate the intervals empirically. It should be noted here that the confidence intervals computed for τ_q and S_q will be conservative due to the fact that the numerator and denominator terms in both (50) and (51) are positively correlated. By not including the correlation in the estimation of the confidence intervals, the estimate of the lower endpoint will be less than the true lower endpoint and the estimate of the upper endpoint will be greater than the true upper endpoint. This is due to the fact that the numerator in each equation must be less than or equal to the denominator in each equation, which leads to conservative intervals when the positive correlation term is not included. Results from applying this method to the AEM sensitivity indices for the total emissions of NO_x are presented in the following subsection.

IV.D. Results

A typical analysis run of the AEM consists of all operations conducted on a particular day that is considered a reasonable representative of all operations from a particular year. These days are referred to as *representative days*. The full AEM run for the representative day for the year 2005, which is the AEM model we consider here, has $N_o = 68,343$ operations. Each of these operations requires a Monte Carlo simulation to calculate operation-level emissions outputs that are then aggregated, as shown in (33) to produce the overall AEM output, y_{tot}. For the computational resources available for this study, a single model evaluation for one operation takes approximately 2.31×10^{-4} seconds. To perform both uncertainty and sensitivity analysis for a single operation requires 13 separate Monte Carlo simulations ($2n + 1$ simulations, where n, which is the dimension of the input space, is 6 for the AEM), each of which consisted of 10,000 model evaluations in this study. Thus, running each operation of the AEM representative day for 2005 to perform uncertainty and sensitivity analysis on y_{tot}, would take approximately 570 hours, which is computationally expensive, especially if many different policy scenarios are to be considered. As will be shown in the following subsections, the methods presented in Section III can be used to perform both uncertainty and sensitivity analyses on the AEM representative day with a surrogate model consisting of a randomly chosen subset of operations, while maintaining quantitative rigor in the analyses in a manner that is computationally efficient.

AEM Surrogate Results for Decision-making Uncertainty Analysis

To estimate the confidence intervals for the mean and variance of the total NO_x emissions from the representative day, 9,914 operations were chosen randomly, without replacement, from the full set of operations. As noted previously, a 10,000-iteration Monte Carlo simulation was run for each of the sampled operations, the results of which were used to compute operation-level NO_x means and sample variances. These values were then used to estimate the expected value of the distribution of operation-level expected values of NO_x emissions, $\mu_{\mathbb{E}}$; the variance of the expected value of the distribution of operation-level expected values of NO_x emissions, $\sigma_{\mathbb{E}}^2$; the expected value of the distribution of operation-level variances of NO_x emissions, μ_{σ^2}; and the variance of the distribution of operation-level variances of NO_x emissions, $\sigma_{\sigma^2}^2$. As was noted in Section III, these estimates are necessary for constructing the confidence intervals for the expected value and variance of the total NO_x emissions of the full AEM and uncertainty in these estimates has been neglected. Figure 4 shows the behavior of these estimates as the number of operations in the subset, n_o, is increased from 2,500 to 9,914 operations.
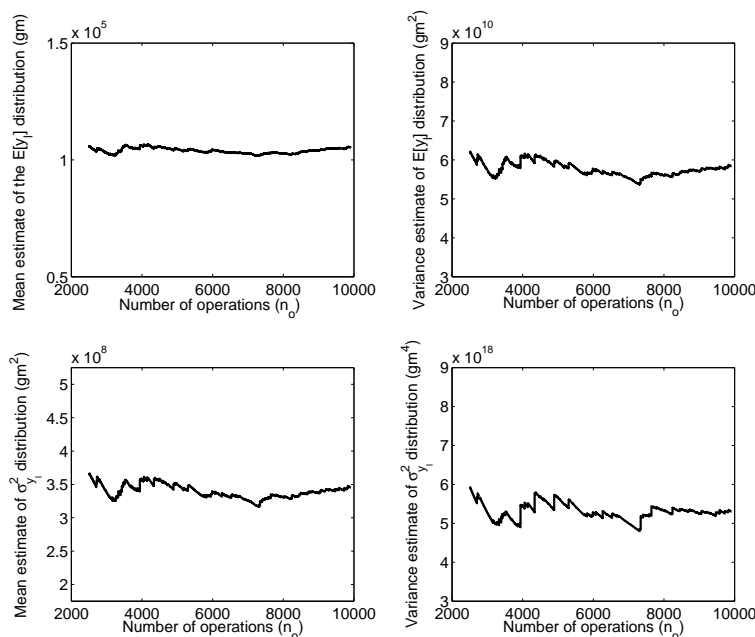

Figure 4. Estimates of $\mu_{\mathbb{E}}, \sigma_{\mathbb{E}}^2, \mu_{\sigma^2}$, and $\sigma_{\sigma^2}^2$ as the number of operations in the surrogate model, n_o, increases from 2,500 to 9,914.

The confidence intervals (95%) for the mean and variance of the total NO_x emissions computed at values of n_o of 2,500, 5,000, 7,500, and 9,914, are presented in Figure 5. Figure 6 presents the dependence of the confidence interval widths, in terms of percentage $\pm$ of the surrogate model estimated values, for the mean and variance of total NO_x emissions for a full run of the representative day as n_o increases from 2,500 to 9,914. These results show that by applying the surrogate modeling methodology described in Section III for
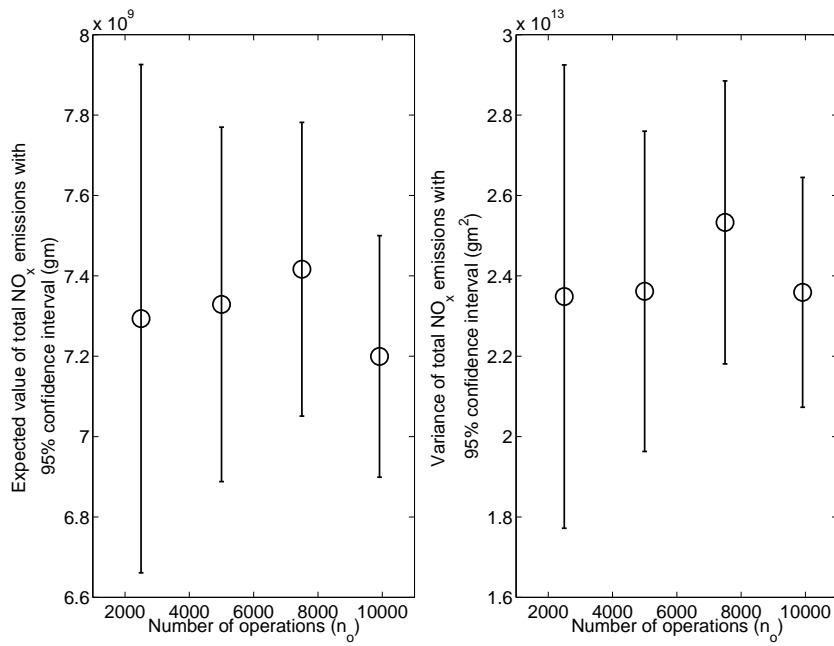
American Institute of Aeronautics and Astronautics

Figure 5. 95% confidence intervals of the mean and variance of total NO$_x$ emissions computed with surrogate models of 2,500, 5,000, 7,500, and 9,914 operations.
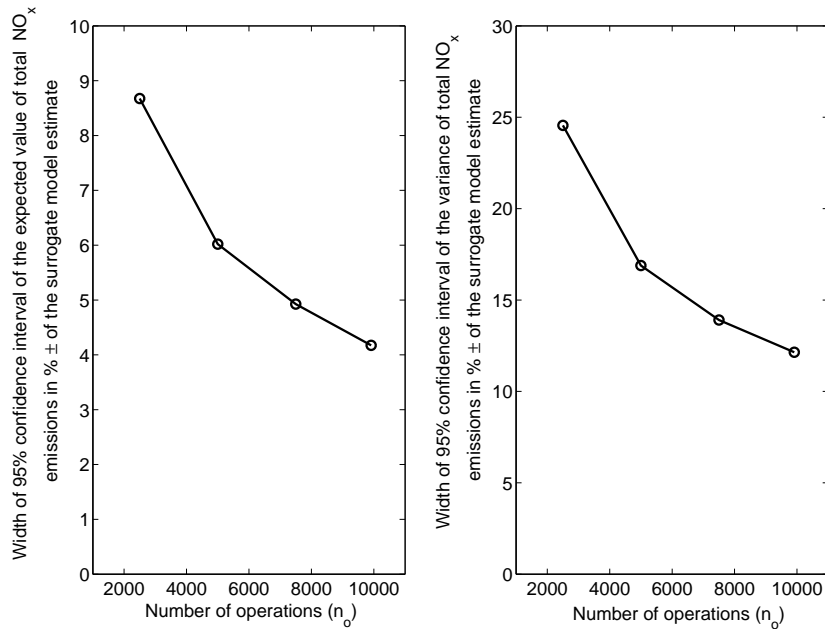

Figure 6. 95 percent confidence interval widths, in terms of percentage $\pm$ of the estimated value, for the mean and variance of total NO$_x$ emissions for a full run of the representative day as n_o increases from 2,500 to 9,914.

uncertainty analysis in support of decision-making, confidence intervals for the mean and variance of total NO$_x$ emissions for the representative day can be constructed. These confidence intervals are quantitatively rigorous and display predictable convergence behavior that can be used to determine optimum tradeoffs

American Institute of Aeronautics and Astronautics

between tighter intervals and longer run times.

AEM Surrogate Results for Model Development Sensitivity Analysis

The total and main effect sensitivity indices were computed using the Sobol' method described in Section II, applied to a surrogate model of $n_o = 5,000$ operations sampled from the representative day. The resulting total and main effect sensitivity indices are shown in Figure 7. These results reveal that factors
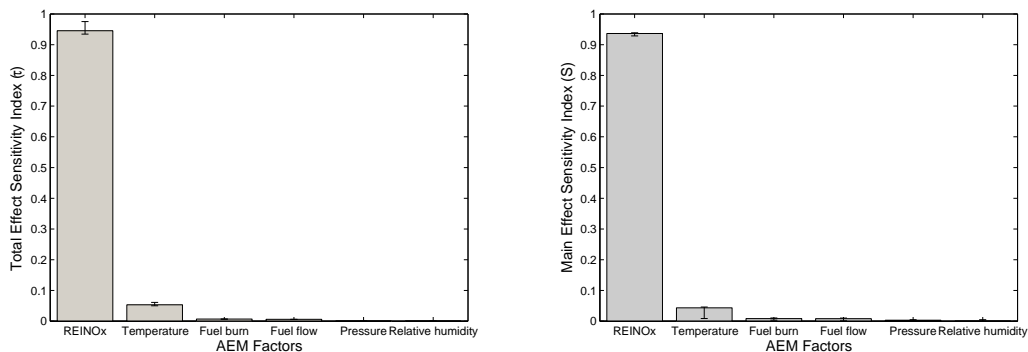

Figure 7. Total and main effect sensitivity indices for the AEM NO_x output as evaluated using a surrogate model consisting of 5,000 operations. The error bars show the 95% confidence intervals for each index.

such as pressure and relative humidity can potentially be fixed for certain analyses since their total effect sensitivity indices are low, and that factors such as the reference emissions index of NO_x and temperature should be the focus of any future research aimed at trimming the variability in total NO_x emissions estimates from the AEM, since their main effect sensitivity indices are highest.

These sensitivity results give valuable insight to guide model development; however, the question arises whether different conclusions might be drawn if the full model were used in place of the surrogate. In this situation, it is computationally impractical to use the full $N_o = 68,343$ operations; however, the sensitivity results computed with the surrogate of $n_o = 5,000$ operations can be rigorously bounded with confidence intervals using the methodology of Section III.B. These confidence intervals were constructed by using a 10,000-iteration Monte Carlo simulation to compute each operation-level global sensitivity index required in (50) and (51). The intervals are shown for each sensitivity index in Figure 7.

Figure 8 shows the convergence behavior of the total effect sensitivity index of the temperature input. The convergence behavior of the other sensitivity indices is similar. Table 2 gives confidence intervals (95%) for the total and main effect sensitivity indices for each input of the AEM for a full run of the representative day computed with a surrogate model of 9,914 operations.

These results show that by applying the surrogate modeling methodology described in Section III for model development sensitivity analysis, confidence intervals for the global sensitivity indices of total NO_x emissions for the full representative day can be constructed from a subset of operations. Just as for the
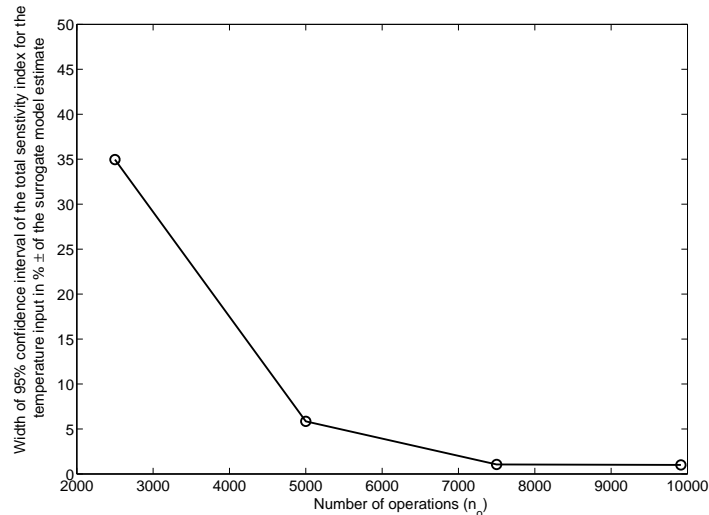
American Institute of Aeronautics and Astronautics

Figure 8. 95 percent confidence interval width, in terms of percentage $\pm$ of the surrogate model sensitivity index estimate, for the total effect sensitivity index of the temperature input as n_o increases from 2500 to 9914.

Table 2. 95 percent confidence intervals of the total and main effect sensitivity indices for each input of the AEM for a full run of the representative day computed with a surrogate model of 9914 operations.

Input	Total effect sensitivity index		Main effect sensitivity index	
	Lower	Upper	Lower	Upper
REINO$_x$	0.935	0.950	0.935	0.940
Temperature	0.050	0.053	0.008	0.060
Fuel burn	0.006	0.007	0.003	0.010
Fuel flow	0.005	0.006	0.001	0.010
Pressure	0.000	0.001	0.000	0.005
Relative Humidity	0.000	0.000	0.000	0.000

confidence intervals constructed to support decision-making uncertainty analysis, these confidence intervals are quantitatively rigorous and display convergence behavior that can be used to determine optimum tradeoffs between tighter intervals and longer run times.

V. Limitations and Additional Sources of Error

The surrogate modeling methodology developed here for the general class of models given by (16) and demonstrated on the AEM is applicable only if several aforementioned assumptions are met. Those assumptions were independence of model inputs, randomness of model inputs, a large number of random quantities being summed, and satisfaction of the Lyapunov condition by those random quantities. As mentioned in Section II.C, the assumptions of independence and randomness in the methods developed here are assumed to be features of the full model, and thus are not restrictions imposed by the surrogate model formulation. If they are not features of the full model, the approach taken here may not be applicable, and analysts may

American Institute of Aeronautics and Astronautics

wish to consider the m-dependent Central Limit Theorem,[22] which for certain situations permits relaxing the independence assumption. For models that fit the form given by (16) but do not involve summing more than about 30 independent random quantities, we must recognize that the assumption of normality in the resulting sum may not be a good one, and confidence intervals should be estimated with recourse to the Student's t-distribution or bootstrapping techniques. The constraint imposed by the Lyapunov condition is easily met in most situations since many examples governed by the model class given by (16), such as those given in Section II.C, consist of random quantities with finite variances whose sum grows without bound as more quantities are added.

Though the surrogate model estimates and their associated confidence intervals presented in Section IV.D are considered rigorous, they are only rigorous in the sense of how they approximate the results that would be obtained using the full model. Additional sources of error that exist and would be quantifiable if comparison to reality were possible, include the error associated with number of Monte Carlo model evaluations used to compute quantities such as the expected value of operation level NO_x emissions, and the fact that the number of random quantities being summed, though large, is not infinite, and thus there is an approximation associated with assuming normality. Since these errors occur for both the surrogate and full models, the errors are not quantified in the surrogate estimates, since the goal of the surrogate modeling methodology was to estimate the outputs of the full model.

VI.   Conclusions

We have presented a novel surrogate modeling methodology designed specifically for supporting decision-making uncertainty analysis and sensitivity analysis for model development for a large-scale aviation environmental policy making model. The surrogate modeling methods developed here allowed for construction of rigorous confidence intervals for metrics that are useful for supporting decision-making (e.g. output means and variances), and for global sensitivity indices, which are useful for informing future research efforts aimed at furthering the development of a model, for a situation where running the analyses on a full model was impractical. Furthermore, the methodology provides predictable convergence behavior of confidence interval widths from the surrogate model estimates, which allows for informed tradeoffs between computation time and uncertainty in the estimation of the various metrics. Here, model structure was exploited to invoke the Central Limit theorem to derive the confidence intervals; however, the method is more generally applicable using bootstrapping techniques.

American Institute of Aeronautics and Astronautics

Acknowledgments

References

[1] Cacuci, D., *Sensitivity and Uncertainty Analysis: Volume I, Theory*, Chapman & Hall/CRC, Boca Raton, Florida, 2003.

[2] Chan, K., Saltelli, A., and Tarantola, S., "Sensitivity Analysis of Model Output: Variance-Based Methods Make the Difference," Proceedings of the 1997 Winter Simulation Conference, 1997.

[3] Homma, T. and Saltelli, A., "Importance Measures in Global Sensitivity Analysis of Nonlinear Models," *Reliability Engineering and System Safety*, Vol. 52, 1996, pp. 1–17.

[4] Saltelli, A., "Global Sensitivity Analysis: An Introduction," Sensitivity Analysis of Model Output, 2005.

[5] Grimmet, G. and Stirzaker, D., *Probability and Random Processes, 3rd Ed.*, Oxford University Press, Oxford; New York, 2006.

[6] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S., *Global Sensitivity Analysis: The Primer*, John Wiley & Sons, Ltd., West Sussex, England, 2008.

[7] Sobol', I., "Theorems and Examples on High Dimensional Model Representation," *Reliability Engineering and System Safety*, Vol. 79, 2003, pp. 187–193.

[8] Saltelli, A. and Bolado, R., "An Alternative Way to Compute Fourier Amplitude Sensitivity Test (FAST)," *Computational Statistics and Data Analysis*, Vol. 26, 1998, pp. 445–460.

[9] Morris, M., "Factorial Sampling Plans for Preliminary Computational Experiments," *Technometrics*, Vol. 33, No. 2, 1991, pp. 161–174.

[10] Saltelli, A., Tarantola, S., and Chan, K., "A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output," *Technometrics*, Vol. 41, No. 1, 1999, pp. 39–56.

[11] Eldred, M., Giunta, S., and Collis, S., "Second-Order Corrections for Surrogate-Based Optimization with Model Hierarchies," In Proceedings of the 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2004.

[12] Venter, G., Haftka, R., and Starnes, J., "Construction of Response Surface Approximations for Design Optimization," *AIAA Journal*, Vol. 36, No. 12, 1998, pp. 2242–2249.

[13] Antoulas, A. C. and Sorensen, D., *Approximation of Large-Scale Dynamical Systems*, SIAM, Philadelphia, Pennsylvania, 2005.

[14] Alexandrov, N., Lewis, R., Gumbert, C., Green, L., and Newman, P., "Approximation and Model Management in Aerodynamic Optimization with Variable-Fidelity Models," *Journal of Aircraft*, Vol. 38, No. 6, 2001, pp. 1093–1101.

[15] Alexandrov, N., Nielsen, E., Lewis, R., and Anderson, W., "First-Order Model Management in Aerodynamic Optimization with Variable-Fidelity Models," In Proceedings of the 8th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, 2000.

[16] Choi, S., Grandhi, R., Canfield, R., and Pettit, C., "Polynomial Chaos Expansion with Latin Hypercube Sampling for Estimating Response Variability," *AIAA Journal*, Vol. 42, No. 6, 2004, pp. 1191–1198.

[17] Robinson, T., Eldred, M., Willcox, K., and Haimes, R., "Strategies for Multi-Fidelity Optimization with Variable Dimensional Hierarchical Models," AIAA paper 2006-1819, presented at the 2nd MDO Specialists Conference, Newport, RI, 2006.

[18] Simpson, T., Peplinski, J., Koch, P., and Allen, J., "Metamodels for Computer-based Engineering Design: Survey and Recommendations," *Engineering with Computers*, Vol. 17, 2001, pp. 129–150.

[19] Messac, A. and Mullur, A., "A Computationally Efficient Metamodeling Approach for Expensive Multiobjective Optimization," *Optim Eng*, Vol. 9, 2008, pp. 37–67.

[20] Karr, A., *Probability*, Springer, New York, 1993.

[21]Lipschutz, S., Schiller, J., and Srinivasan, R., *Schaum's Outline of Theory and Problems of Beginning Finite Mathematics*, McGraw-Hill Professional, New York; London, 2004.

[22]Berk, K., "A Central Limit Theorem for m-Dependent Random Variables with Unbounded m," *The Annals of Probability*, Vol. 1, No. 2, 1973, pp. 352–354.

[23]Horvitz, D. and Thompson, D., "A Generalization of Sampling Without Replacement from a Finite Universe," *Journal of the American Statistical Association*, Vol. 47, No. 260, 1952, pp. 663–685.

[24]Kutner, M., Nachtsheim, C., Neter, J., and Li, W., *Applied Linear Statistical Models, Fifth Edition*, McGraw-Hill, Boston, Massachusetts, 2005.

[25]Kim, B., Fleming, G., Lee, J., Waitz, I., Clarke, J., Balasubramanian, S., Malwitz, A., Klima, K., Locke, M., Holsclaw, C., Maurice, L., and Gupta, M., "System for Assessing Aviation's Global Emissions (SAGE), Part 1: Model Description and Inventory Results," *Transportation Research Part D: Transport and Environment*, Vol. 12, No. 5, 2007, pp. 325–346.

[26]"Partnership for AiR Transportation Noise and Emissions Reduction (PARTNER)," http://web.mit.edu/aeroastro/partner/index.html.

[27]Baughcomb, S., Henderson, S., and Tritz, T., "Scheduled Civil Aircraft Emissions Inventories for 1976 and 1984: Database Development and Analysis," NASA CR-4722, 1996.

American Institute of Aeronautics and Astronautics