AIAA 2016-1440

AIAA SciTech
4-8 January 2016, San Diego, California, USA
18th AIAA Non-Deterministic Approaches Conference

Importance Sampling-based Post-Processing Method for Global Sensitivity Analysis

Daniel M. Sparkman[1]
Universal Technology Corporation, Dayton, Ohio, 45432

Jose E. Garza[2]
University of Texas at San Antonio, San Antonio, Texas, 78249

Harry R. Millwater, Jr.[3]
University of Texas at San Antonio, San Antonio, Texas, 78249

Benjamin P. Smarslok[4]
Air Force Research Laboratory, Aerospace Systems Directorate, Wright-Patterson Air Force Base, Ohio, 45433

A methodology is presented for the calculation of global sensitivity indices from an existing sample of simulations. A sample weighting scheme similar to kernel regression is proposed for estimating the moments of the conditional distributions in the calculation of the sensitivity indices. The weights are developed based on importance sampling, but with the roles of the target distribution and the sampling distribution reversed. The method is demonstrated with benchmark examples from the sensitivity analysis literature.

Nomenclature

$\boldsymbol{x}$	=	vector of random input random variables
$\boldsymbol{x}^{(j)}$	=	j^{th} random input random variable
y	=	output of model $f(\boldsymbol{x})$ evaluated at $\boldsymbol{x}$
pdf	=	probability density function
$p(\boldsymbol{x})$	=	probability density function of variables $\boldsymbol{x}$
$E[\cdot]$	=	expected value operator
$V[\cdot]$	=	variance operator
N	=	sample size (number of observations)
$q(\boldsymbol{x})$	=	sampling probability density function of variables $\boldsymbol{x}$
$t(\boldsymbol{x})$	=	target probability density function of variables $\boldsymbol{x}$
$L(\boldsymbol{x})$	=	likelihood ratio function
$\hat{m}_i$	=	estimated mean of the response conditioned on $\boldsymbol{x}=\boldsymbol{x}_i$
$S^{(j)}$	=	first order sensitivity index of $\boldsymbol{x}^{(j)}$
$T^{(j)}$	=	total effects sensitivity index of $\boldsymbol{x}^{(j)}$

[1] Research Engineer, Universal Technology Corporation, AIAA Member.
[2] PhD, Mechanical Engineering, University of Texas at San Antonio.
[3] Professor and Associate Dean of Research and Graduate Studies, Mechanical Engineering, University of Texas at San Antonio, AIAA Member.
[4] Aerospace Engineer, Aerospace Systems Directorate, AFRL, AIAA Member.

1. Introduction

The design and analysis of physical systems often makes use of a numerical model that predicts the system's response for certain inputs. As part of an effort to better understand the physical system, an analyst will study the numerical model and draw conclusions. Real physical systems are subject to uncertainty in their inputs and will have uncertainty in the system's response. Uncertainty analysis and sensitivity analysis are considered an essential component of analysis of complicated models [7]. In an uncertainty quantification analysis, the uncertainty of the model's response is estimated by propagating the uncertainty in the inputs through the model. A sensitivity analysis identifies the significant sources of uncertainty and quantifies their contributions. Sensitivity analyses can aid in decision-making for the physical system.

Sensitivity analyses are typically classified into one of two categories, local or global. Local sensitivity analyses examine how the model behaves in a region near a point in the space of the inputs, often using the partial derivatives of the response with respect to the inputs. Typically the derivatives are estimated with finite differencing from perturbations of the inputs one-factor-at-a-time (OAT) about some expansion point. The number of model evaluations required is on the order of $O(k)$, where k is the number of inputs. Some examples include the differential importance measure [40]. These metrics can be useful for determining the driving factors of the response uncertainty in the local region about the expansion point. The local sensitivities will agree with the global sensitivities only under conditions that the response is smooth, continuous, and approximately linear with the inputs and the distribution of the inputs is unimodal, symmetric, and centered about the expansion point. However, for nonlinear responses and general input probability distributions the local sensitivities are not necessarily useful over the whole space of the inputs.

Global sensitivity analysis (GSA) on the other hand is "model free" and thus is applicable to response models of any form and input variables with any distribution type. Also they are global as they account for the influence of the full distribution of the input variables [1]. The measures computed in GSA are called sensitivity indices. The most popular indices are variance-based. They describe the fraction of the total variance of the response that is due to subsets of the input variables [1]. GSA methods have a significantly larger computational cost than local sensitivity methods. The simplest method to implement involves a Monte Carlo Simulation (MCS) nested within another MCS. For each realization of the variable of interest in the 'outer loop' MCS, an 'inner loop' MCS for all of the other variables is performed. The total cost is on the order of $O(2kN^2)$ where N is the size of a single MCS. Practical applications of GSA use more efficient calculation methods. The popular method of [1, 15] requires $O(N(k+2))$ model evaluations. The Fourier Amplitude Sensitivity Test (FAST) method [22] has a cost on the order of $O(k^2)$. These methods are very efficient but require specially constructed experiment designs *a priori*. Other examples using special designs involve the use of Latin Hypercube Sampling, Quasi-Monte Carlo Sampling, or importance sampling [6, 9] to reduce the computational cost of computing GSA indices. It is noted that in [9] a simple function is used to approximate the model and weighted with importance sampling. In [6], importance sampling is used in the traditional manner to reduce the variance of the probability of failure estimates for the calculation of reliability sensitivity indices.

There are a few methods for estimation of the variance-based global sensitivity indices from an existing sample of model evaluations obtained in a previous uncertainty quantification analysis such as MCS. In [8] the sample is sorted in a way to approximate the method in FAST. In [10] a state dependent parameter approach was proposed to estimate the conditional moments for sensitivity analysis. The method allowed for the estimation of higher conditional moments such as the conditional variance and conditional skewness. Methods using kernel regression [17] have been applied successfully. The binning approaches in [1, 18, 35] can be viewed as a special case of kernel regression.

Besides the variance-based indices, other global sensitivity metrics have been proposed. Some indices have been proposed based on a disparity metric or 'distance' between conditional and unconditional distributions of the response. In [11], influential input variables are identified using different statistics of the Kolmogorov-Smirnov statistic between the response unconditional cumulative density function (cdf) and the conditional cdf. In [13], kernel estimates of the conditional and unconditional response probability density functions (pdfs) are used to estimate the delta index in [40]. In the influential variable identification setting, [12] examined many non-parametric regression methods and suggest using the Predicted Residual Sum of Squares (PRESS) statistic to identify the subset of important variables in a stepwise fashion. The elementary effects method of [20] and further developed in [21] amount to estimating statistics of partial derivatives computed from finite differences.

Other sensitivity methods commonly employed include scatter plots, correlation coefficients, regression coefficients, the score function method, and localized sensitivities [4, 7]. Scatter plots are a visual tool for identifying relationships between the inputs and output, but are not quantitative. Correlation coefficients and regression coefficients are simple to compute, but may not provide useful metrics for nonlinear models. Score function methods compute the derivatives of probabilistic responses such as the mean, standard deviation, and event probabilities with respect to parameters of the distribution of the input variables. A recently proposed localized sensitivity method in [3] uses a special score function to obtain probabilistic sensitivities with respect to shifts of probability density in local regions of the inputs. The approach in [14] divides the input and output space into subdomains and a regional global sensitivity is computed for each section. Common to all of these methods is that the sensitivity analysis is performed in a post-processing step using a single existing sample of realizations from an uncertainty quantification analysis.

There are a large number of methods and measures available in the literature for the sensitivity analysis of computer models. Some of the factors to consider when selecting an approach are the objective of the sensitivity analysis [1] (i.e. model corroboration, factor prioritization, model simplification, etc.), the desired quality of the sensitivity results, the computational cost of the sensitivity analysis, and the difficulty of implementation of the sensitivity analysis procedures.

This work focuses on variance-based sensitivity indices in the context of GSA, and proposes a simple method to compute the sensitivities based on importance sampling using an existing sample of simulations; hence the computational cost is significantly less than traditional approaches. Section 2 recalls the formal definitions of the variance-based sensitivity indices. Section 3 provides necessary context of kernel regression and importance sampling, then presents the proposed importance sampling kernel (ISK) method for computing the sensitivity indices, and discusses implementation of the method. Section 4 illustrates the method with two classical benchmark examples from the literature and a new problem that highlights certain features of the method.

2. Global Sensitivity Indices

Consider the output Y of a system is dependent on the k input variables $\boldsymbol{X} = X^{(1)}, X^{(2)}, \ldots, X^{(k)}$ and is predicted by a mathematical model $Y = f(\boldsymbol{X})$. In the following, superscripts in parenthesis refer to a specific input variable or set of inputs. When the inputs are subject to uncertainty according to the joint distribution $p_x(\boldsymbol{x})$, the response will also be uncertain. Global sensitivity analysis focuses on the calculation of sensitivity indices which are developed from the law of total variance. The first order effects sensitivity index, $S^{(j)}$, describes the fraction of the variance of the output due only to the variation of $X^{(j)}$, sometimes called the main effect contribution. Similar to the notation in [1], the superscripts (j) and $(\sim j)$ denote the set of inputs $X^{(j)}$ and all of the inputs except $X^{(j)}$, respectively. The first order effects sensitivity index is defined in Eq. (1).

$$S^{(j)} = \frac{V[E(Y \mid X^{(j)})]}{V[Y]} \tag{1}$$

The total effect index, $T^{(j)}$, is a measure of the total contribution of the variation of the random variable $X^{(j)}$ to the variation of the output. It includes interactions with other random variables and is determined by subtracting the portion of the variance due to the other variables [1]. The total effects index is defined in Eq. (2).

$$T^{(j)} = 1 - \frac{V[E(Y \mid \boldsymbol{X}^{(\sim j)})]}{V[Y]} \tag{2}$$

If the total effect index for a variable is large relative to the first order index, then interactions with the other variables significantly affect the variance of the response. Conversely, if the total effect index for a variable is zero, then the variance of that variable is of no importance in the problem. The first order indices and total effect indices can be compared between inputs to identify their relative importance. Larger values of $S^{(j)}$ and $T^{(j)}$ denote greater relative importance.

3. Importance Sampling-based Kernel Functions for GSA

At the core of the computation of the global sensitivity indices is the estimation of $E(Y \mid X^{(j)})$, the expected value of the response Y conditioned on the input $X^{(j)} = x^{(j)}{}_0$. The conditional expected value is defined as $E[Y \mid X^{(j)}] = \int y\, p_{Y,X^{(\sim j)}|X^{(j)}}(y, \boldsymbol{x}^{(\sim j)} \mid x^{(j)}) dy dx^{(\sim j)}$. Since the model $y = f(\boldsymbol{x})$ is a deterministic function of the inputs, the integral simplifies to $E[Y \mid X^{(j)}] = \int f(\boldsymbol{x}) p_{X^{(\sim j)}|X^{(j)}}(\boldsymbol{x}^{(\sim j)} \mid x^{j}) dx^{(\sim j)}$, where $p_{X^{(\sim j)}|X^{(j)}}(\boldsymbol{x}^{(\sim j)} \mid x^{(j)})$ is the conditional pdf. From $p_{X^{(\sim j)}|X^{(j)}}(\boldsymbol{x}^{(\sim j)} \mid x^{(j)}) = p_X(\boldsymbol{x}) \big/ p_{X^{(j)}}(x^{(j)})$, where $p_X(\boldsymbol{x})$ is the joint density and $p_{X^{(j)}}(x^{(j)})$ is the marginal density of $X^{(j)}$, the modified expression for the conditional expectation is given in Eq. (3).

$$E[Y \mid X^{(j)}] = \int f(\boldsymbol{x}) \frac{p_X(\boldsymbol{x})}{p_{X^{(j)}}(x^{(j)})} d\boldsymbol{x}^{(\sim j)} \tag{3}$$

In the context of working with an existing sample of size N from an unconditional MCS, the available data are $(\boldsymbol{x}_i, y_i)$; $i = 1, \ldots, N$. The $\boldsymbol{x}_i$ are N i.i.d. random draws from $p_x(\boldsymbol{x})$ and $y_i = f(\boldsymbol{x}_i)$ are the corresponding model outputs. The unconditional realizations of $X^{(j)}$ will not be equal to $x^{(j)}{}_0$ and Eq. (3) cannot be directly estimated. To compute $E(Y \mid X^{(j)})$, additional evaluations of the model would be needed with $X^{(j)}$ held fixed (such as in nested MCS). Kernel regression is a method that uses the values of the response *in the neighborhood* of a realization to estimate an approximate local average without requiring additional evaluations of the model. This work proposes a modified form of kernel regression, in which the kernel function is developed from the likelihood ratio in importance sampling for better accuracy. Kernel regression and importance sampling are both briefly reviewed in the following before introducing the proposed method.

3.1 Background: Kernel Regression

Kernel regression is a nonparametric technique that uses values of the output in the neighborhood of a sample point to estimate the conditional mean. The method was introduced in [23] and [44]. The integral in Eq. (3) is estimated by replacing the joint and marginal pdfs with their kernel density estimates. After some simplification, the kernel regression estimator takes the form in Eq. (4),

$$\hat{E}(Y \mid X = x) = \sum_{i=1}^{N} y_i w_i \tag{4}$$

where the weights, $w_i = K\left(\frac{x_i - x}{h}\right) \Big/ \sum_{j=1}^{N} K\left(\frac{x_j - x}{h}\right)$, are computed from the kernel function $K(u)$ and sum to one. The kernel function determines how much weight is given to samples in the neighborhood of the kernel's center. Kernel functions often used in practice include the uniform kernel, Epanechnikov [24], and Gaussian. These kernels are all symmetric about $u = 0$. The uniform kernel gives equal weight to each point within the kernel's width, the Epanechnikov kernel assigns weight that decays quadratically with distance from the center, and the Gaussian kernel decays as a squared exponential. The choice of kernel function impacts the estimated mean but is less important than the selection of a proper bandwidth, h, particularly as the sample size N increases. The bandwidth parameter determines the size of the local neighborhood over which the averaging occurs which affects the smoothness of the regression. Consistent estimators require that h decreases as N increases. Optimal values for h in the mean integrated squared error (MISE) sense depend on the scatter in the conditioning input variable, the sample size, and the form of the kernel function. Methods for calculating a suitable bandwidth [5] include the normal rule of thumb, plug-in estimators, and cross validation [30].

Kernels most commonly used in practice are symmetric. Usually the same kernel function and bandwidth are used over the entire input space. This is referred to as a local constant kernel. Optimal asymmetric kernels [42] and variable bandwidth kernels [31, 32] have been developed. The nearest neighbor-based kernels [41] are akin to the variable bandwidth kernels, and both methods attempt to adapt to the local density. The works in [31] and [41] focused on kernel density estimation, but are easily extended for kernel regression. In [33], variable bandwidth kernel regression was combined with a local linear smoother for an improved estimator.

3.2 Background: Importance Sampling

With traditional MCS a sample is generated according to $p_x(\boldsymbol{x})$, the model response $y = f(\boldsymbol{x})$ is evaluated for each sample, and then the results are post-processed to estimate the statistics of interest. For the remainder of this work the subscript $\boldsymbol{x}$ of the pdf is dropped for clarity of presentation. Importance sampling (IS) can place more samples in the 'important' or event region that contributes most to the integral but requires an *a priori* specification of a distribution of the input variables $q(\boldsymbol{x})$, called the sampling distribution, that is different than the actual distribution $p(\boldsymbol{x})$, called the target distribution. In typical applications, $q(\boldsymbol{x})$ is chosen from the same parametric family of distributions as $p(\boldsymbol{x})$, but with the parameters adjusted in such a way (such as translation of the mean, or scaling the variance) to place more samples in the event region.

The mean of a function is estimated with importance sampling by multiplying the integrand in the definition of the expected value, $E(Y) = \int f(\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}$, by $q(\boldsymbol{x})/q(\boldsymbol{x})$ and rearranging, yielding Eq. (5).

$$E(Y) = \int f(\boldsymbol{x})\frac{p(\boldsymbol{x})}{q(\boldsymbol{x})}q(\boldsymbol{x})d\boldsymbol{x} \tag{5}$$

As shown in Eq. (6), the ratio of the target pdf over the sampling pdf is sometimes called the likelihood ratio and is given the label $L(\boldsymbol{x})$. It serves as a weighting function to 'correct' for the fact that the sample was drawn from the importance sampling distribution instead of the actual distribution.

$$L(\boldsymbol{x}) = \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} \tag{6}$$

Substitution into the MCS discrete estimator gives Eq. (7), the formulation of importance sampling most often encountered in the literature [2, 35]. Note $y_i = f(\boldsymbol{x}_i)$ and $L_i = L(\boldsymbol{x}_i)$.

$$\hat{E}(Y) = \frac{1}{N}\sum_{i=1}^{N} y_i L_i \tag{7}$$

In Eq. (8), the leading 1/*N* is moved inside the summand and $w_i = \frac{1}{N}L_i$ is introduced to develop an expression equivalent to Eq. (7) in a form similar to the kernel regression estimator in Eq. (4). It is emphasized that the form in Eq. (8) estimates the unconditional mean. In the next subsection it will be modified to estimate a conditional mean.

$$\hat{E}(Y) = \sum_{i=1}^{N} y_i w_i \tag{8}$$

Importance sampling is typically employed to reduce the sampling variance of the estimate, thus smaller sample sizes are required than with traditional MCS. The efficiency gains may be significant, with a theoretical limit of only one sample and corresponding evaluation of $y = f(\boldsymbol{x})$ necessary. Appropriately selecting $q(\boldsymbol{x})$ is the central 'art' of applying importance sampling. Experience and prior knowledge of the problem under study usually guide this step.

The sample weights w_i are calculated after the evaluation of the responses. As a post-processing step, the results for a sample drawn from $q(\boldsymbol{x})$ are in essence corrected by the weights to provide estimated statistics for $p(\boldsymbol{x})$ without running a new MCS. A more general interpretation of importance sampling was introduced in [2] in which the roles of $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are reversed. In this representation, $q(\boldsymbol{x})$ takes on the role of the actual distribution and

$p(\boldsymbol{x})$ takes on the role of a 'different' distribution. For a single sample from $q(\boldsymbol{x})$, the resulting data $(\boldsymbol{x}_i, y_i)$; $i = 1,\ldots,N$ are used to estimate statistics for different distributions $p(\boldsymbol{x})$ without running a new MCS. The method in [2] enables investigation of what-if scenarios, studying the effect of changing the input variables' distribution using a single existing MCS.

3.3 Proposed Importance Sampling-based Kernel (ISK) Estimator of *E*(*Y*|*X*=*x*$_i$)

In this work, the generalization of importance sampling in [2] is applied to estimate the conditional expectation using a sample from the joint distribution in which all inputs are allowed to vary. This is accomplished by selecting a target density that is tightly distributed around the value of the input variables the expectation is conditioned on. Global sensitivity indices are then computed. The method is intended to be applied to problems in which a traditional MCS analysis has already been performed, and the sensitivities are to be estimated in a post-processing step.

The estimate for the conditional expectation based on importance sampling in integral form is

$$E(Y \mid \boldsymbol{X} = \boldsymbol{x}_i) \approx \int f(\boldsymbol{x}) \frac{t(\boldsymbol{x})}{q(\boldsymbol{x})} q(\boldsymbol{x}) d\boldsymbol{x} . \tag{9}$$

For $E(Y \mid \boldsymbol{X} = \boldsymbol{x}_i) \equiv m_i$, the MCS discrete form of the proposed estimator of the conditional expectation is given in Eq. (10).

$$\hat{m}_i = \sum_{b=1}^{N} y_b w_b \tag{10}$$

For each realization $\boldsymbol{x}_b$ in the sample, $b = 1,\ldots,N$, the likelihood ratio is computed with Eq. (11).

$$L_b = \frac{t(\boldsymbol{x}_b)}{q(\boldsymbol{x}_b)} \tag{11}$$

The sample weights are obtained with Eq. (12) by normalizing the likelihood ratio to sum to one as in [27]. This makes the estimator in Eq. (10) more robust and brings it into a form with a close resemblance to the kernel regression estimator. Note that the sample indices *b* and *c* are introduced to avoid confusion with the conditioning sample index *i* and the random variable index *j*.

$$w_b = L_b \bigg/ \sum_{c=1}^{N} L_c \tag{12}$$

The proposed estimator in Eq. (10) will be referred to as an importance sampling kernel (ISK) estimator since the likelihood ratio from importance sampling is used as the kernel function in kernel regression.

Additional notation for the target density $t(\boldsymbol{x})$ is introduced in Eq. (11) to maintain the distinction from the actual distribution of the inputs $p(\boldsymbol{x})$. While a wide variety of models are valid pdfs for $t(\boldsymbol{x})$, in this work a Gaussian model is used because it is symmetric, unimodal, and samples far from the center have low density. To condition on $X^{(j)} = x_i^{(j)}$, the mean of $t(\boldsymbol{x})$ is set equal to $x_i^{(j)}$. The standard deviation for $t(\boldsymbol{x})$ takes on the same role as the bandwidth parameter in kernel regression and will be referred to as the bandwidth interchangeably in the following. In this work the normal reference rule-of-thumb[5] is applied to estimate the bandwidth *h* in Eq. (13). It is a function of the marginal standard deviation of $X^{(j)}$ and the sample size N.

$$h = 1.06 \sigma_{X^{(j)}} N^{-1/5} \tag{13}$$

Collecting the parameters $\mu = x_i^{(j)}, \sigma = h$, gives the target pdf in Eq. (14) to estimate $E(Y \mid \boldsymbol{X} = \boldsymbol{x}_i)$.

$$t(\boldsymbol{x}) = t(\boldsymbol{x}; \mu, \sigma) \tag{14}$$

Figure 1 illustrates the methodology using a scatter plot with a simple example. The example is for variable 4 in section 4.1, but at this point only general characteristics of the method will be discussed with specifics reserved for section 4.1. The sample of inputs and corresponding output (x_i, y_i) from MCS with N=1000 are plotted as black points. The samples were drawn from the actual pdf $p^{(j)}(x^{(j)})$ and thus the sampling pdf $q(x) = p^{(j)}(x^{(j)})$. Visually the output increases as the input increases, but with scatter about the trend. This example shows how the method works to estimate the conditional expectation of the output at one of the sample points $x_{i=39}$=3.56. This sample point is identified as a black star beneath the trend line. The target pdf $t(x)$ is centered at $x_{i=39}$=3.56 with a bandwidth computed based on the sample with Eq. (13). It is plotted as a magenta dashed line, and has a tighter distribution than $p^{(j)}(x^{(j)})$, the marginal pdf of the input given by the black curve. Note that the density curves are scaled vertically to have equal maxima for easy comparison. The corresponding weights from Eq. (12) are scaled and plotted as the solid red curve labeled $w\,(\mathrm{ISK})$. The weights are larger for samples near $x_{i=39}$ than for samples farther away. The estimated conditional expectation for each sample point is plotted as red dots, labeled $m\,(\mathrm{ISK})$, and they follow the trend of the data. For comparison, the analytically computed exact value of the conditional expectation is plotted as the solid blue line and close agreement is observed between the predicted estimate and the exact values.

Additionally, the conditional expectation was computed using traditional kernel regression for comparison. A local constant kernel equal to the target density $t(x)$ for each x_i was used to illustrate the difference between the ISK estimator and traditional methods. The scaled weights appear as $w\,(\mathrm{GK})$ and the computed conditional expectation is plotted as magenta dots labeled $m\,(\mathrm{GK})$ denoting estimation with a Gaussian Kernel. The difference between the two estimates is apparent in the tails of the input. These are regions where the sampling density is lower. Further examination of the weights $w\,(\mathrm{ISK})$ and $w\,(\mathrm{GK})$ reveals the ISK estimator amounts to an asymmetric and adaptive kernel function for kernel regression. At $x_{i=39} = 3.56$ the sampling pdf $q(x)$ has a negative slope. The samples that fall just on either side of $x_{i=39} = 3.56$ are weighted equally by the traditional kernel $w\,(\mathrm{GK})$ but unequally by the ISK method $w\,(\mathrm{ISK})$. Other works have noted the non-uniformity of sample density can bias the estimated local average for non-adaptive kernel functions [12, 34, 45]. However, the ISK directly accounts for the fact that more samples are available in areas of higher sampling density which results in improved accuracy.

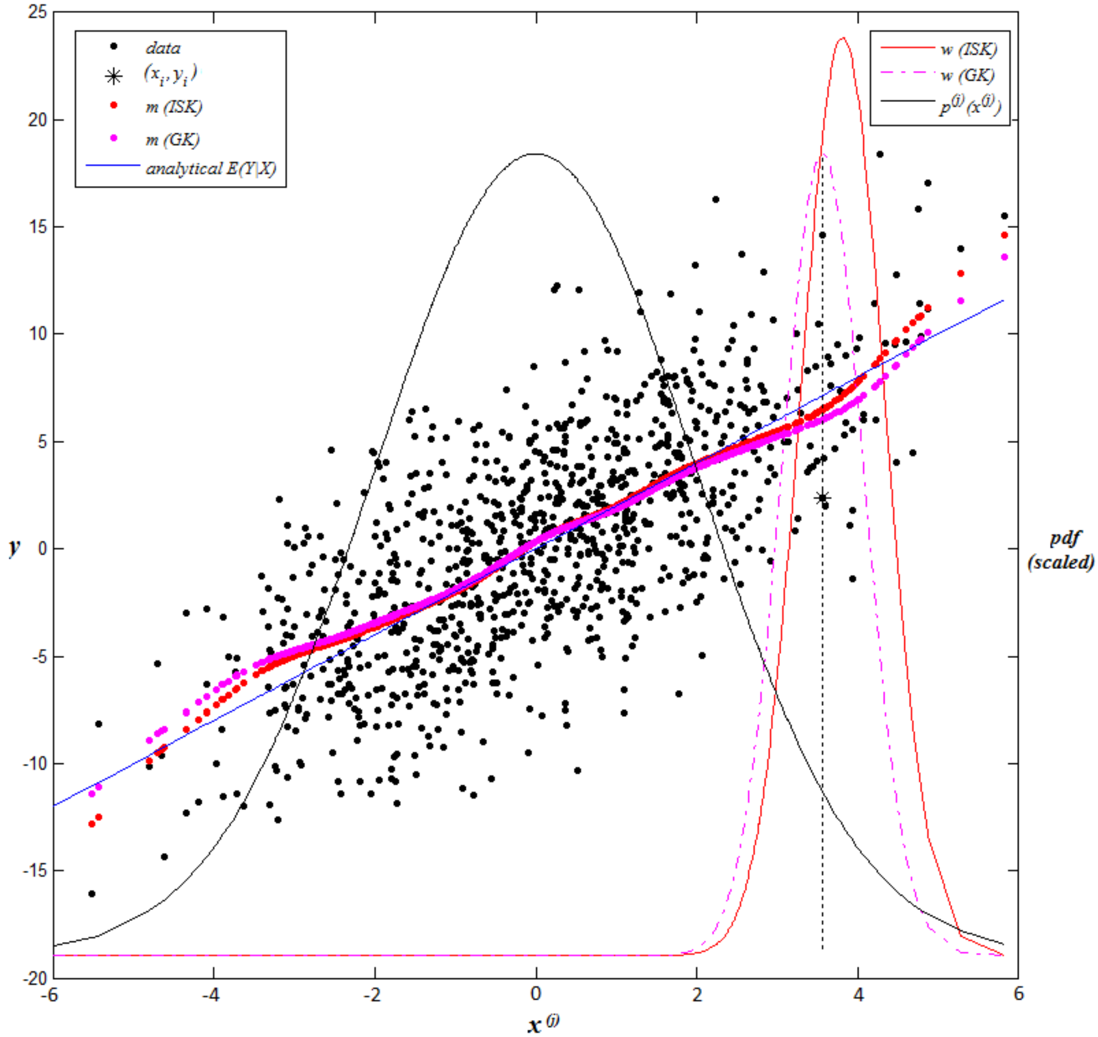

Figure 1. Importance Sampling-based Conditional Mean Estimation

Using Eq. (10), the conditional expectation is estimated at each sample point. It is then straightforward to compute the variance of the conditional expectation and calculate the global sensitivity indices with Eq. (1) and Eq. (2).

3.4 Implementation of the Method

This section gives details pertinent to implementing the ISK method in computer codes. The method was developed for the case that the input variables are independent. For independent variables, the joint density is simply the product of each variable's marginal density, $p(\boldsymbol{x}) = p^{(1)}(x^{(1)}) \cdot p^{(2)}(x^{(2)}) \cdot \ldots \cdot p^{(k)}(x^{(k)})$ and thus $q(\boldsymbol{x}) = q^{(1)}(x^{(1)}) \cdot q^{(2)}(x^{(2)}) \cdot \ldots \cdot q^{(k)}(x^{(k)})$. The joint target density may also be developed as a product of marginal target densities, $t(\boldsymbol{x}) = t^{(1)}(x^{(1)}) \cdot t^{(2)}(x^{(2)}) \cdot \ldots \cdot t^{(k)}(x^{(k)})$. Applying the method to estimate the first order effects sensitivity index amounts to conditioning on the input variable of interest. The total order effects index is estimated by conditioning on all other inputs except the variable of interest. The general formulation in Eqs. (10)-(14) is applicable to both first order and total effects indices, but the computational effort involved may be reduced using simplifications based on the independence of the inputs. These steps are presented in the following subsections.

3.4.1 First Order Effects Sensitivity Index

To approximate conditioning on a single variable taking a specific value, $X^{(j)} = x_i^{(j)}$, only the j^{th} variable's marginal target density is changed from the actual density. The bandwidth *h* used in the target density function is computed using the sample standard deviation of $X^{(j)}$ in Eq. (13). The other variables' target marginal densities remain equal to their actual densities to form the joint target density in Eq. (15).

$$t(\boldsymbol{x}) = p^{(1)}(x^{(1)}) \cdot p^{(2)}(x^{(2)}) \cdot \ldots \cdot t^{(j)}(x^{(j)}) \cdot \ldots \cdot p^{(k)}(x^{(k)}) \tag{15}$$

When conditioning on a single variable, the likelihood ratio of Eq. (11) simplifies as the remaining marginal densities divide out, resulting in the ratio of the target density for just $X^{(j)}$ divided by the sampling density of just $X^{(j)}$ in Eq. (16).

$$L_b{}^{(j)} = \frac{t^{(j)}(x_b{}^{(j)})}{q^{(j)}(x_b{}^{(j)})} \tag{16}$$

The weights are computed with Eq. (12) and then Eq. (10) is used to obtain the estimated conditional mean $\hat{m}_i$. This process is repeated to obtain $\hat{m}_i$ for $i = 1, \ldots, N$. The partial variance $V[E(Y \mid X^{(j)})] \equiv V^{(j)}$ due to variable j is estimated with the variance of $\hat{m}_i$.

$$V^{(j)} = V[\hat{m}_i] \tag{17}$$

The first order sensitivity indices are computed with Eq. (18) noting that $V_Y = V[Y]$ is the total variance of the response.

$$S^{(j)} = \frac{V^{(j)}}{V_Y} \tag{18}$$

Table 1 summarizes the steps to compute all of the first order effects sensitivity indices.

Table 1. Methodology Steps for First Order Effects Sensitivity Indices

1. Compute output total variance V_Y
2. for variable $X^{(j)}$
    - 2.1. Compute bandwidth parameter h (Eq. (13))
    - 2.2. for realization X_i: $i=1,\ldots,N$
        - 2.2.1. for realization X_b: $b=1,\ldots,N$
            - 2.2.1.1. Compute likelihood ratios $L_b{}^{(j)}$ (Eq. (16))
            - 2.2.1.2. Compute weights w_b (Eq. (12))
        - 2.2.2. Compute conditional expected value $\hat{m}_i$ (Eq. (10))
    - 2.3. Compute output partial variance $V^{(j)}$ (Eq. (17))
    - 2.4. Compute global sensitivity index $S^{(j)}$ (Eq. (18))

3.4.2 Total Order Effects Sensitivity Index

To approximate conditioning on all other variables except $X^{(j)}$ taking specific values, $\boldsymbol{X}^{(\sim j)} = \boldsymbol{x}_i{}^{(\sim j)}$, all the other variables' marginal target densities are changed from their actual densities. Only the target marginal density of

$X^{(j)}$ remains equal to its actual density. The bandwidth used for each other variable is computed separately with Eq. (13). The joint target density is then formed in Eq. (19).

$$t(\boldsymbol{x}) = t^{(1)}(x^{(1)}) \cdot t^{(2)}(x^{(2)}) \cdot \ldots \cdot p^{(j)}(x^{(j)}) \cdot \ldots \cdot t^{(k)}(x^{(k)}) \tag{19}$$

Conditioning on all of the other variables except $X^{(j)}$ does not simplify as significantly as for conditioning on a single variable. The likelihood ratio is given here without simplifying to retain a compact expression.

$$L_b^{(\sim j)} = \frac{t(\boldsymbol{x}_b)}{q(\boldsymbol{x}_b)} \tag{20}$$

The results from Eq. (20) are substituted into Eq. (12) to obtain the sample weights. The conditional mean is estimated for each sample with Eq. (10). The partial variance $V[E(Y \mid X^{(j)})] \equiv V^{(j)}$ due to all of the variables except $X^{(j)}$ is calculated with Eq. (21).

$$V^{(\sim j)} = V[\hat{m}_i] \tag{21}$$

The total effects sensitivity indices are calculated with Eq. (22).

$$T^{(j)} = 1 - \frac{V^{(\sim j)}}{V_Y} \tag{22}$$

The steps to estimate the total effects sensitivities are summarized in Table 2. The primary difference from the steps for the first order effects indices (section 3.4.1) is that the bandwidth and marginal target density must be computed for *k*-1 variables just to compute the index for a single input.

Table 2. Methodology Steps for Total Order Effects Indices

1. Compute response total variance V_Y
2. for variable $X^{(j)}$
   - 2.1. Compute bandwidth parameter h for each variable $\boldsymbol{X}^{(\sim j)}$
   - 2.2. for realization X_i: $i=1,\ldots,N$
     - 2.2.1. for realization X_b: $b=1,\ldots,N$
       - 2.2.1.1. Compute likelihood ratios $L_b^{(\sim j)}$ - Eq. (20)
       - 2.2.1.2. Compute weights w_b - Eq. (12)
     - 2.2.2. Compute conditional expected value $\hat{m}_i$ - Eq. (10)
   - 2.3. Compute response partial variance $V^{(\sim j)}$ - Eq. (21)
   - 2.4. Compute global sensitivity index $T^{(j)}$ - Eq. (22)

3.5 Relation to Other Nonparametric Approaches to Global Sensitivity Analysis

Nonparametric methods have been applied to the estimation of the global sensitivities in previous works. In [17], asymptotic results for sensitivities using kernel regression and bandwidths estimated from cross validation are given. In [1] a binning approach is discussed, in which the space of the inputs is partitioned into discrete bins, the mean is computed within each bin, and then the variance is computed across all bins. The input space is partitioned in an adaptive manner in [18]. In another work [35], several partitioning schemes are examined. The binning approaches and quantile approach can both be interpreted as a special case of kernel methods. In the binning approach, samples are in essence given a weight of 0 or 1, depending on whether they fall within a given bin or not. Within a bin, all of the samples are weighted equally. The binning weights of 0 or 1 can also be achieved with an indicator/uniform-type of kernel function. Binning approaches suffer from bin edge effects and convergence can be slow.

Kernel methods are more general than binning methods. The kernel function can be selected such that it decays smoothly which removes the bin edge effects and improves the rate of convergence. The matter of selecting an appropriate bin spacing is replaced with selecting an appropriate bandwidth for the kernel function. A mixture of the two approaches can be found in [26], where kernel density estimates and bins of the input space are used to estimate a density-based sensitivity index.

Other works have applied the generalization of importance sampling introduced by [2] to sensitivity analysis. The recent work in [19] discusses this under the name 'extended MCS' in the context of the fraction of the variance of the failure probability due to variation of the input variable's distribution parameters. Methods using importance sampling and rejection sampling are compared. Other work [39] has identified rejection sampling as a special case of importance sampling. The practical sensitivity method in [29] applies rejection sampling in the manner in [2] to estimate a variance-based sensitivity. It describes the average reduction of variance if an input is not held fixed but has a new distribution. Importance sampling could be used in place of rejection sampling to estimate the indices in [29]. The indices proposed in [36] describe the average reduction in variance if an input is held fixed after its range is reduced. The method [36, 37] is essentially importance sampling where the target pdfs have reduced range based on quantiles of the sampling pdfs. The ISK method proposed here is different because it uses importance sampling to estimate the conditional mean rather than the variance of the conditional mean.

The nonparametric regression approaches for sensitivity analysis in [12, 25] differ from the proposed method here. In [12] variables are ranked according to the order in which they are selected in a stepwise regression procedure. First order and total effects indices are not computed. In [25], nonparametric regression methods are used to develop a metamodel and new samples are generated to estimate the sensitivity indices from responses computed with the metamodel. The ISK method proposed here is different because it uses kernel regression with subsets of the inputs to directly estimate the primary quantity (conditional expectation) in the computation of the global sensitivity indices. This bypasses any additional simulation with the metamodel.

4. Numerical Examples

The proposed method for GSA was applied to several illustrative example problems. The first two example problems are taken from [1] and [28], respectively, and are commonly used in the GSA literature as benchmark problems. The third example was developed to highlight the differences between the proposed ISK method and traditional kernel regression methods.

4.1 Example 1: Linear Function of Normal Inputs

In the first example, the response $Y(X)$ is a linear function [1] of four independent normally distributed random variables, $Y(\boldsymbol{X}) = X^{(1)} + 2X^{(2)} + 3X^{(3)} + 4X^{(4)}$, where $X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)} \sim N(\mu = 0, \sigma = 2)$. The model output is a linear, monotonic, and additive function of the input variables. Because the function is additive, there are no interaction effects among the inputs and the exact values of the total effects indices are equal to the first order effects indices. The coefficient of $X^{(4)}$ in the model is the largest and as such $X^{(4)}$ has the largest influence on the variance of the response. Therefore, the corresponding sensitivity indices are also larger.

An uncertainty analysis with MCS generated a sample with N=10^4 evaluations of $Y(\boldsymbol{X})$. Scatterplots are presented in in Figure 2. Visual inspection reveals each input contributes to the variance of the response. A global sensitivity analysis was performed with the existing sample and no additional evaluations of the model using the ISK estimator and steps in section 3.4. For this sample, the normal rule of thumb estimates of the bandwidths using Eq. (13) were $h^{(1)}=0.3375$, $h^{(2)}=0.3368$, $h^{(3)}=0.3308$, and $h^{(4)}=0.3362$. The ISK-computed conditional expected value is plotted as red points in Figure 2.

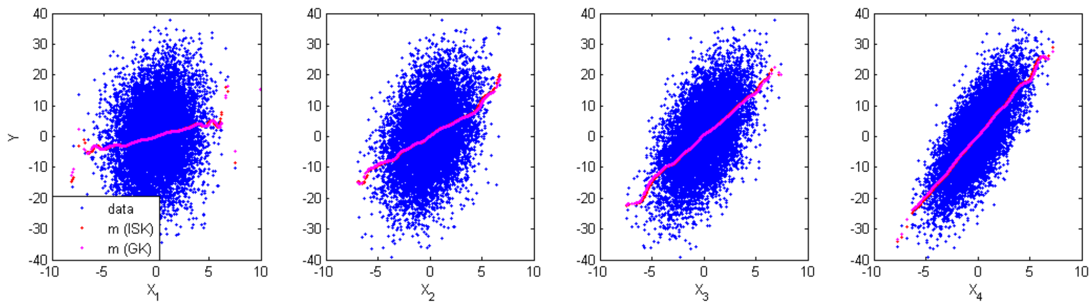

Figure 2. Scatter Plots for Linear Function Example.

The computed sensitivity indices based on the existing sample are given in Table 3. The ISK-computed sensitivity results agree well with the exact values. The sensitivity indices for $X^{(4)}$ are larger than those of the other variables.

Table 3. Sensitivity Indices for Linear Function from Importance Sampling (N = 10^4)

Variable	$S^{(j)}$ (ISK)	$S^{(j)}$ exact [1]	$T^{(j)}$ (ISK)	$T^{(j)}$ exact [1]
$X^{(1)}$	0.029	0.036	0.032	0.036
$X^{(2)}$	0.137	0.14	0.123	0.14
$X^{(3)}$	0.293	0.31	0.275	0.31
$X^{(4)}$	0.541	0.56	0.493	0.56

Convergence of the estimated sensitivities with increasing sample size was investigated. A MCS uncertainty analysis was repeated 100 times for samples sizes of 10^2, 10^3, and 10^4. A global sensitivity analysis was performed as a post-processing step for each repetition at each sample size using the ISK estimator. The mean absolute error (MAE) of the estimated sensitivities decreases with increasing sample size, as seen in Figure 3. The MAE of the first order effects index is smaller than the MAE for the total effects index for the important variables, $X^{(4)}$ and $X^{(3)}$.

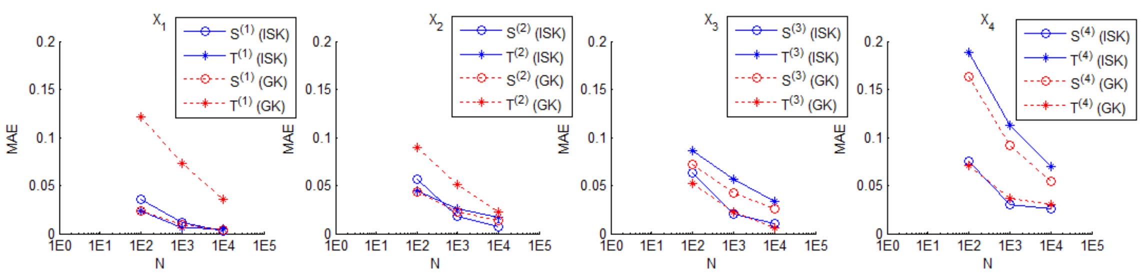

Figure 3. Mean Absolute Error of Sensitivity Estimates wrt Sample Size.

4.2 Example 2: Ishigami Function of Uniform Inputs

In the second example, the response $Y(\boldsymbol{X})$ is the Ishigami function [28], $Y(\boldsymbol{X}) = \sin(X^{(1)}) + 7\sin^2(X^{(2)}) + 0.1(X^{(3)})^4 \sin(X^{(1)})$, a nonlinear function of three independent uniformly distributed inputs, $X^{(1)}, X^{(2)}, X^{(3)} \sim U(-\pi, \pi)$. The model output is non-linear, non-monotonic, and non-additive, making it a useful benchmark problem for sensitivity methods and it appears in many studies [10, 28, 43]. In particular the variable $X^{(1)}$ has a smaller first order effects index than $X^{(2)}$ but through interaction with $X^{(3)}$ has the largest total effects index. Because the actual distributions of the inputs are uniform, when applying the proposed method, the importance sampling based kernel is equal to the target density, and thus reduces to a local constant kernel estimator.

An uncertainty analysis with MCS using a sample size of N=10^5 was performed. Scatterplots are presented in in Figure 4. Visual inspection reveals each input has a non-negligible contribution to the variance of the response. A global sensitivity analysis of the model using the ISK estimator and steps in section 3.4 was performed with the existing sample and no additional model evaluations. For this sample, the normal rule of thumb estimate of the bandwidths using Eq. (13) was $h^{(1)}=0.1918$, $h^{(2)}=0.1923$, and $h^{(3)}=0.1924$. The ISK-computed conditional expected value is plotted as red points in Figure 4.

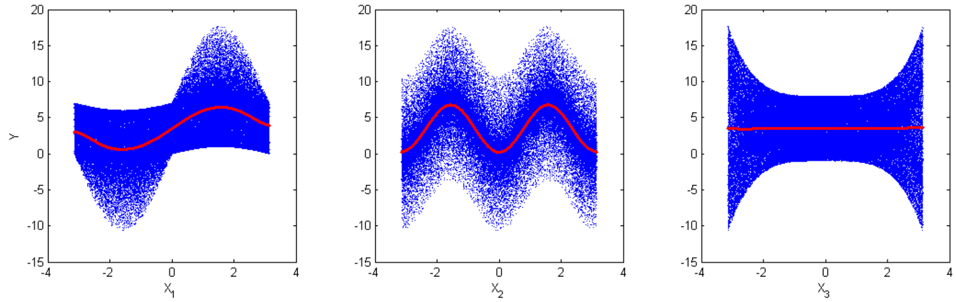

Figure 4. Scatter Plots for Ishigami Function.

Table 4 presents the estimated sensitivity indices. The results for ranking the inputs based on the contribution to the output variance agree with the published analytical values in [28]. However the sample size was rather large and the results are not as accurate as desired. Testing with other values of the bandwidths was able to achieve more accurate results, which leads to the conclusion that since the inputs are uniformly distributed, the normal rule of thumb approach gave unsatisfactory and non-optimal bandwidth estimates.

Table 4. Sensitivity Indices for Ishigami Function from Importance Sampling (N = 10^5)

Variable	$S^{(j)}$ (ISK)	$S^{(j)}$ exact [28]	$T^{(j)}$ (ISK)	$T^{(j)}$ exact [28]
$X^{(1)}$	0.307	0.3138	0.623	0.5532
$X^{(2)}$	0.376	0.4424	0.488	0.4338
$X^{(3)}$	0.0	0.0	0.317	0.2380

Convergence of the estimated sensitivities with increasing sample size was investigated. A MCS uncertainty analysis was repeated 100 times for samples sizes of 10^2, 10^3, and 10^4. A global sensitivity analysis was performed as a post-processing step for each repetition at each sample size using the ISK estimator. The MAE of the estimated sensitivities decreases with increasing sample size, as seen in Figure 5. The MAE of the total order effects index tended to be larger for variables with a significant interaction, $X^{(1)}$ and $X^{(3)}$. It is anticipated that estimation of a better bandwidth will reduce the MAE.

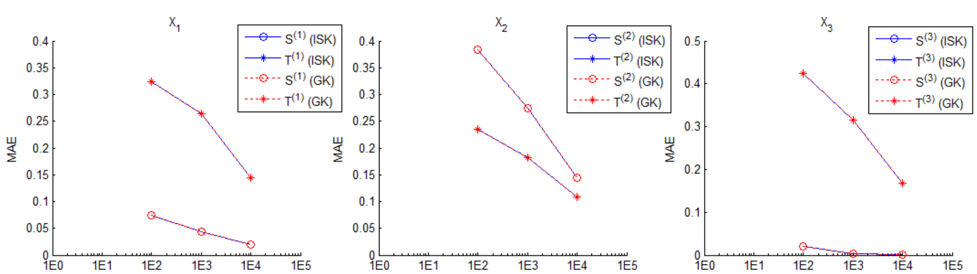

Figure 5. Mean Absolute Error of Sensitivity Estimates wrt Sample Size.

4.3 Example 3: Nonlinear Function of Normal Inputs and a Mixture Input

In the third example, the response $Y(\boldsymbol{X})$ is a nonlinear function, $Y(\boldsymbol{X}) = X_1 + 2X_2 + 100X_3^{\,2}$, of three independent random variables. Variables $X^{(1)}$ and $X^{(2)}$ have a standard normal distribution, $X^{(1)}, X^{(2)} \sim N(0,1)$. The pdf of variable $X^{(3)}$ is a mixture distribution, $X^{(3)} \sim 0.5N(-0.25,0.1) + 0.5N(0.25,0.1)$, consisting of two component normal pdfs, and is bimodal and symmetric about zero.

An MCS uncertainty analysis for a sample size of N=10^3 model evaluations was performed. Figure 6 contains scatter plots of the model output with respect to each of the inputs with the data plotted as blue points. The bimodal distribution of $X^{(3)}$ is apparent as well as the quadratic response of the output with respect to $X^{(3)}$. The minor influence of the other inputs is visually obvious. A global sensitivity analysis of the model using the ISK estimator and steps in section 3.4 was performed with the existing sample and no additional model evaluations. In the importance sampling kernel function, the target pdf was Gaussian. For this sample, the normal rule of thumb sample estimates for the bandwidths using Eq. (13) were $h^{(1)}$*=0.2635*, $h^{(2)}$*=0.2733*, and $h^{(3)}$*=0.0727*. The computed conditional expectation for each point is plotted as red points and labeled *m (ISK)*. Traditional kernel regression with a Gaussian kernel generated the estimates plotted as magenta points and labeled *m (GK).*

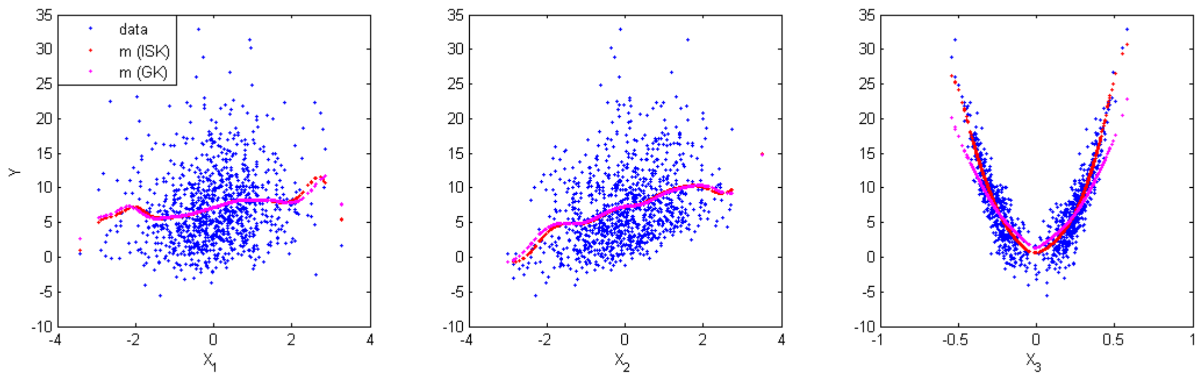

Figure 6. Scatter Plots for Nonlinear Function of 2 Normal Inputs and 1 Mixture Input.

The computed sensitivities are given in Table 4. Variable $X^{(3)}$ drives most of the variance of the output. Analytically-derived exact values of the sensitivities are not available for this problem for comparison. In their place estimates were developed using the methods and estimators in [15] with a base sample of N=10^7 simulations. Note these “truth” estimates required an *a priori* sample design of $N(3+2)$ samples, where N was increased to N=10^7 until the values of the sensitivity indices converged. The sensitivities estimated with the ISK agree well with the values computed using the method in [1].

Table 4. Sensitivity Indices for Nonlinear Function from Importance Sampling (N = 10^3)

Variable	$S^{(j)}$ (ISK)	$S^{(j)}$ ~exact	$T^{(j)}$ (ISK)	$T^{(j)}$ ~exact
$X^{(1)}$	0.030	0.031	0.049	0.031
$X^{(2)}$	0.151	0.125	0.164	0.125
$X^{(3)}$	0.813	0.845	0.798	0.844

Using 100 repetitions of the MCS and GSA using ISK, the mean absolute error (MAE) of the sensitivity estimates was investigated at sample sizes of N=10^2, 10^3, and 10^4. The MAE decreases as the sample size increases as seen in Figure 7. The sensitivities were also estimated with a local constant Gaussian kernel function and the MAE is plotted as red lines in Figure 7. The MAE of the first order effects index for $X^{(3)}$ is significantly larger for the Gaussian kernel (GK) estimate than the ISK estimate.

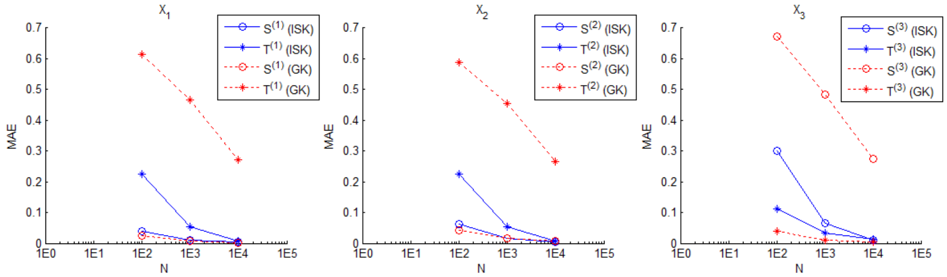

Figure 7. Mean Absolute Error of Sensitivity Estimates wrt Sample Size.

Figure 8 gives a more detailed view of the scatter plot for $X^{(3)}$ in a manner similar to Figure 1. The estimated conditional expectation using the importance sampling kernel agrees much more closely with the analytical exact values than using the traditional Gaussian kernel, particularly in the tails. The corresponding estimated global sensitivities are more accurate using the ISK than the local constant Gaussian kernel.

The importance sampling-based kernel function for conditioning on $x^{(3)} = -0.017$ is plotted. This point lies between the two modes of the input variable. Due to the small sample size, the traditional Gaussian kernel *w (GK)* has a wide bandwidth and the samples near each of the modes contaminate the estimated local average. The importance sampling kernel *w (ISK)* adapts to the local density and in this case assigns less weight away from the center than the traditional kernel, particularly in the ranges of $-0.2 < x < -0.1$ and $0.1 < x < 0.2$.

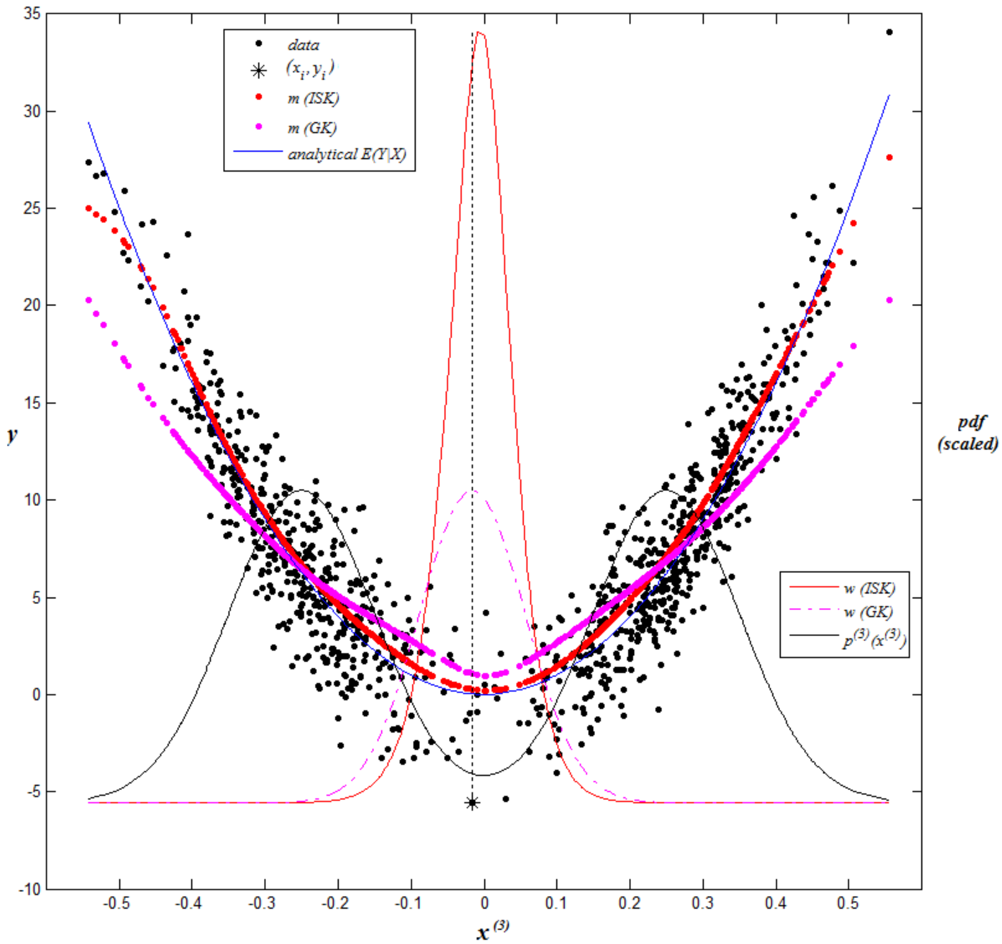

Figure 8. Comparison of Importance Sampling-based Kernel and Gaussian Kernel for Example 3 -
The IS kernel places more weight on samples closer to $x^{(3)}$ *=-0.017 than the traditional Gaussian kernel.*

4.4 Discussion

The variance-based global sensitivity indices are calculated with proposed importance sampling-based estimator in the manner of kernel regression. The likelihood ratio function is in fact a particular type of kernel function and the intent of both is to estimate a local expected value. Drawing on ideas from importance sampling, a theoretically ideal model for the kernel does exist, but is problem dependent and generally unknown as it depends on the unknown value m_i we are trying to estimate. If this value was known, it could be used to generate a kernel function that could return the exact value of the conditional expectation from just one realization. Even though the local mean is unknown, in the context of simulation the distribution that the sample was drawn from is known. With previous methods, the burden of carrying this information into the analysis is supported only by the samples themselves. The ISK incorporates knowledge of the sampling distribution into the kernel function. This additional information is carried external to the samples and later leveraged in the sensitivity analysis. This can result in an improvement in the accuracy of the estimates, as seen in the third example problem. Also, the ISK provides an attractive way to adapt to the local density due to its simplicity.

The ISK method was developed for GSA in the case in which an existing sample of simulations is available from a previous MCS uncertainty analysis. In this case other GSA methods that require an *a priori* sample design are not applicable. The ISK post processing method was demonstrated on several example problems. The ISK method is generally less accurate for computing the total effects index than the first order effects index. This is because when estimating the total effects index, the response is conditioned on all of the variables except one. Due to the curse of dimensionality, local averaging thus takes place over a much smaller space, there are fewer samples available to contribute significantly to the estimated mean, and the estimate is less accurate. The problem becomes worse as the number of input variables increases. The estimate for the first order effects index is not affected as much by the number of input variables.

The examples show that the accuracy of both the first order effects index and the total effects index improves as the size of the existing sample increases. A related aspect of the method that impacts the accuracy of the estimated sensitivities is the computed bandwidth. The second example shows how the normal rule of thumb estimator for the bandwidth (Eq. (13)) can provide poor estimates of a suitable bandwidth when the inputs are not normally distributed. The accuracy of the corresponding computed sensitivities is adversely affected. Future work will address this issue for estimation of optimal bandwidths. The third example highlights how computing sensitivities from an existing sample with the ISK estimator can be more accurate than using existing kernel functions with kernel regression. The sensitivities for some datasets for which a local constant kernel function would not be accurate enough may still be accessible with the proposed ISK estimator.

5. Conclusion

The importance sampling-based post processing method allows for efficient estimation of global sensitivity indices using an existing sample without requiring additional evaluations of the response function. It is presented as an extension of kernel regression methods that incorporates the knowledge of the distribution the samples were drawn from. In particular, the importance sampling kernel is a locally-adaptive kernel that is simple to implement.

The technique was demonstrated with several illustrative examples, including two benchmark problems from the literature. The examples give empirical evidence for more accurate estimation of variance-based global sensitivity indices for small sample sizes than kernel regression using a local constant kernel. The estimator uses some of the same construction as the local constant kernel, with a bandwidth that shrinks as the sample size increases. This makes the estimator a consistent estimator, and the examples give empirical evidence that the accuracy of the estimated sensitivity indices improves with larger samples sizes.

Future efforts will make the ISK estimator more data driven. The normal rule of thumb estimator for the bandwidth is attractive due to its simplicity, but is most useful when the input data is univariate and known to be normally distributed. In the kernel regression literature, bandwidth estimation is of prime importance and methods including plug-in estimators and cross validation are used to obtain optimal bandwidths. The same principle should still apply for the ISK, and more accurate sensitivity estimates are anticipated.

Acknowledgments

Authors DMS and HRM acknowledge the financial support provided by AFOSR Grant FA9550-09-1-0452, and DMS acknowledges the financial support of TOPS IV FA8650-11-D-5800 TO 009. DMS thanks Eric Tuegel, Pam Kobryn, Steven Turek, and Brenchley Boden II for mentoring and support on this work.

References

[1] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., Global Sensitivity Analysis: The Primer. John Wiley & Sons Ltd, England. 2008, pp. 160-165.

[2] Beckman, R., and McKay, M., "Monte Carlo Estimation under Different Distributions Using the Same Simulation", *Technometrics*, 29(2), 1987. pp. 153-160.

[3] Millwater, H., Singh, G., and Cortina, M., "Development of a localized probabilistic sensitivity method to determine random variable regional importance," *Reliability Engineering and System Safety*, 2011.

[4] Millwater, H., "Universal properties of kernel functions for probabilistic sensitivity analysis," *Probabilistic Engineering Mechanics*, 24, 2009. pp. 89-99.

[5] Li, Q., and Racine, J., Nonparametric Econometrics: theory and practice. Princeton University Press. 2007.

[6] Wei, P., Lu, Z., Hao, W., Feng, J. and Wang, B., "Efficient sampling methods for global reliability sensitivity analysis", Computer Physics Communications, 183, 2012, pp. 1728-1743.

[7] Helton J, Johnson JD, Sallaberry CJ, Storlie CB. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliability Engineering and System Safety, 91. 2006: pp. 1175-1209.

[8] Plischke, E., An effective algorithm for computing global sensitivity indices (EASI), Reliability Engineering and System Safety. 95. 2010: pp. 354-360.

[9] Sobol, I.M., Levitan, Y.L., On the use of variance reducing multipliers in Monte Carlo computations of a global sensitivity index. Computer Physics Communications, 117. 1991: pp. 52-61.

[10] Ratto, M., Pagano, A., Young, P.C., Non-parametric estimation of conditional moments for sensitivity analysis. Reliability Engineering and System Safety, 94. 2009: pp. 237-243

[11] Pianosi, F., Wagener, T., A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, Environmental Modelling & Software, 67. 2015: pp. 1-11.

[12] Storlie, C.B., Helton, J.C., Multiple predictor smoothing methods for sensitivity analysis: Description of techniques, Reliability Engineering and System Safety, 93. 2008: pp. 28-54.

[13] Wei, P., Lu, Z., Yuan, X., Monte Carlo simulation for moment-independent sensitivity analysis, Reliability Engineering and System Safety, 110. 2013: pp. 60-67.

[14] Pannier, S., Graf, W., Sectional global sensitivity measures, Reliability Engineering and System Safety, 134. 2015: pp. 110-117.

[15] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Computer Physics Communications, 181. 2010: pp. 259-270.

[16] Luo, X., Lu, Z., Xu, X., Non-parametric kernel estimation for the ANOVA decomposition and sensitivity analysis, Reliability Engineering and System Safety, 130. 2014: pp. 140-148.

[17] Doksum, K., Samarov, A., Nonparametric estimation of global functionals and a measure of the explanatory power of covariates in regression, The Annals of Statistics, Vol. 23, No. 5. 1995: pp. 1443-1473.

[18] Dimov, I., Georgieva, R., Monte Carlo algorithms for evaluating Sobol' sensitivity indices, Mathematics and Computers in Simulation, 81. 2010: pp. 506-514.

[19] Cheng, L., Lu, Z., Zhang, L., Application of Rejection Sampling based methodology to variance based parametric sensitivity analysis, Reliability Engineering and System Safety, 142. 2015: pp. 9-18.

[20] Morris, M., Factorial Sampling Plans for Preliminary Computational Experiments, Technometrics, Vol. 33, No. 2. 1991: pp. 161-174

[21] Campolongo, F. Cariboni, J., Saltelli, A., An effective screening design for sensitivity analysis of large models, Environmental Modelling & Software, 22. 2007: pp. 1509-1518.

[22] Cukier, R.I, Fortuin, C.M., Shuler, K.E., Petschek, A.G., Schaibly, J.H., Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory, The Journal of Chemical Physics, Vol. 59, No. 8. 1973: pp. 3873-3878.

[23] Nadaraya, E.A., On estimating regression, Theory of Probability and its Application, 9. 1964: pp. 141-142.

[24] Epanechnikov, V.A., Nonparametric estimation of a multivariate probability density, Theory of Probability and its Application, 14. 1969: pp. 153-158.

[25] Storlie, C.B., Swiler, L.P., Helton, J.C., Sallaberry, C.J., Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models, Reliability Engineering and System Safety, 94. 2009: pp. 1735-1763.

[26] Plischke, E., Borgonovo, E., Smith, C., Global sensitivity measures from given data, European Journal of Operational Research, 226. 2013: pp. 536-550.

[27] Hesterberg, T.C., Estimates and Confidence Intervals for Importance Sampling Sensitivity Analysis, Mathematical Computer Modeling, Vol. 23, No. 8/9. 1996: pp. 79-85.

[28] Homma, T., Saltelli, A., Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering and System Safety, 52. 1996: pp. 1-17.

[29] Allaire, D.L., Willcox, K.E., A variance-based sensitivity index function for factor prioritization, Reliability Engineering and System Safety, 107. 2012: pp. 107-114.

[30] Craven, P., Wahba, G., Smoothing Noisy Data with Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation, Numerische Mathematik, 31. 1979: pp. 377-403.

[31] Abramson, I., On Bandwidth Variation in Kernel Estimates – A Square Root Law, The Annals of Statistics, Vol. 10, No. 4. 1982: pp. 1217-1223.

[32] Terrell, G., Scott, D., Variable Kernel Density Estimation, The Annals of Statistics, Vol. 20, No. 3. 1992: pp. 1236-1265.

[33] Fan, J., Gijbels, I., Variable Bandwidth and Local Linear Regression Smoothers, The Annals of Statistics, Vol. 20, No. 4. 1992: pp. 2008-2036.

[34] Hastie, T., Loader, C., Local Regression: Automatic Kernel Carpentry, Statistical Science, Vol. 8, No. 2. 1993: pp. 120-129.

[35] Zhai, Q., Yang, J., Zhao, Y., Space-partition method for the variance-based sensitivity analysis: Optimal partition scheme and comparative study, Reliability Engineering and System Safety, 131. 2014: pp. 66-82.

[36] Wei, P., Lu, Z., Song, J., A new variance-based global sensitivity analysis technique, Computer Physics Communications, 184. 2013: pp. 2540-2551.

[37] Wei, P., Lu, Z., Song, J., Regional and parametric sensitivity analysis of Sobol' indices, Reliability Engineering and System Safety, 137. 2015: pp. 87-100.

[38] Ibrahim, Y., Observations on Applications of Importance Sampling in Structural Reliability Analysis, Structural Safety, 9. 1991: pp. 269-282.

[39] Chen, Y., Another look at rejection sampling through importance sampling, Statistics & Probability Letters, 72. 2005: pp. 277-283.

[40] Borgonovo, E., Apostolakis, G.E., A new importance measure for risk-informed decision making, Reliability Engineering and System Safety, 72. 2001: pp. 193-212.

[41] Loftsgaarden, D.O., Quesenberry, C.P., A Nonparametric Estimate of a Multivariate Density Function, Annals of Mathematical Statistics, Vol. 36, No. 3. 1965: pp. 1049-1051.

[42] Abadir, K.M., Lawford, S., Optimal asymmetric kernels, Economics Letters, 83. 2004: pp. 61-68.

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

[44] Watson, G.S., Smooth regression analysis, Sankhya, Series A, 26. 1964: pp. 359-372.

[45] Fan, J., Design-adaptive Nonparametric Regression, Journal of the American Statistical Association, Vol. 87, No. 420. 1992: pp. 998-1004.