

Max M. J. Opgenoord¹

Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
Cambridge, MA 02139
e-mail: mogg@mit.edu

Douglas L. Allaire

Assistant Professor
Department of Mechanical Engineering,
Texas A&M University,
College Station, TX 77843
e-mail: dallaire@tamu.edu

Karen E. Willcox

Professor
Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
Cambridge, MA 02139
e-mail: kwillcox@mit.edu

Variance-Based Sensitivity Analysis to Support Simulation-Based Design Under Uncertainty

Sensitivity analysis plays a critical role in quantifying uncertainty in the design of engineering systems. A variance-based global sensitivity analysis is often used to rank the importance of input factors, based on their contribution to the variance of the output quantity of interest. However, this analysis assumes that all input variability can be reduced to zero, which is typically not the case in a design setting. Distributional sensitivity analysis (DSA) instead treats the uncertainty reduction in the inputs as a random variable, and defines a variance-based sensitivity index function that characterizes the relative contribution to the output variance as a function of the amount of uncertainty reduction. This paper develops a computationally efficient implementation for the DSA formulation and extends it to include distributions commonly used in engineering design under uncertainty. Application of the DSA method to the conceptual design of a commercial jetliner demonstrates how the sensitivity analysis provides valuable information to designers and decision-makers on where and how to target uncertainty reduction efforts.

[DOI: 10.1115/1.4034224]

1 Introduction

During the initial stages of the design of a system, the inputs to and parameters of that system are still very much uncertain, while at the same time decisions made here have the largest effect on risk, performance, and cost of the final design [1–3]. It is therefore of utmost importance to take these uncertainties into account, both when making design decisions and also when considering the allocation of resources (people, time, financial budget, computer resources, experimental resources, etc.) throughout the design process. Techniques for the quantitative evaluation of the impacts of these uncertainties in engineering design are increasingly being studied and developed. Two significant aspects of uncertainty quantification in engineering design are uncertainty analysis and sensitivity analysis. Uncertainty analysis approaches, such as sampling-based techniques [4–7], surrogate modeling techniques [8–10], and surrogate uncertainty representation techniques [11–14], are all employed to ensure the efficient propagation of uncertainty from model inputs to model outputs. Once this has been achieved, sensitivity analysis in the form of screening methods [15], global sensitivity analysis [4,16–18], entropy based methods [4,19], and moment independent methods [20–22] are often conducted to quantitatively evaluate the impact of a given input on output quantities of interest.

Sensitivity analysis plays an important role in engineering decisions under uncertainty, by quantifying the effects of uncertainties on the output quantities of interest. For example, sensitivity analysis informs a ranking of the most influential uncertain input parameters and it yields insight into how reducing uncertainty in those inputs might correspondingly reduce uncertainty in the output quantities of interest [16]. This paper explores these ideas and extends the so-called distributional sensitivity analysis (DSA) that targets decision-making for the allocation of resources throughout the early stages of engineering design.

One goal of sensitivity analysis of model output is “to ascertain how a given model (numerical or otherwise) depends on its input factors” [23]. One way of performing sensitivity analysis is to find the linearized sensitivity of an output with respect to an input

around a baseline, the so-called *local approach* [16]. These methods are usually cheap and based on computing derivatives; however, they do not paint a complete picture, because they are inherently local. Instead, a *global sensitivity analysis* (GSA) explores the full space of the input factors [16]. A variance-based GSA apportions the variance of a quantity of interest among the contributions from each uncertain factor and from interactions among factors.

GSA results are often used in factor prioritization, which ranks the input parameters based on the expected influence that uncertainty reduction in those parameters has on the system output [17,18]. In this method, one ranks the factors based on which factor, if fixed to its true value, yields the largest expected reduction in output uncertainty; however, Oakley and O’Hagan [24] noted that it is rarely possible to fix a factor to its true value. Allaire and Willcox addressed this limitation of GSA by developing a *distributional sensitivity analysis* (DSA) [25], which computes the influence on the output variance as a function of the amount of variance reduction in the input parameters. This results in a variance-based sensitivity index function. This function can be used directly to inform decision-making, or alternatively, by considering the amount of variance reduction in each factor as a random variable, the mean value of the variance-based sensitivity index function provides an average main effect sensitivity index used for ranking purposes.

This paper expands and extends the DSA approach to create a sensitivity-based methodology to support resource allocation decisions in engineering design under uncertainty. Specifically, this paper contributes methods to improve the quality and efficiency of the sampling within the DSA, which is essential for application to computational expensive models. Section 2 presents the problem setup. Section 3 describes the overall methodology. We apply the method to an engineering system—the conceptual design of a commercial jetliner—in Sec. 4. Section 5 concludes the paper.

2 Problem Setup

We represent a general engineering system of interest as

$$y = g(\mathbf{x}) \quad (1)$$

where y is the quantity of interest (QoI) and \mathbf{x} is the vector of system input variables. For clarity of exposition, we consider a scalar

¹Corresponding author.

Contributed by the Design Automation Committee of ASME for publication in the JOURNAL OF MECHANICAL DESIGN. Manuscript received March 5, 2016; final manuscript received June 30, 2016; published online September 12, 2016. Associate Editor: Samy Missoum.

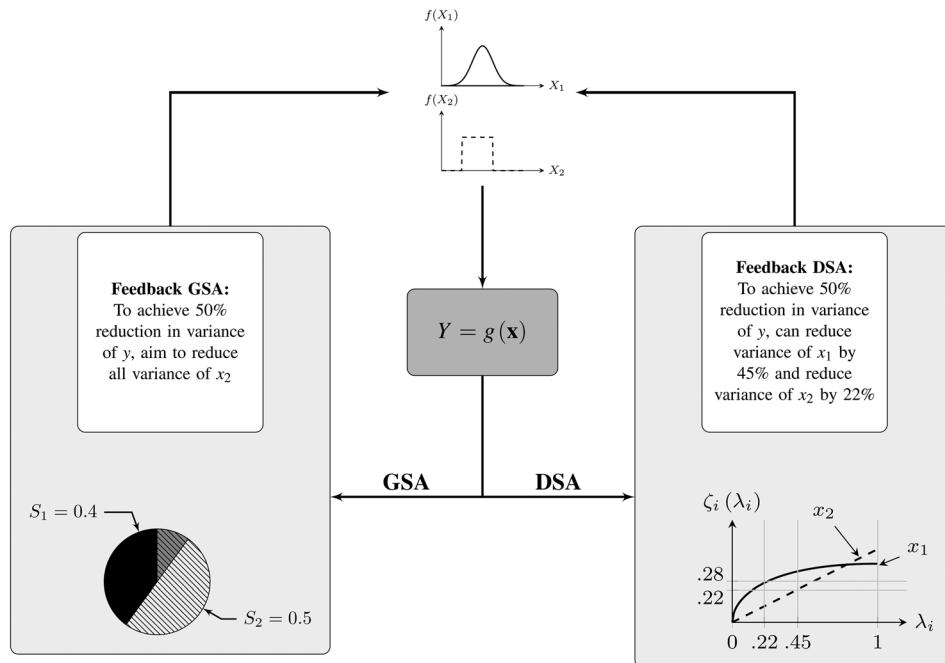


Fig. 1 Comparison of design process using global sensitivity analysis and distributional sensitivity analysis

quantity of interest, although the methodology is straightforward to apply when there are multiple quantities of interest. These quantities typically represent design objectives and constraints, such as the fuel consumption, maximum takeoff weight, and environmental performance of an aircraft. We consider n input variables, $\mathbf{x} = [x_1, \dots, x_n]^\top$, where x_i denotes the i th input variable. The input parameters represent quantities within the designer's control, such as aircraft flight speed, aircraft geometry, density of the materials used, etc. We consider the model that maps the inputs to the output quantity of interest, $g(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$, to be a black box function that can be used to evaluate y at any point in the input space.

We consider the input parameters \mathbf{x} to be uncertain, which in turn induces uncertainty in the output quantity of interest y . In this paper, our focus is on the task of deciding how to allocate resources in the design process. That is, we pose the question: *How should the designer allocate design effort to reduce uncertainty in the input parameters, so as to appropriately manage uncertainty in the output quantity of interest?* For example, in the conceptual design of an aircraft, the uncertain input parameters represent different aspects of the aircraft design, including aerodynamic performance, structures, weights, stability, controls, propulsion, and environmental performance. At any current state of the design, uncertainty in the inputs for each of these disciplines results in uncertainty in the corresponding system-level metrics (the quantities of interest). Our goal is to provide quantitative guidance to the designer on which disciplines to focus design effort and by how much to reduce uncertainty in each discipline, so as to achieve desired confidence levels on the quantities of interest. This information is especially important in early-stage design where time and money resources are limited, yet critical decisions are made that lock in later options.

This paper builds on the work of Allaire and Willcox [25] on the distributional sensitivity analysis (DSA) methodology, which is proposed as a basis for providing this kind of quantitative guidance. A DSA provides not only an average main effect sensitivity index that can be used to prioritize efforts but also a function that estimates uncertainty reduction in y as a function of the uncertainty reduction in each component of \mathbf{x} . This information can be translated directly into deciding how much effort to expend on

reducing uncertainty in a given parameter (e.g., deciding what fidelity of analyses to run, whether to conduct an experiment, whether to task a disciplinary team with conducting more detailed design work, etc.).

A simple notional example to illustrate these ideas is presented in Fig. 1. On the left, a standard GSA computes the sensitivity indices that apportion the variance in the output y : 50% from input parameter x_2 , 40% from input parameter x_1 , and the remaining 10% from the interaction between x_1 and x_2 . These sensitivity indices are based on the premise that all variance in each parameter is reduced. On the right, a DSA computes the sensitivity index function, which apportions the expected reduction in output variance as a function of the amount of input variance reduction. For illustration, consider the case that the goal is to reduce the uncertainty of the output y by 50%. A prioritization of effort based on the GSA would suggest focusing on x_2 : reducing all the variance in x_2 is expected to yield the desired reduction in output variance. One could perhaps decide to allocate effort between x_2 and x_1 , but the GSA results provide no guidance on how to do this and no indication of the expected result on uncertainty reduction in the output. Using the DSA results, however, we can combine partial variance reductions of each input parameter and choose our reductions accordingly. For example, the DSA results show that reducing variance in x_1 by 45% and reducing variance in x_2 by 22% together provide an expected reduction in output variance of 50%. From Sec. 3.2 onward, we further explain the methodology that permits these ideas to be applied in the engineering design setting.

3 Methodology

Global sensitivity analysis and distributional sensitivity analysis are introduced in Secs. 3.1 and 3.2, respectively. Methodologies for computing the sensitivity index function for DSA are introduced in Sec. 3.3, while Sec. 3.4 focuses on the surrogate modeling technique used throughout this work.

3.1 Global Sensitivity Analysis (GSA). Variance-based global sensitivity analysis computes a variance decomposition of the output variance to apportion it among the input variables. This provides information about the influence of these variables on the

output variance, if their respective uncertainties are entirely reduced. The decomposition of the variance of the random variable Y can be written as

$$\text{var}(Y) = \mathbb{E}[\text{var}(Y|X_i)] + \text{var}(\mathbb{E}[Y|X_i]) \quad \text{for } i = 1, \dots, n \quad (1)$$

where X_i denotes the random variable representing the uncertain input x_i .

The quantity $\mathbb{E}[\text{var}(Y|X_i)]$ is the unexplained variance—the part of the variance in Y that remains after X_i is fixed somewhere in its domain. GSA exploits this definition to compute the main effect sensitivity indices, defined as

$$S_i = \frac{\text{var}(\mathbb{E}[Y|X_i])}{\text{var}(Y)} \quad (2)$$

where S_i is the main effect sensitivity index of input variable X_i . These main effect sensitivity indices are usually computed using either the Fourier amplitude sensitivity test (FAST) or the Sobol' method [18,26,27]. The FAST method uses Fourier transforms, while the Sobol' method employs Monte Carlo simulation. The Sobol' method is computationally more intensive and depends on the chosen sample size, but it can more easily compute the higher-order sensitivity indices, which is required when there are interactions between the parameters [28]. A different approach to compute a GSA, employed in this work, is to use a high-dimensional model representation (HDMR) and build up the ANOVA-HDMR.

High-Dimensional Model Representation (HDMR) is used to represent high-dimensional relations between inputs and outputs [29–31]. HDMR is a representation that decomposes the model $y = g(\mathbf{x})$ into component functions corresponding to the input variables acting alone, the interactions between two input variables, and so on [32]

$$\begin{aligned} g(\mathbf{x}) = g_0 + \sum_{j=1}^n g_i(x_i) + \sum_{1 \leq i < j \leq n} g_{ij}(x_i, x_j) + \dots \\ + \sum_{1 \leq i_1 < \dots < i_l \leq n} g_{i_1 i_2 \dots i_l}(x_{i_1}, x_{i_2}, \dots, x_{i_l}) + \dots \\ + g_{12 \dots n}(x_1, x_2, \dots, x_n) \end{aligned} \quad (3)$$

Here, g_0 represents the mean response to $g(\mathbf{x})$ over the input space. The first-order component function $g_i(x_i)$ captures the independent contribution of the i th input variable alone. The component subfunction $g_{ij}(x_i, x_j)$ represents the effect of the interaction between the i th and j th input variable on $g(\mathbf{x})$, and so on, for the higher-order terms.

The decomposition as written in Eq. (3) is not unique [33]. The vanishing condition [26,33–35] states that the integral of an HDMR component function with respect to any of its own variables is zero

$$\int f_s(x_s) g_{i_1 i_2 \dots i_k}(x_{i_1}, x_{i_2}, \dots, x_{i_k}) dx_s = 0 \quad \forall s \in \{i_1, i_2, \dots, i_k\} \quad (4)$$

where $f_s(x_s)$ is the probability density function of the s th input variable.

Due to the orthogonality constraints in Eq. (4), the variance of the QoI can be decomposed as

$$V(y) = \sum_i V_i + \sum_{1 \leq i < j \leq n} V_{ij} + \dots + V_{12 \dots n} \quad (5)$$

where V_i is the contribution to the QoI variance of input variable X_i

$$V_i = V(g_i(X_i)) = V[\mathbb{E}(Y|X_i)] \quad (6)$$

and V_{ij} the contribution of the interactions between X_i and X_j

$$V_{ij} = V(g_{ij}(X_i, X_j)) = V[\mathbb{E}(Y|X_i, X_j)] - V[\mathbb{E}(Y|X_i)] - V[\mathbb{E}(Y|X_j)] \quad (7)$$

Similar expressions can be built up for the third-order terms V_{ijk} and higher-order terms. The decomposition of sensitivity indices is then obtained by dividing through the total variance

$$\sum_i S_i + \sum_{1 \leq i < j \leq n} S_{ij} + \dots + S_{12 \dots n} = 1 \quad (8)$$

where $S_i = V_i/V$ are the main effect sensitivity indices, as defined in Eq. (2). The second-order sensitivity index is $S_{ij} = V_{ij}/V$, and so on, for the higher-order sensitivity indices.

3.2 Distributional Sensitivity Analysis. In GSA, the factor prioritization is based on the main effect sensitivity indices in Eq. (2), and sometimes also on the total sensitivity indices—the sum of all sensitivity indices to which the i th variable contributes. Such a factor prioritization strategy relies on the assumption that all uncertainties of a particular input variable can be reduced. Distributional sensitivity analysis relaxes this assumption by treating the amount of variance reduction in a particular input variable as a random variable instead of assuming it can be reduced completely to zero [25]. For a given amount of variance reduction in the input variable X_i , DSA considers a family of reasonable distributions and calculates an average change in the variance of Y over this family. In the design setting this is particularly useful, because reducing the uncertainty in an input parameter to zero is generally an impractical option.

In DSA, we define a variance-based sensitivity index function ζ_i for variable X_i . When reducing the variance of X_i by a percentage $100(1 - \lambda_i)\%$, this variance-based sensitivity index function ζ_i is defined as

$$\zeta_i(\lambda_i) = \frac{\text{var}(Y^0)S_i^0 - \mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i]}{\text{var}(Y^0)} \quad (9)$$

Here, $\text{var}(Y^0)$ and S_i^0 denote the variance of Y and the main effect sensitivity index of X_i , respectively, corresponding to the original input distribution used to represent input variable x_i . X'_i is an updated random variable for input x_i , which results from design effort to reduce that variable's uncertainty (e.g., further research, experiments, higher-fidelity modeling, etc.). $\text{var}(Y')$ is the updated variance of Y and S'_i is the updated main effect sensitivity index, both associated with the updated random variable X'_i . The quantity $\lambda_i = \text{var}(X'_i)/\text{var}(X_i)$ is the ratio of the variance of X_i that is not reduced to the variance of the original distribution. The variance-based sensitivity index function ζ_i is therefore the relative difference between the global sensitivity index and the expected value of the updated variance with the sensitivity index if we were to only reduce the uncertainty partially. When $\lambda_i = 1$, this expected value is zero and therefore the variance-based sensitivity index function reduces to the global sensitivity index.

Since it is not known how we can reduce the uncertainty in the input variables, we take the expected value of the reduction in variance of Y over reasonable distributions. $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i]$ is the expected value of the product of the updated sensitivity indices and the updated variance of Y , taken over reasonable distributions. Those reasonable distributions are further defined below for different distribution families. However, it is unknown how much uncertainty can be reduced before doing further research, therefore the parameter λ_i is considered a uniform random variable Λ_i on $[0, 1]$. We note that further research could indicate that the uncertainty of the input variable increases, but that would mean the original input distribution was flawed. Therefore, $\lambda_i > 1$ is not considered in this work. Other distributions for λ could also be considered; for instance, one might wish to bias λ toward smaller reductions in uncertainties. This work could easily accommodate

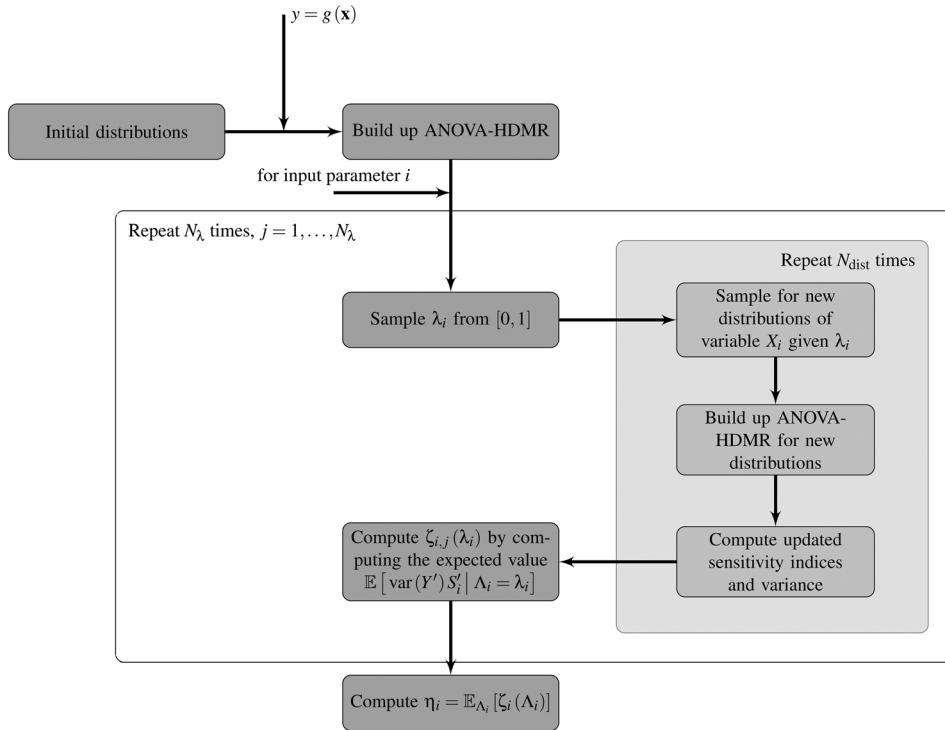


Fig. 2 Flow chart for distributional sensitivity analysis

such a model, but it is not included in this paper because choosing such a model is problem dependent.

For factor prioritization using DSA, we consider the expected value of $\zeta_i(\lambda_i)$ to obtain the *average main effect sensitivity index* (η) as

$$\eta_i = \mathbb{E}_{\Lambda_i} [\zeta_i(\Lambda_i)] \quad (10)$$

While the variance-based sensitivity index function provides more insight about how reducing a factor's uncertainty impacts variance in the output, the average main effect sensitivity index is useful because it permits us to perform factor prioritization. In particular, the average main effect sensitivity index can be used to rank input variables based on the average amount of variance of Y that can be reduced by doing further research on a particular input variable. This is analogous to the typical ranking in a conventional GSA. Note that as written, these average main effect sensitivity indices do not add up to one, whereas the variance-based sensitivity indices do. However, the values of these η_i indices are important in an absolute sense, because they relate to the expected reduction of variance.

Figure 2 shows the overall DSA approach, where N_λ is the number of samples used to compute η_i and N_{dist} is the number of times a new distribution is drawn for a given λ_i . ζ_i is the variance-based sensitivity index function for the i th input parameter corresponding to the j th sample of λ_i .

3.3 Modeling Variance Reduction. This section defines the reasonable distributions over which we sample to obtain ζ_i . We focus on distributions that are relevant in an engineering design setting. The uniform distribution is often used when limited information is available about the input parameter, but its range is available (or can be estimated). The normal distribution is often used for error modeling (e.g., model error in a simulation code). A triangular distribution is often used in an engineering setting when the designer has insight about a most-likely value in addition to the variable's range. The general DSA methodology is extensible to other distributions; here, we present the details for these three

distributions. We build off the work in Refs. [25,36]. Algorithm 1 is taken directly from Ref. [25], whereas Algorithm 1 is a variation where we also compute the variance-based sensitivity function with quadrature. For the normal distribution, Algorithms 3 and 4 deviate from Ref. [25] because we allow the mean of the updated distribution to vary as well. Finally, for the triangular distribution, Algorithm 5 uses the same idea as in Ref. [25], but uses a different implementation.

3.3.1 Uniform Distribution. The variance of a uniformly distributed random variable $X \sim \mathcal{U}[a, b]$ is $\text{var}(X) = (b - a)^2/12$. Consider the original distribution to be $X \sim \mathcal{U}[a^0, b^0]$ and the updated distribution to be $X' \sim \mathcal{U}[a', b']$. Then we have $\lambda_i = [(b' - a')/(b^0 - a^0)]^2$. All updated distributions for a given λ_i have the same width, as $\lambda_i^{1/2}(b^0 - a^0)$. In order to compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i]$, we sample over different distributions using Algorithm 1.

Algorithm 1: Computing the variance-based sensitivity index function for a uniform distribution using sampling (from Ref. [25]).

- 1 Sample λ_i from a uniform distribution on the interval $[0, 1]$;
- 2 **for** $j \leftarrow 1$ **to** N_{dist} **do**
- 3 Sample b' from a uniform distribution on the interval $[a^0 + \sqrt{\lambda_i}(b^0 - a^0), b^0]$;
- 4 Let $a' = b' - \sqrt{\lambda_i}(b^0 - a^0)$;
- 5 Compute $\text{var}(Y')_j$ and S'_i for the distribution $\mathcal{U}[a', b']$;
- 6 **end**
- 5 Compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i] = \frac{1}{N_{\text{dist}}} \sum_j \text{var}(Y')_j S'_i$;
- 6 Compute variance-based sensitivity index function using Eq. (9) for λ_i .

An improved approach uses Gauss-Legendre quadrature to compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i]$, instead of sampling over new distributions for every λ_i . By using Gauss-Legendre quadrature, we can specify the mean and standard deviation of the new distribution. This is not possible with a sampling-based method, because one is dependent on the original samples and may therefore not get the exact mean and variance. To solve for the

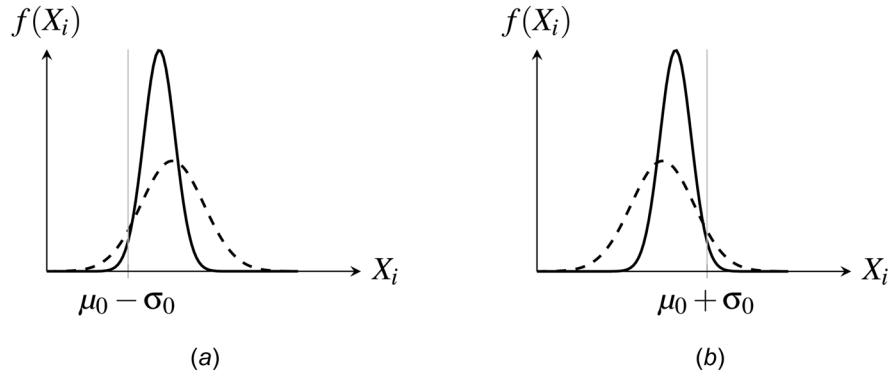


Fig. 3 Reasonable distributions for the normal distribution. In this example $\lambda_i = 0.5$: (a) lower bound for the updated distribution and (b) upper bound for the updated distribution.

expected value over the mean of the updated distributions, we rewrite this expected value as an integral over μ' , the mean of the updated distribution

$$\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i] = \frac{\int \text{var}(Y')S'_i d\mu'}{\int d\mu'} \Big|_{\Lambda_i = \lambda_i} \quad (11)$$

A method for computing the variance-based sensitivity index function ζ_i using Gauss-Legendre quadrature is given in Algorithm 2.

Algorithm 2: Computing the variance-based sensitivity index function for a uniform distribution using Gauss-Legendre quadrature.

- 1 Sample λ_i from a uniform distribution on the interval $[0, 1]$;
- 2 Let lower bound for the mean μ_l be $a^0 + ((1/2)\sqrt{\lambda_i}(b^0 - a^0))$ and the upper bound μ_u be $b^0 - ((1/2)\sqrt{\lambda_i}(b^0 - a^0))$;
- 3 **for** $j \leftarrow 1$ **to** N_{Gauss} **do**
- 4 Let μ_j be Gauss-Legendre quadrature node on the interval $[\mu_l, \mu_u]$;
- 5 Let $d' = \mu_j - \frac{1}{2}\sqrt{\lambda_i}(b^0 - a^0)$ and $b' = \mu_j + \frac{1}{2}\sqrt{\lambda_i}(b^0 - a^0)$;
- 6 Compute $\text{var}(Y')_j$ and $S'_{i,j}$ for the distribution $\mathcal{U}[d', b']$;
- 7 **end**
- 8 Compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i] = \sum_j w_j \text{var}(Y')_j S'_{i,j} / (\mu_u - \mu_l)$, where w_j are the Gauss-Legendre quadrature weights on the domain $[\mu_l, \mu_u]$;
- 9 Compute variance-based sensitivity index function using Eq. (9) for λ_i

3.3.2 Normal Distribution. The procedure for computing the variance-based sensitivity ζ_i for a normal distribution is similar to that for a uniform distribution. Let the original distribution be given by $\mathcal{N}(\mu_0, \sigma_0^2)$ and the updated distribution by $\mathcal{N}(\mu', \sigma'^2)$, where $\sigma'^2 = \lambda_i \sigma_0^2$. A method for computing ζ_i by sampling over possible distributions is given in Algorithm 3. Considering the support of the normal distribution is infinite, we need to bind the mean of the updated distribution to sample over it. This is similar to the method for the uniform distribution where we only allow for the updated distributions to be confined to the support of the original distribution. We therefore only allow the mean of the updated distribution to vary within the interval $\mu' \in [\mu'_l, \mu'_u]$, where μ'_l is the minimum allowable mean and μ'_u is the maximum allowable mean of the updated distributions. Here we choose μ'_l and μ'_u such that $\mu'_l - \sigma' = \mu_0 - \sigma_0$ and $\mu'_u + \sigma' = \mu_0 + \sigma_0$. This is also illustrated in Fig. 3. One could choose wider bounds, but we do not want the updated distribution to extend too far into the tails of the original distribution, because these regions had a low probability in the original distribution.

Algorithm 3: Computing the variance-based sensitivity index function for a normal distribution using sampling.

- 1 Sample λ_i from a uniform distribution on the interval $[0, 1]$;
- 2 **for** $j \leftarrow 1$ **to** N_{dist} **do**
- 3 Sample μ' from a uniform distribution on the interval $[\mu_0 - (1 - \sqrt{\lambda_i})\sigma_0, \mu_0 + (1 - \sqrt{\lambda_i})\sigma_0]$;
- 4 Compute $\text{var}(Y')_j$ and $S'_{i,j}$ for the distribution $\mathcal{N}[\mu', \lambda_i \sigma_0^2]$;
- 5 **end**
- 6 Compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i] = \frac{1}{N_{\text{dist}}} \sum_j \text{var}(Y')_j S'_{i,j}$;
- 7 Compute variance-based sensitivity index function using Eq. (9) for λ_i

Again, instead of sampling over the distributions we can also solve for $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i]$ using Gauss-Legendre quadrature. The algorithm then becomes as given in Algorithm 4.

Algorithm 4: Computing the variance-based sensitivity index function for a normal distribution using Gauss-Legendre quadrature.

- 1 Sample λ_i from a uniform distribution on the interval $[0, 1]$;
- 2 Let lower bound for the mean μ_l be $\mu_0 - (1 - \sqrt{\lambda_i})\sigma_0$ and the upper bound $\mu_0 + (1 - \sqrt{\lambda_i})\sigma_0$;
- 3 **for** $j \leftarrow 1$ **to** N_{Gauss} **do**
- 4 Let μ_j be Gauss-Legendre quadrature node on the interval $[\mu_l, \mu_u]$;
- 5 Compute $\text{var}(Y')_j$ and $S'_{i,j}$ for the distribution $\mathcal{N}[\mu_j, \lambda_i \sigma_0^2]$;
- 6 **end**
- 7 Compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i] = \sum_j w_j \text{var}(Y')_j S'_{i,j} / (\mu_u - \mu_l)$, where w_j are the Gauss-Legendre quadrature weights on the domain $[\mu_l, \mu_u]$;
- 8 Compute variance-based sensitivity index function using Eq. (9) for λ_i

3.3.3 Triangular Distribution. The triangular distribution has three parameters, which makes the sampling more complicated than for the normal and uniform distributions. We present two different ways of sampling and it depends on the problem at hand which strategy makes most sense. Let $\mathcal{T}(a^0, b^0, c^0)$ be the original input distribution and $\mathcal{T}(a', b', c')$ the updated distribution. For a triangular distribution $\mathcal{T}(a, b, c)$ ² the variance is $(a^2 + b^2 + c^2 - ab - ac - bc)/18$. Therefore, λ_i is given by

$$\lambda_i = \frac{a'^2 + b'^2 + c'^2 - a'b' - a'c' - b'c'}{a^{02} + b^{02} + c^{02} - a^0b^0 - a^0c^0 - b^0c^0} \quad (12)$$

The first strategy considers distributions whose shape does not change. That is, if all distributions were mapped to the $[0, 1]$ domain, they would overlay one another. The maximum allowable shift in mean is then constrained by the fact that the support of the

²Following standard notation, a denotes the minimum value in the distribution, b the maximum value of the distribution, and c the most likely value of the distribution.

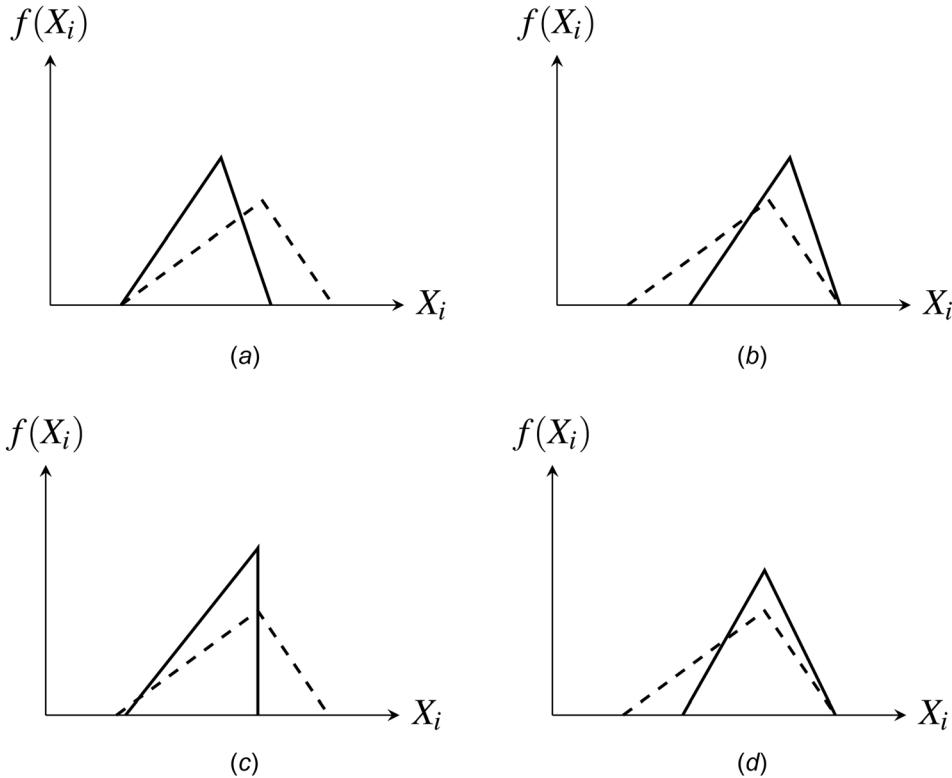


Fig. 4 Reasonable distributions for different strategies for triangular distributions. In this example, $\lambda_i = 0.5$: (a) Lower bound for strategy where shape is kept fixed, (b) upper bound for strategy where shape is kept fixed, (c) lower bound for strategy where most-likely value is kept fixed, and (d) upper bound for strategy where most-likely value is kept fixed.

updated distribution should be within the support of the original distribution, i.e., $a' \geq a^0$ and $b' \leq b^0$. This strategy is illustrated in Figs 4(a) and 4(b) and the algorithm for it is shown in Algorithm 5.

Algorithm 5: Computing the variance-based sensitivity index function for a triangular distribution using Gauss–Legendre quadrature where the shape of the distribution is kept constant.

- 1 Sample λ_i from a uniform distribution on the interval $[0, 1]$;
- 2 Let lower bound for a' be $a'_l = a^0$ and the upper bound $a'_u = a^0 + (1 - \sqrt{\lambda_i})(b^0 - a^0)$;
- 3 for $j \leftarrow 1$ to N_{Gauss} do
 - 4 Let a' be $x_j(a'_u - a'_l)/2 + (a'_l + a'_u)/2$, where x_j is the Gauss–Legendre quadrature node on the interval $[0, 1]$;
 - 5 Compute $\text{var}(Y')_j$ and $S'_{i,j}$ for the distribution $\mathcal{T}[a', b', c']$;
 - 6 end
- 7 Compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i] = \sum_j w_j \text{var}(Y')_j S'_{i,j}$, where w_j are the Gauss–Legendre quadrature weights on the domain $[0, 1]$;
- 8 Compute variance-based sensitivity index function using Eq. (9) for λ_i

In the second strategy, the most-likely value of the distribution is kept fixed and we sample distributions around that, as proposed by Allaire [36]. Thus, we consider $c^0 = c'$. At the same time, the support of the updated distribution has to be within the support of the original distribution, i.e., $a' \geq a^0$ and $b' \leq b^0$. The algorithm for computing the variance-based sensitivity index function using this strategy is shown in Algorithm 6.

Algorithm 6: Computing the variance-based sensitivity index function for a triangular distribution where the most-likely value is kept constant.

- 1 Sample λ_i from a uniform distribution on the interval $[0, 1]$;
- 2 %Compute minimum a'
- 3 Solve for a'_l using Eq. (12) with $b' = c^0$ and $c' = c^0$;
- 4 if $a'_l < a^0$ then
- 5 3 $a'_l = a^0$;
- 6 end
- 7 %Compute maximum a'
- 8 Solve for b'_u using Eq. (12) with $a' = c^0$ and $c' = c^0$;
- 9 if $b'_u > b^0$ then
- 10 5 Solve for a'_u using Eq. (12) with $b' = b^0$ and $c' = c^0$;
- 11 else
- 12 6 $a'_u = c^0$;
- 13 end
- 14 for $j \leftarrow 1$ to N_{dist} do
- 15 7 Sample a' from a uniform distribution on the interval $[a'_l, a'_u]$ and set $c' = c^0$;
- 16 8 Solve for a' using Eq. (12);
- 17 9 Compute $\text{var}(Y')_j$ and $S'_{i,j}$ for the distribution $\mathcal{T}[a', b', c']$;
- 18 end
- 19 Compute $\mathbb{E}[\text{var}(Y')S'_i | \Lambda_i = \lambda_i] = \frac{1}{N_{\text{dist}}} \sum_j \text{var}(Y')_j S'_{i,j}$;
- 20 Compute variance-based sensitivity index function using Eq. (9) for λ_i

3.4 HDMR-Based Surrogate. Computing the sensitivity indices for every new distribution using the full model $g(\mathbf{x})$ is prohibitively expensive. Reference [25] addressed this challenge by reusing the original Monte Carlo samples from the global sensitivity analysis by means of rejection sampling. We instead use the ANOVA-HDMR both to reduce the initial cost of performing the GSA and to evaluate updated distributions where rejection sampling is inaccurate. One such example is the triangular distribution: when the most-likely value of the updated distribution is

almost at the tail of the original distribution, there are very few samples one can use to compute the sensitivity indices.

Building up the full ANOVA-HDMR as in Eq. (3) is often unnecessary. For many systems and models, the third-order and higher terms are negligibly small [29]. That allows us to truncate the HDMR by neglecting these higher order terms, yielding the approximation [30]

$$g(\mathbf{x}) \approx g_0 + \sum_{j=1}^n g_i(x_i) + \sum_{1 \leq i < j \leq n} g_{ij}(x_i, x_j) \quad (13)$$

Following Ref. [33], an accurate and fast surrogate model can be built up by approximating the remaining component functions as expansions of an appropriate set of basis functions

$$g_i(x_i) \approx \sum_{r=1}^{\ell} \alpha_r^i \varphi_r(x_i) \quad (14)$$

$$g_{ij}(x_i, x_j) \approx \sum_{p=1}^{\ell} \sum_{q=1}^{\ell} \beta_{pq}^{ij} \varphi_{pq}(x_i, x_j) \quad (15)$$

where $g_i(x_i)$ is approximated using ℓ basis functions $\varphi_1, \dots, \varphi_{\ell}$, and α_r^i is the coefficient for the r th basis function, corresponding to the i th input variable. In the same way, β_{pq}^{ij} is the coefficient corresponding to the pq th basis function and the ij th second-order component function. Here, we take $\varphi_{pq}(x_i, x_j) = \varphi_p(x_i) \varphi_q(x_j)$.

In this work, we choose orthonormal polynomials as basis functions. The choice of orthonormal basis functions leads to the following evaluations of the coefficients of the component functions:

$$\alpha_r^i = \int g(\mathbf{x}) \varphi_r(x_i) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (16)$$

$$\beta_{pq}^{ij} = \int g(\mathbf{x}) \varphi_p(x_i) \varphi_q(x_j) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (17)$$

Because we need orthonormal basis functions with respect to the distribution, different basis functions are needed for every distribution considered. For the uniform distribution, we use normalized shifted Legendre polynomials, and the integrals in Eqs. (16) and (17) are then solved efficiently using Gauss-Legendre quadrature. For the normal distribution, we use the normalized probabilists' Hermite polynomials, and we solve the integrals for every α_r^i and β_{pq}^{ij} in Eqs. (16) and (17) efficiently using (probabilists') Gauss-Hermite quadrature. In building up the ANOVA-HDMR for the triangular distributions, one has to deal with the discontinuous derivative of the probability density function. This complicates finding the correct basis functions and also requires splitting the integral into multiple parts. Furthermore, the basis functions now also depend on the shape of the probability density function directly. This in contrast to the uniform and normal distribution, where we map $\mathcal{U}(a, b)$ to $\mathcal{U}(0, 1)$ and $\mathcal{N}(\mu, \sigma^2)$ to $\mathcal{N}(0, 1)$, respectively. Here, we map $\mathcal{T}(a, b, c)$ to $\mathcal{T}(0, 1, \mu)$, which means that the basis functions depend on μ and therefore depend directly on the shape of the probability density function. The first three basis functions for the triangular distributions are derived by Wang et al. [34] and are used here.

Following the strategy outlined in Fig. 2, we build up the ANOVA-HDMR for the original input distributions and then reuse that surrogate, instead of the black-box model, to find the new standard deviation of the QoI for every updated input distribution. In order to find the sensitivity indices for this updated input distribution, we build up a new ANOVA-HDMR for the new distribution by computing an integral over the new input space using the original ANOVA-HDMR. In other words, we re-solve for α_r^i and β_{pq}^{ij} using Eqs. (16) and (17) with the new distributions and replacing $g(\mathbf{x})$ with the original ANOVA-HDMR.

We note that other choices of surrogate modeling technique are possible. Our choice of HDMR surrogate model is based on our goal of estimating variance-based sensitivity indices. An HDMR surrogate model approximates the very subfunctions needed to compute these sensitivities, thus it is a natural choice in this setting. In particular, a polynomial chaos expansion might be a suitable alternative choice [37,38], particularly when tailored for computation of global sensitivity indices [39]. One limitation of HDMR is that for higher-dimension problems (e.g., greater than dimension five or six), evaluation of the integrals to estimate the coefficients in the HDMR expansion will be computationally expensive, even with quadrature rules. One way to address this is to use the cut-HDMR [30], which we employ in a high-dimensional example in Sec. 4. A polynomial chaos expansion could address high-dimensional problems by using a sparse polynomial basis and evaluating the coefficients with regression.

4 Results

The proposed DSA method is employed on a test function in Sec. 4.1 to assess the convergence and the accuracy of the surrogate. Section 4.2 presents results for an engineering system—the conceptual design of a commercial jetliner aircraft.

4.1 Test Function Analysis. To assess the convergence of the method, we consider the Ishigami function, which is also used in Ref. [25]

$$Y = \sin X_1 + a \sin^2 X_2 + b X_3^4 \sin X_1 \quad (18)$$

with $a = 5$ and $b = 0.1$, as used in Ref. [40]. Here, the X_i 's are considered to be independent and uniformly distributed on $\mathcal{U}[-\pi, \pi]$. The main effect sensitivity indices for this function and these distributions are $S_1 = 0.40$, $S_2 = 0.29$ and $S_3 = 0.00$ and the only nonzero interaction term is $S_{1,3} = 0.31$. These were computed using the analytical expressions in Ref. [16].

The variance-based sensitivity index function as computed using Algorithm 2 is shown in Fig. 5. These results were generated using an ANOVA-HDMR consisting of up to eighth-order basis functions and using nine quadrature points in one dimension. With three input variables, we therefore need a total of 729 function evaluations of Eq. (18). This is a substantial reduction in the number of function evaluations required, since, in comparison, the sampling approach in Ref. [25] needed 4096 function evaluations. We compare the accuracy of our method versus the method in Ref. [25] with the analytical average main effect sensitivity index in Table 1. We see that our method provides more accurate results with considerably less samples required.

To test the method for normal distributions, we consider the same additive function as in Ref. [25]

$$Y = 100X_1 + 4 \exp(X_2) + 350 \sin X_3 \quad (19)$$

with $X_i \sim \mathcal{N}(0, 4)$, $i = 1, 2, 3$. To build up the ANOVA-HDMR, we use up to eighth-order basis functions and ten quadrature points in each dimension. For three input parameters, this results in a total of 1000 function evaluations required to build up the ANOVA-HDMR, a substantial reduction from the 65,536 function evaluations used in the rejection sampling method of [25]. 1000 samples are still a substantial number of samples for a three parameter problem, but we note that this is an inherently complicated function, and that the infinite support of the normal distribution requires us to approximate the function over a wide range. In practice, one could choose the number of samples and the order of the basis functions adaptively by monitoring changes in the approximation; here, we analyze the convergence for different number of samples and basis functions by estimating the error between the surrogate and the actual function. The error is defined as

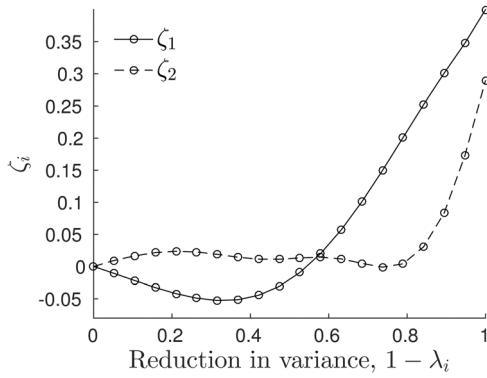


Fig. 5 Variance-based sensitivity index functions for X_1 , X_2 , and X_3 for the Ishigami test function

$$\delta(g, \tilde{g}) = \frac{1}{D} \int [g(\mathbf{x}) - \tilde{g}(\mathbf{x})]^2 f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (20)$$

with $D = \int [g(\mathbf{x}) - g_0]^2 f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}$

Here, g is the truth model, \tilde{g} is the truncated ANOVA-HDMR model, and D is the variance of the truth output quantity of interest Y . The integral in Eq. (20) is evaluated using quasi Monte Carlo sampling with 65,536 samples (the same number of samples as used in Ref. [25]).

Using an ANOVA-HDMR surrogate with up to eighth-order basis functions and nine quadrature points in each dimension (which yields an error of $\mathcal{O}(10^{-2})$ as shown in Fig. 6), the main effect sensitivity indices are found to be $S_1 = 0.27$, $S_2 = 0.31$, and $S_3 = 0.41$. GSA-based factor prioritization would then lead to

the conclusion to focus research on uncertainty reduction in X_3 . The distributional sensitivity analysis results in Fig. 7(a), however, show that it might be more worthwhile to invest in uncertainty reduction in X_2 . This is also indicated by the average main effect sensitivity indices in Fig. 7(b), again found by considering $\lambda_i \sim \mathcal{U}[0, 1]$. Those indicate a different ranking for factor prioritization compared to the GSA results. These results were generated using Algorithm 4.

Note that the results in Fig. 7(a) differ slightly from the results in Ref. [25]. This is because in that work the mean of the updated distribution was constrained to be the same as the mean of the original distribution.

4.2 Aircraft Conceptual Design. The conceptual sizing of a commercial jetliner is considered using the transport aircraft sizing and optimization tool (TASOPT). This tool uses low-order physics-based aircraft sizing models with minimal reliance on empirical and historical data, making it appropriate to use over a wide range of possible designs [41]. In this work, the tool is used in sizing mode, not in the optimization mode. That means that the overall configuration (i.e., sweep angle, material properties, etc.) are fixed, but that other quantities such as the subcomponent weights, overall weight, span, and surface area of wings and tails, are allowed to vary. Two different aircraft configurations are considered and the different impacts of uncertainty on each are investigated.

4.2.1 Problem Setup. Two different aircraft are considered: the Boeing 737-800 and the D8.6 double-bubble conceptual design. For the Boeing 737-800, we consider a sizing mission with a range of 2950 nautical miles, with 180 passengers, at a cruise altitude of 35,000 ft. The D8.6 flies in the same class as the Boeing 737-800, thus we consider the same mission—carrying

Table 1 Comparison of DSA results for the sampling-based method from Ref. [25] (denoted “MC”), our proposed method (denoted “HDMR”), and the exact solution

	X_1 MC	HDMR	Exact	X_2 MC	HDMR	Exact
Main effect sensitivity index	0.4074	0.3999	0.4007	0.2870	0.2886	0.2881
Average main effect sensitivity index	0.06818	0.06861	0.06871	0.03211	0.03162	0.03085

Table 2 Uniform distribution parameters ($\mathcal{U}[a, b]$) of the uncertain design variables considered for the Boeing 737-800 from Refs. [43] and [44].

Parameter	Units	Definition	Lower bound	Upper bound
St_h	—	Turbine cooling Stanton number	0.094	0.096
t_{film}	—	Turbine cooling film effectiveness factor	0.315	0.325
$\eta_{\text{pol},c}$	—	Low pressure compressor efficiency	0.936	0.937
$\eta_{\text{pol},h}$	—	High pressure compressor efficiency	0.903	0.905
$\eta_{\text{pol},l}$	—	Low pressure turbine efficiency	0.875	0.877
$\eta_{\text{pol},h}$	—	High pressure turbine efficiency	0.870	0.872
T_{metal}	K	Turbine metal temperature	1 172	1 272
$(T_{14})_{\text{TO}}$	K	Takeoff turbine inlet total temperature	1 783	1 883
$(T_{14})_{\text{CR}}$	K	Cruise turbine inlet total temperature	1 541.5	1 641.5
OPR_D	—	Overall pressure ratio	24.2	28.2
π_{J_D}	—	Fan pressure ratio	1.609	1.611
$\sigma_{\text{fus,skin}}$	Psi	Maximum allowable fuselage skin pressurization stress	14 250	15 750
$\sigma_{\text{fus,bend}}$	Psi	Maximum allowable fuselage shell bending stress	28 500	31 500
$\sigma_{\text{wt,cap}}$	Psi	Maximum allowable wing and tail spar cap stress	29 500	30 500
$\tau_{\text{wt,web}}$	Psi	Maximum allowable wing and tail spar web shear stress	19 000	21 000
$E_{\text{wt,cap}}$	Psi	Young's modulus wing and tail spar cap	9.50×10^6	10.5×10^6
$\rho_{\text{fus,skin}}$	kg/m ³	Fuselage pressure-skin material density	2 672	2 726
$\rho_{\text{fus,bend}}$	kg/m ³	Fuselage bending-material density	2 672	2 726
$\rho_{\text{wt,cap}}$	kg/m ³	Wing and tail spar cap material density	2 672	2 726
$\rho_{\text{wt,web}}$	kg/m ³	Wing and tail spar web material density	2 672	2 726
$C_{L_{\perp,\text{max}}}$	—	Maximum lift coefficient perpendicular to the chord of the wing	2.2	2.3
C_L	—	Cruise aircraft lift coefficient	0.56714	0.58714

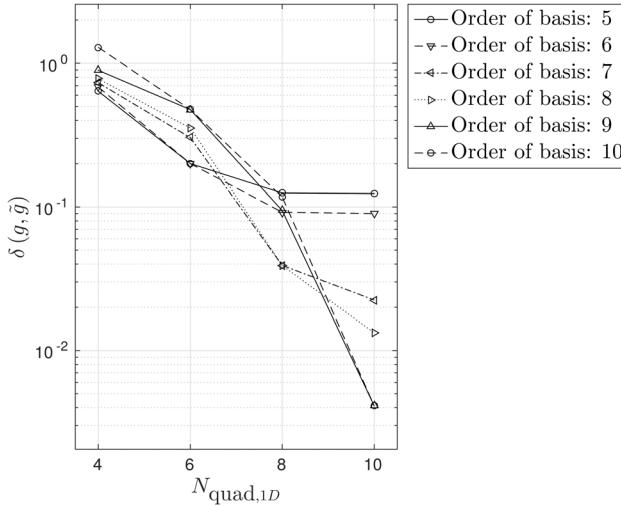


Fig. 6 Convergence of $\delta(g, \tilde{g})$ for the additive function in Eq. (19) for different numbers of quadrature points and basis functions using quasi Monte Carlo simulation with 65,536 samples

180 passengers over a range of 2950 nautical miles. The D8.6 is part of the D8 family developed as part of the NASA N+3 project [42]. The goal of that project was to come up with a design which has, among other attributes, 70% less fuel burn and 75% less landing and take-off NOx exhausts, all relative to the current Boeing 737-800. In order to achieve these ambitious goals, advanced new technologies are employed in the design, many of which are still in development. This implies uncertainties in the expected gain from these developments, and therefore in the performance of the aircraft.

In our study, two factors are considered to be uncertain: overall pressure ratio of the engine, OPR, and the cruise Mach number, M . OPR represents engine performance, allowing for a trade-off between engine technology and mission parameters. For the Boeing 737-800, OPR and M are modeled as uniform random variables, with $OPR \sim \mathcal{U}[24.2, 28.2]$ and $M \sim \mathcal{U}[0.77, 0.79]$, see Refs. [43,44] for details on how those distributions were derived. For the D8.6, OPR and M are modeled as triangular random variables, with $OPR \sim \mathcal{T}[45, 52, 50]$ and $M \sim \mathcal{T}[0.73, 0.75, 0.74]$. These distributions are derived using a combination of historical data and expert opinion, as described in Refs. [44,45]. As the quantity of interest, we consider PFEI, the fuel energy consumption per payload-range.

This case study provides an opportunity to study the influence of nonlinear responses on the DSA results. The Boeing 737-800

flies at a high Mach number (~ 0.78) and we expect there to be a nonlinear effect on the fuel efficiency as we increase Mach number. The D8.6, however, is designed to fly at a lower Mach number (~ 0.74) for higher efficiency and we therefore expect a more linear response.

4.2.2 Sensitivity Analysis. In order to perform the sensitivity analysis, an ANOVA-HDMR surrogate is created for both aircraft. For the Boeing 737-800, nine quadrature points per dimension and up to seventh-order basis functions are used. Therefore, a total of 81 TASOPT function evaluations are required for the Boeing 737-800. The reason we need a fairly large number of samples is because we expect the response of PFEI with respect to Mach number to be nonlinear. For the D8.6, we use a five-point quadrature scheme with up to third-order basis functions. This combination results in a low error with the true model (as also shown in Fig. 10(b)) and ensures that the first-order sensitivity indices are converged. Because the domain is split in four parts to account for the discontinuous derivative in the triangular distribution, we need a total of 100 TASOPT function evaluations to create the surrogate for the D8.6.

A GSA on the Boeing 737-800 model gives the main effect sensitivity indices for OPR as 49% and for M as 51%. For the D8.6, we find that the main effect sensitivity index for OPR is 27% and the main effect sensitivity index for M is 73%. The variance-based index function and the average main effect sensitivity indices for the Boeing 737-800 and D8.6—generated using Algorithm 6—are shown in Figs. 8 and 9, respectively. The main observation in Fig. 8 is that the variance-based index function for M is nonlinear. The global sensitivity indices indicate that OPR and M are equally important in uncertainty reduction of the QoI. However, the DSA sensitivity index function indicates that for partial variance reductions, OPR is more important. In this case, DSA allows a designer to make a more informed decision about where to reduce uncertainty. For the D8.6, the variance-based sensitivity indices are linear functions and therefore the ranking is the same between GSA and DSA.

The nonlinear variance-based sensitivity index function for the Mach number of the Boeing 737-800 is the result of the nonlinear response surface of PFEI as a function of Mach number and OPR, as shown in Fig. 10. The response surface of the D8.6 is much smoother, resulting in more linear variance-based sensitivity index functions. The nonlinear response of PFEI as a function of Mach number can be explained by looking at the configuration of the Boeing 737-800. At the design configuration of the Boeing 737-800—designed for $M=0.78$ —the airplane is designed to have the highest efficiency. This means that sweep angle of the wing is chosen

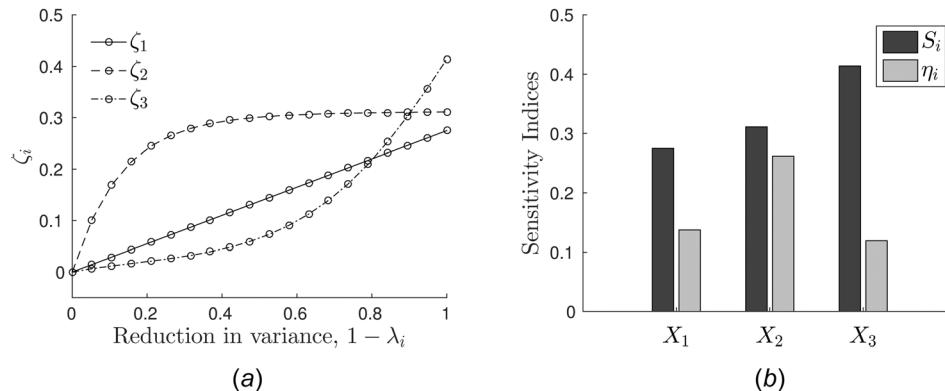


Fig. 7 Distributional sensitivity analysis results for the additive function with normal distributions: (a) variance-based sensitivity index functions for X_1 , X_2 , and X_3 with $\lambda_i = 0.0, 0.05, \dots, 1.0$ and (b) comparison between main-effect sensitivity indices S_i and average main effect sensitivity indices η_i

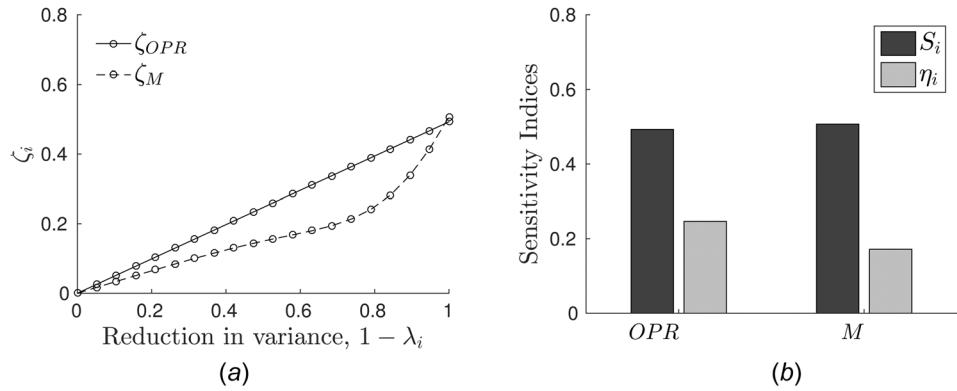


Fig. 8 Distributional sensitivity analysis results for the Boeing 737-800 (a) variance-based sensitivity index function and (b) average main effect sensitivity indices compared to global sensitivity indices

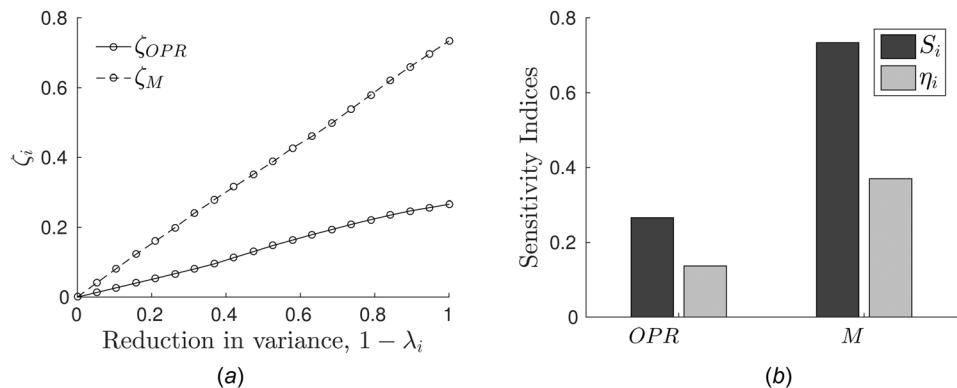


Fig. 9 Distributional sensitivity analysis results for the D8.6: (a) variance-based sensitivity index function and (b) average main effect sensitivity indices compared to global sensitivity indices

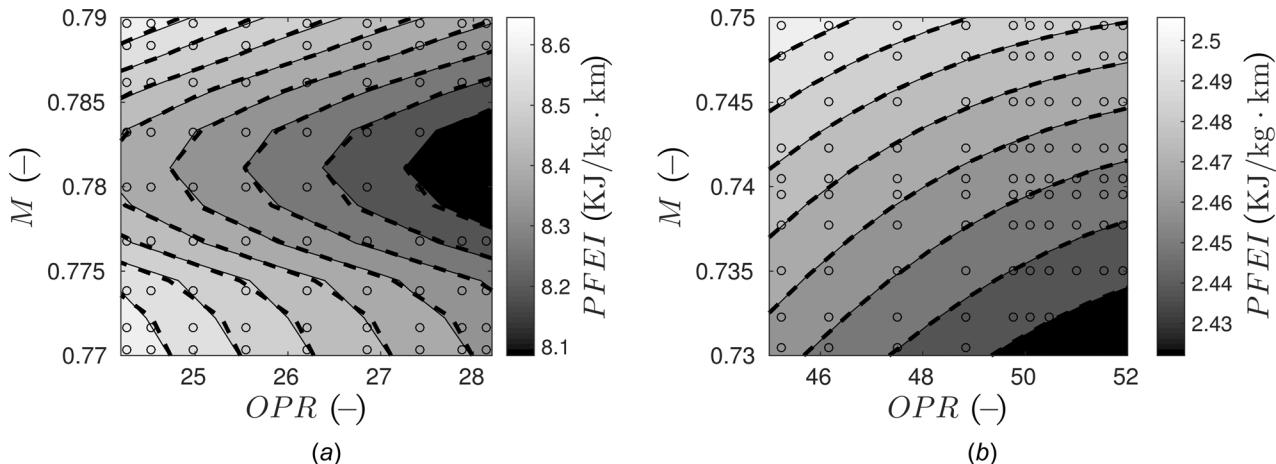


Fig. 10 Contour of PFEI as a function of OPR and M for both aircraft. The shaded contour is the TASOPT "truth" output, while the dashed lines represent the ANOVA-HDMR surrogate and the circles represent the quadrature points: (a) Boeing 737-800 and (b) D8.6.

such that enough lift can be provided while avoiding large drag increases due to high Mach number flow along the chord line of the airfoil. Therefore when one sizes an aircraft for a different Mach number, but does not change the configuration (e.g., C_L or sweep angle), the drag and corresponding fuel consumption are expected to go up for both an increase and decrease in Mach number. For a decrease in Mach number, the

sweep angle is too large for that condition, resulting in more drag in the spanwise direction, increasing fuel burn. For higher Mach numbers, the flow along the chord line of the airfoil is highly transonic, resulting in a drag increase and hence fuel consumption increase. The response of PFEI as a function of Mach number would be closer to linear if the C_L and Mach number were allowed to vary in the aircraft sizing.

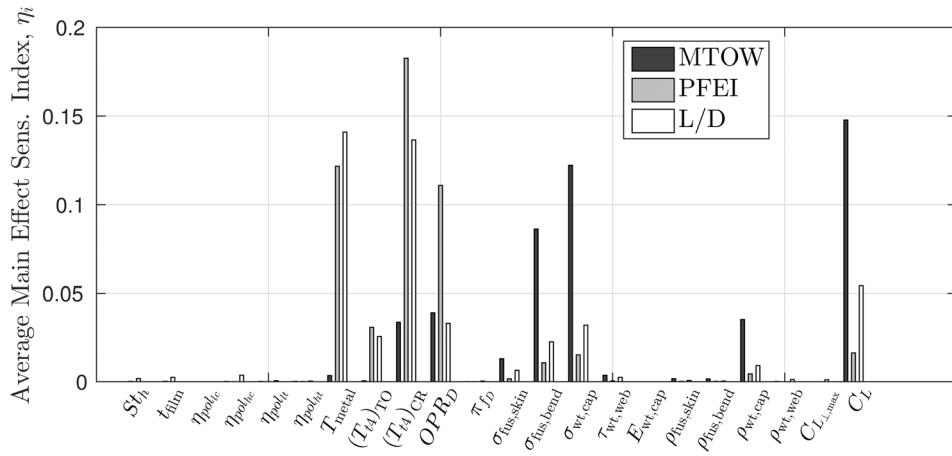


Fig. 11 Average main effect sensitivity indices from DSA of all uncertain design input parameters according to the distributions for the Boeing 737-800 in Refs. [43] and [44]

4.3 High-Dimensional Aircraft Design Problem. Finally, we apply the method to a 22-dimensional aircraft design problem. Specifically, we determine the influence of 22 uncertain design variables on quantities of interest for the conceptual design of the Boeing 737-800 from Sec. 4.2. However, building up a 22-dimensional ANOVA-HDMR using the methods described so far would be computationally intractable, because it would require $(N_{\text{quad}})^{22}$ sample points. Instead, if we neglect higher-order terms in the ANOVA-HDMR using the so-called *cut-HDMR* [30], the number of required samples reduces to $\binom{22}{2}(N_{\text{quad}})^2$. For this problem, we build up a surrogate using $N_{\text{quad}} = 7$ quadrature points in each dimension and up to fifth-order basis functions.

We use a uniform distribution to characterize the uncertainty in all 22 design variables, according to the distribution parameters in Fig. 11, taken from Refs. [43,44]. For reference, the distribution parameters for the 22 design variables are listed in Table 2. In order to find which parameters we should focus uncertainty reduction efforts on, we perform a distributional sensitivity analysis on this problem, using Algorithm 2. Our quantities of interest are MTOW, the maximum take-off weight of the aircraft, PFEI, the fuel energy consumption per payload-range, and L/D, the lift-to-drag ratio of the aircraft during cruise.

The average main effect sensitivity indices from DSA are shown in Fig. 11. The average main effect sensitivity indices provide an informative ranking about the influence of the input parameters on our quantities of interest. We find that propulsion uncertainties ((T_{metal}) , $(T_{14})_{\text{TO}}$, $(T_{14})_{\text{CR}}$, OPR_D), structural uncertainties ($\sigma_{\text{fus},\text{bend}}$, $\sigma_{\text{wt},\text{cap}}$), and an aerodynamic uncertainty (C_L) have the most influence on our three quantities of interest.

We find that the ranking between GSA and DSA does not change for the most important input parameters. However, rather than obtaining just the ranking in the input parameters, we are now able to quantify, using the variance-based sensitivity index functions, the required reduction in uncertainty in each design variable in order for the uncertainty in the quantities of interest to reduce. These variance-based sensitivity index functions could then be combined with cost models for changes in the input distributions and requirements on the uncertainty and cost. This would allow us to come up with new input distributions which meet those requirements.

5 Conclusion

A new formulation of the distributional sensitivity analysis method reduces its computational cost and makes it more widely applicable to probability distributions commonly used in engineering design. The paper formulates and illustrates different

strategies for performing the sampling on these distributions. Application of the method to a case study in aircraft conceptual design demonstrates how the variance-based sensitivity index function provides useful information to the designer on where to target uncertainty reduction efforts. The results also show that when the output of interest depends nonlinearly on the uncertain input parameters, the distributional sensitivity analysis can lead to different conclusions about the relative importance of the inputs, compared to using a standard global sensitivity analysis. This is particularly important when design resources are limited and must be directed as effectively as possible.

A potential area of future work is to extend the methodology in this paper to perform distributional sensitivity analysis for correlated input variables. This would require use of an HDMR framework that incorporates correlated inputs, an example being the *Generalized ANOVA-HDMR* by Rahman [46]. Furthermore, the way in which we sample over new input distributions in the DSA framework would also need to be adapted to handle correlated inputs.

Acknowledgment

This work was supported in part by the NASA LEARN program through Grant No. NNX14AC73A, and by the United States Department of Energy Applied Mathematics Program, Awards DE-FG02-08ER2585 and DE-SC0009297, as part of the DiaMonD Multifaceted Mathematics Integrated Capability Center.

References

- [1] Smith, N., and Mahadevan, S., 2003, "Probabilistic Methods for Aerospace System Conceptual Design," *J. Spacecr. Rockets*, **40**(3), pp. 411–418.
- [2] Antonsson, E. K., and Otto, K. N., 1995, "Imprecision in Engineering Design," *ASME J. Mech. Des.*, **117**(B), pp. 25–32.
- [3] Nikolaidis, E., Mourelatos, Z. P., and Pandey, V., 2011, *Design Decisions Under Uncertainty With Limited Information: Structures and Infrastructures Book Series*, Vol. 7, CRC Press, Boca Raton, FL.
- [4] Helton, J. C., Johnson, J. D., Sallaberry, C. J., and Storlie, C. B., 2006, "Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis," *Reliab. Eng. Syst. Saf.*, **91**(10), pp. 1175–1209.
- [5] Helton, J. C., and Davis, F. J., 2003, "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems," *Reliab. Eng. Syst. Saf.*, **81**(1), pp. 23–69.
- [6] McKay, M. D., Beckman, R. J., and Conover, W. J., 2000, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, **42**(1), pp. 55–61.
- [7] Niederreiter, H., 1992, *Random Number Generation and Quasi-Monte Carlo Method* (SIAMCBMS-NSF Regional Conference Series in Applied Mathematics), Vol. 63, SIAM, Philadelphia, PA.
- [8] Simpson, T. W., Lin, D. K., and Chen, W., 2001, "Sampling Strategies for Computer Experiments: Design and Analysis," *Int. J. Reliab. Appl.*, **2**(3), pp. 209–240.
- [9] Eldred, M. S., Giunta, A. A., and Collis, S. S., 2004, "Second-Order Corrections for Surrogate-Based Optimization With Model Hierarchies," *AIAA* 2004-4457.

[10] Antoulas, A. C., 2005, *Approximation of Large-Scale Dynamical Systems*, Vol. 6, SIAM, Philadelphia, PA.

[11] Gu, X. S., Renaud, J. E., and Penninger, C. L., 2006, "Implicit Uncertainty Propagation for Robust Collaborative Optimization," *ASME J. Mech. Des.*, **128**(4), pp. 1001–1013.

[12] McDonald, M., and Mahadevan, S., 2008, "Uncertainty Quantification and Propagation for Multidisciplinary System Analysis," *AIAA* Paper No. 2008-6038.

[13] Kokkolaras, M., Mourelatos, Z. P., and Papalambros, P. Y., 2006, "Design Optimization of Hierarchically Decomposed Multilevel Systems Under Uncertainty," *ASME J. Mech. Des.*, **128**(2), pp. 503–508.

[14] Mahadevan, S., and Smith, N., 2006, "Efficient First-Order Reliability Analysis of Multidisciplinary Systems," *Int. J. Reliab. Saf.*, **1**(1–2), pp. 137–154.

[15] Morris, M. D., 1991, "Factorial Sampling Plans for Preliminary Computational Experiments," *Technometrics*, **33**(2), pp. 161–174.

[16] Saltelli, A., 2008, *Global Sensitivity Analysis: The Primer*, Wiley, Hoboken, NJ.

[17] Saltelli, A., and Tarantola, S., 2002, "On the Relative Importance of Input Factors in Mathematical Models: Safety Assessment for Nuclear Waste Disposal," *J. Am. Stat. Assoc.*, **97**(459), pp. 702–709.

[18] Homma, T., and Saltelli, A., 1996, "Importance Measures in Global Sensitivity Analysis of Nonlinear Models," *Reliab. Eng. Syst. Saf.*, **52**(1), pp. 1–17.

[19] Saltelli, A., and Marivoet, J., 1990, "Non-Parametric Statistics in Sensitivity Analysis for Model Output: A Comparison of Selected Techniques," *Reliab. Eng. Syst. Saf.*, **28**(2), pp. 229–253.

[20] Borgonovo, E., 2006, "Measuring Uncertainty Importance: Investigation and Comparison of Alternative Approaches," *Risk Anal.*, **26**(5), pp. 1349–1361.

[21] Borgonovo, E., 2007, "A New Uncertainty Importance Measure," *Reliab. Eng. Syst. Saf.*, **92**(6), pp. 771–784.

[22] Chun, M.-H., Han, S.-J., and Tak, N.-I., 2000, "An Uncertainty Importance Measure Using a Distance Metric for the Change in a Cumulative Distribution Function," *Reliab. Eng. Syst. Saf.*, **70**(3), pp. 313–321.

[23] Saltelli, A., Tarantola, S., and Chan, K.-S., 1999, "A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output," *Technometrics*, **41**(1), pp. 39–56.

[24] Oakley, J. E., and O'Hagan, A., 2004, "Probabilistic Sensitivity Analysis of Complex Models: A Bayesian Approach," *J. R. Stat. Soc. Ser. B (Stat. Methodol.)*, **66**(3), pp. 751–769.

[25] Allaire, D. L., and Willcox, K. E., 2012, "A Variance-Based Sensitivity Index Function for Factor Prioritization," *Reliab. Eng. Syst. Saf.*, **107**, pp. 107–114.

[26] Sobol, I. M., 2003, "Theorems and Examples on High Dimensional Model Representation," *Reliab. Eng. Syst. Saf.*, **79**(2), pp. 187–193.

[27] McRae, G. J., Tilden, J. W., and Seinfeld, J. H., 1982, "Global Sensitivity Analysis Computational Implementation of the Fourier Amplitude Sensitivity Test (Fast)," *Comput. Chem. Eng.*, **6**(1), pp. 15–25.

[28] Saltelli, A., and Bolado, R., 1998, "An Alternative Way to Compute Fourier Amplitude Sensitivity Test (Fast)," *Comput. Stat. Data Anal.*, **26**(4), pp. 445–460.

[29] Rabitz, H., Aliş, Ö. F., Shorter, J., and Shim, K., 1999, "Efficient Input–Output Model Representations," *Comput. Phys. Commun.*, **117**(1), pp. 11–20.

[30] Rabitz, H., and Aliş, Ö. F., 1999, "General Foundations of High-Dimensional Model Representations," *J. Math. Chem.*, **25**(2–3), pp. 197–233.

[31] Aliş, Ö. F., and Rabitz, H., 2001, "Efficient Implementation of High Dimensional Model Representations," *J. Math. Chem.*, **29**(2), pp. 127–142.

[32] Smith, R., 2014, *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM, Philadelphia, PA.

[33] Li, G., Wang, S.-W., and Rabitz, H., 2002, "Practical Approaches to Construct RS-HDMR Component Functions," *J. Phys. Chem. A*, **106**(37), pp. 8721–8733.

[34] Wang, S.-W., Georgopoulos, P. G., Li, G., and Rabitz, H., 2003, "Random Sampling-High Dimensional Model Representation (RS-HDMR) With Nonuniformly Distributed Variables: Application to an Integrated Multimedia/Multi-pathway Exposure and Dose Model for Trichloroethylene," *J. Phys. Chem. A*, **107**(23), pp. 4707–4716.

[35] Li, G., and Rabitz, H., 2012, "General Formulation of HDMR Component Functions With Independent and Correlated Variables," *J. Math. Chem.*, **50**(1), pp. 99–130.

[36] Allaire, D. L., 2009, "Uncertainty Assessment of Complex Models With Application to Aviation Environmental Systems," *Ph.D thesis*, Massachusetts Institute of Technology, Cambridge, MA.

[37] Ghanem, R., and Spanos, P., 1990, "Polynomial Chaos in Stochastic Finite Elements," *ASME J. Appl. Mech.*, **57**(1), pp. 197–202.

[38] Ghanem, R. G., and Spanos, P. D., 2003, *Stochastic Finite Elements: A Spectral Approach*, Dover Publications, Mineola, NY.

[39] Sudret, B., 2008, "Global Sensitivity Analysis Using Polynomial Chaos Expansions," *Reliab. Eng. Syst. Saf.*, **93**(7), pp. 964–979.

[40] Ratto, M., Pagano, A., and Young, P. C., 2009, "Non-Parametric Estimation of Conditional Moments for Sensitivity Analysis," *Reliab. Eng. Syst. Saf.*, **94**(2), pp. 237–243.

[41] Drela, M., 2010, "N3 Aircraft Concept Designs and Trade Studies—Appendix," Technical Report No. NASA CR-2010-216794.

[42] Uranga, A., Drela, M., Greitzer, E. M., Titchener, N. A., Lieu, M. K., Siu, N. M., Huang, A. C., Gatlin, G. M., and Hannon, J. A., 2014, "Preliminary Experimental Assessment of the Boundary Layer Ingestion Benefit for the D8 Aircraft," *AIAA* Paper No. 2014-0906.

[43] Opgenoord, M. M. J., 2016, "Uncertainty Budgeting Methods for Conceptual Aircraft Design," *SM thesis*, Massachusetts Institute of Technology, Cambridge, MA.

[44] Amaral, S., 2015, "A Decomposition-Based Approach to Uncertainty Quantification of Multicomponent Systems," *Ph.D thesis*, Massachusetts Institute of Technology, Cambridge, MA.

[45] Ng, L. W.-T., 2013, "Multifidelity Approaches for Design Under Certainty," *Ph.D thesis*, Massachusetts Institute of Technology, Cambridge, MA.

[46] Rahman, S., 2014, "A Generalized ANOVA Dimensional Decomposition for Dependent Probability Measures," *SIAM/ASA J. Uncertainty Quantif.*, **2**(1), pp. 670–697.