

Estimating Sensitivity Indices Based on Gaussian Process Metamodels with Compactly Supported Correlation Functions

Joshua Svenson ^a, Thomas Santner ^{*,b}, Angela Dean ^{b,c}, and
Hyejung Moon ^d

^a*JPMorgan Chase & Co., 1111 Polaris Parkway,
Columbus, OH, 43240, USA*

^b*Department of Statistics, The Ohio State University
1958 Neil Avenue, Columbus, OH 43210, USA*

^c*Department of Mathematics, University of Southampton
Southampton, SO17 1BJ, UK*

^d*The Bank of Korea, 110, 3-Ga, Namdaemun-Ro, Jung-Gu, Seoul, 100-794, Korea*

Abstract

Specific formulae are derived for quadrature-based estimators of global sensitivity indices when the unknown function can be modeled by a regression plus stationary Gaussian process using the Gaussian, Bohman, or cubic correlation functions. Estimation formulae are derived for the computation of process-based Bayesian and empirical Bayesian estimates of global sensitivity indices when the observed data are the function values corrupted by noise. It is shown how to restrict the parameter space for the compactly supported Bohman and cubic correlation functions so that (at least) a given proportion of the training data correlation entries are zero. This feature is important in the situation where the set of training data is large. The estimation methods are illustrated and compared via examples.

AMS Subject Classification: 62K05

Key words: Bayesian estimation, Computer experiments, Global sensitivity indices, Main-effect sensitivity indices, Process-based estimator, Quadrature-based estimator, Total sensitivity indices,

1 Introduction

A *computer experiment* uses a computer simulator based on a mathematical model of a physical process as an experimental tool to determine “responses” or “outputs” at a set of user-specified input sites. These input sites constitute the design for the computer experiment. Sophisticated computer codes may take hours or even days to produce an output and, therefore, a flexible and rapidly-computable predictor, sometimes called a code *emulator* or *metamodel*, is often fitted to the inputs/outputs of the design, which are then called *training data*. An emulator allows the detailed, albeit approximate, exploration of the output over the entire experimental region (see, for example, Sacks, Welch, Mitchell, and Wynn 1989b; Santner, Williams, and Notz 2003). A *sensitivity analysis*, based on the outputs of either the simulator or emulator, enables the researcher to assess the variation in the output due to changes in individual inputs or groups of inputs (see, for example Saltelli, Tarantola, and Campolongo 2000; Helton, Johnson, and Storlie 2006; Oakley and OHagan 2004).

In this paper, we assume that the computer simulator has d continuous input variables denoted by the vector $\mathbf{x} = (x_1, \dots, x_d)$ and that the (one-dimensional) output of the simulator, denoted by $y(\mathbf{x}) = y(x_1, \dots, x_d)$, can be determined for \mathbf{x} in the hyper-rectangle $\mathcal{X} = \prod_{j=1}^d [l_j, u_j]$, but is computationally expensive. The sensitivity of $y(\mathbf{x})$ to the input values \mathbf{x} can be measured locally or globally. A *local sensitivity index* is based on the change in $y(\cdot)$ at a specified $\mathbf{x}^0 = (x_1^0, \dots, x_d^0)$ as the j^{th} input varies by a small amount parallel to the x_j axis and this can be measured by the partial derivatives of $y(\cdot)$ with respect to x_j . In contrast, a first (or higher) order *global sensitivity index* measures the change in $y(\cdot)$ as one (or more) inputs vary over their entire range, when the remaining inputs are fixed (see, for example Saltelli 2002). Homma and Saltelli (1996) further defined the j th *total sensitivity index* as a measure of the change in $y(\cdot)$ due to the j th input, both through its main effect and its joint effect with other inputs. Chen, Jin, and Sudjianto (2005, 2006) defined *subset sensitivity indices* based on non-overlapping partitions of the inputs. One popular definition of global sensitivity indices is in terms of the variability of the (weighted) average output $y(\mathbf{x})$ over $\mathbf{x} \in \mathcal{X} = \prod_{j=1}^d [l_j, u_j]$, as reviewed in Section 2.

As well as providing an understanding of the input/output relationship, sensitivity analysis provides a tool for “screening”, that is for selecting the inputs that have major impacts on

* Corresponding author. Tel: +01-614-292-2866; fax: +01-614-292-2096

Email addresses: `joshua.d.svenson@chase.com` (Joshua Svenson), `santner.1@osu.edu` (Thomas Santner), `dean.9@osu.edu` (Angela Dean), `hymoonoh@gmail.com` (Hyejung Moon).

an input-output system, thereby allowing researchers to restrict attention to these important inputs while setting the others to nominal values in their computational simulator. For various discussions and applications of sensitivity analysis and screening, see for example, Welch, Buck, Sacks, Wynn, Mitchell, and Morris (1992), Linkletter, Bingham, Hengartner, Higdon, and Ye (2006), Moon, Santner, and Dean (2012), and the references cited therein.

For estimating local sensitivity indices, Morris (1991) proposed the use of “elementary effects” calculated directly from the simulator output, with inputs selected according to a “one-at-time” sampling design. This methodology was extended by Campolongo, Cariboni, and Saltelli (2007). Sampling designs for estimating global sensitivity indices were presented and discussed by, for example, Saltelli (2002), Morris, Moore, and McKay (2008), Da Viegua, Wahl, and Gamboa (2009), and Saltelli, Annoni, Azzini, Campolongo, Ratto, and Tarantola (2010). In the case when the simulator is expensive to run, such estimation methods may require more simulator runs than is feasible to produce accurate global sensitivity index estimates. Chen et al. (2005), Oakley and O’Hagan (2004), Marrel, Iooss, Lauren, and Roustant (2009), and Storlie, Reich, Helton, Swiler, and Sallaberry (2013) gave alternative estimation methods based on analytical and probabilistic methods using emulators.

In this paper, we use the popular $y(\mathbf{x})$ emulator based on a Gaussian process model as proposed, for example, by Sacks et al. (1989b), and which has the form

$$Y(\mathbf{x}) = \mathbf{f}^\top(\mathbf{x})\boldsymbol{\beta} + Z(\mathbf{x}), \quad (1.1)$$

where $\mathbf{f}^\top(\mathbf{x})\boldsymbol{\beta}$ is a linear function of an unknown regression parameter vector $\boldsymbol{\beta}$, and $Z(\mathbf{x})$ is a zero-mean Gaussian process having variance σ^2 . Assuming this type of model, Sacks et al. (1989b), Sacks, Schiller, and Welch (1989a), and Welch et al. (1992) used a $y(\mathbf{x})$ predictor derived from the classical theory of best linear unbiased prediction. Other authors, including Currin, Mitchell, Morris, and Ylvisaker (1991), O’Hagan (1992), Oakley and O’Hagan (2004), have viewed the random function $Y(\mathbf{x})$ as representing prior uncertainty about the true function and adopted a Bayesian approach to estimation.

The purpose of this paper is to give specific formulae for global sensitivity index estimates for a broad class of regression plus Gaussian process models (1.1) with independent inputs in the special case of stationary $Z(\mathbf{x})$ with compactly supported Bohman and cubic (separable) correlation functions. As compared with the often-used Gaussian correlation function, use of compactly supported correlation functions together with a suitably rich mean structure has the potential to provide sparse correlation matrices, thus allowing prediction to be performed

with larger data sets within the Gaussian process framework (see Kaufman, Bingham, Habib, Heitmann, and Frieman 2011).

In Section 3, we give formulae for quadrature-based methods of estimation using Gaussian processes with polynomial mean and either Gaussian or Bohman correlation functions. In the on-line Supplementary Material, we provide the corresponding formulae for the cubic correlation function. In Section 4, together with the Supplementary Material, we derive the specific formulae required to compute both fully Bayesian and empirical (plug-in) Bayesian estimates of sensitivity indices. The formulae in these two sections extend the work of Chen et al. (2005), Oakley and OHagan (2004), Marrel et al. (2009), and others, who provide explicit formulae for global sensitivity estimators for Gaussian process emulators with constant mean and Gaussian correlation function.

In Section 6, it is shown via two examples that sensitivity indices estimated using output from a Gaussian process emulator under the compactly supported Bohman, and cubic correlation functions are similar to the estimates obtained using the Gaussian correlation function, but that the computational times are much shorter. Although the current examples are not extremely large, they illustrate the potential computational savings, described by Kaufman et al. (2011), that can be achieved when handling large data sets and/or large numbers of inputs. In line with previous studies, our examples also illustrate that calculation of sensitivity indices using a moment-based estimation method (based on “permuted column sampling” as described by Morris et al. 2008) is less accurate when using only a moderate number of simulator runs. Finally, Section 7 shows how to restrict the parameter space for the Bohman and cubic correlation functions so that (at least) a given proportion of the training data correlation entries are zero.

2 Calculation of Main Effect and Total Effect Sensitivity Indices

In this section, we review definitions of main effect and total effect global sensitivity indices, as described by Homma and Saltelli (1996); Saltelli (2002); Chen et al. (2005, 2006), for example. Throughout the paper, $Q = \{k_1, \dots, k_s\} \subset \{1, 2, \dots, d\}$ denotes a non-empty subset of the input variables and \mathbf{x}_Q denotes the vector of inputs $(x_{k_1}, \dots, x_{k_s})$ where, for definiteness, it is assumed $1 \leq k_1 < k_2 < \dots < k_s \leq d$. The vector of the remaining inputs will be denoted by \mathbf{x}_{-Q} also arranged in lexicographical order of their input index. By rearranging the order of the entire set of input variables we write the input vector \mathbf{x} as $\mathbf{x} = (\mathbf{x}_Q, \mathbf{x}_{-Q})$ in a slight abuse of notation.

Throughout the paper, we take $[l_j, u_j] = [0, 1]$, for all inputs x_j , $j = 1, \dots, d$, so that $\mathcal{X} = [0, 1]^d$. The formulae can be extended to the more general hyper-rectangle case. Also for simplicity of notation, it is assumed that the weight function can be specified by a joint density function over $\mathcal{X} = [0, 1]^d$ having independent and identically distributed x_j components each with probability density function $g(x)$. For any subset $E \subseteq \{1, \dots, d\}$ the notation $g(\mathbf{x}_E)$ denotes $\prod_{\ell \in E} g_\ell(x_\ell)$. It is a straightforward generalization to allow the weight function to have independent but not identically distributed components.

For any non-empty $Q = \{k_1, \dots, k_s\} \subset \{1, 2, \dots, d\}$, the *uncorrected mean effect* (also known as the *joint effect function*) of the input vector \mathbf{x}_Q on $y(\cdot)$ is defined to be the average $y(\mathbf{x}_Q, \mathbf{x}_{-Q})$ over \mathbf{x}_{-Q} ; that is,

$$u_Q(\mathbf{x}_Q) = \int y(\mathbf{x}_Q, \mathbf{x}_{-Q})g(\mathbf{x}_{-Q})d\mathbf{x}_{-Q} = E_g[y(\mathbf{X})|\mathbf{X}_Q = \mathbf{x}_Q], \quad (2.1)$$

and $u_{\{1, \dots, d\}} = y(\mathbf{x})$. The notation makes clear that the function average can be viewed as an expectation with respect to subcomponents of \mathbf{X} defined by Q . When $Q = \{j\}$ for a given $j \in \{1, \dots, d\}$, then $u_j(x_j)$ is called the (*uncorrected*) *main effect function* of input j associated with $y(\mathbf{x})$. Plots of the main effect functions $u_j(x_j)$ versus x_j , and plots of the joint effect functions $u_{j_1 j_2}(x_{j_1}, x_{j_2})$ versus pairs of inputs (x_{j_1}, x_{j_2}) can be used to provide a visual understanding of the change in the averaged $y(\mathbf{x})$ with respect to each single input or pairs of inputs (see, for example, Jones, Schonlau, and Welch 1998).

To define global sensitivity indices, Sobol' (1990) and Sobol' (1993) advocated the use of a functional analysis of variance (ANVOA) decomposition of $y(\mathbf{x})$ as follows:

$$y(\mathbf{x}) = y_0 + \sum_{j=1}^d y_j(x_j) + \sum_{1 \leq j_1 < j_2 \leq d} y_{j_1, j_2}(x_{j_1}, x_{j_2}) + \dots + y_{1, 2, \dots, d}(x_1, \dots, x_d), \quad (2.2)$$

where

$$y_0 = \int y(\mathbf{x})g(\mathbf{x}) d\mathbf{x} = E_g[y(\mathbf{X})] \quad (2.3)$$

denotes the overall (weighted) mean of $y(\mathbf{x})$, expressing the fact that inputs x_1, \dots, x_d have distribution $g(\cdot)$. The component terms of (2.2), called *corrected mean effect functions*, are defined recursively to be

$$y_Q(\mathbf{x}_Q) = u_Q(\mathbf{x}_Q) - \sum_{E \subset Q} y_E(\mathbf{x}_E) - y_0, \quad (2.4)$$

where the sum is over the collection of all non-empty, *proper* subsets E of $Q \subseteq \{1, \dots, d\}$ (Q non-empty). The components of (2.2) satisfy $E_g[y_Q(\mathbf{X}_Q)] = 0$ with respect to any sub-

component of \mathbf{X}_Q , for $Q \subset \{1, 2, \dots, d\}$, and are *pairwise orthogonal*, meaning that for any $Q_1 \neq Q_2 \subset \{1, \dots, d\}$ and $Q = Q_1 \cup Q_2$, $E_g[y_{Q_1}(\mathbf{X}_{Q_1})y_{Q_2}(\mathbf{X}_{Q_2})] = 0$, (cf. Van Der Vaart 1998, Section 11.4). Using these facts, *global sensitivity indices*,

$$S_Q = v_Q/v, \quad (2.5)$$

are defined as functions of the variances v_Q of the corrected effect functions:

$$v_Q = \text{Var}_g[y_Q(\mathbf{X}_Q)] = \int y_Q^2(\mathbf{x}_Q)g(\mathbf{x}_Q) d\mathbf{x}_Q, \quad (2.6)$$

where v is the *total variance* of $y(\mathbf{x})$; that is,

$$v = \text{Var}_g[y(\mathbf{X})] = \int y^2(\mathbf{x})g(\mathbf{x})d\mathbf{x} - y_0^2, \quad (2.7)$$

(cf. Homma and Saltelli 1996; Saltelli 2002). Due to the pairwise orthogonality of the components of (2.2), the total variance can be partitioned as

$$v = \sum_{j=1}^d v_j + \sum_{1 \leq j_1 < j_2 \leq d} v_{j_1 j_2} + \dots + v_{1,2,\dots,d}, \quad (2.8)$$

with v_Q defined as in (2.6). The quantity $S_j = v_j/v$ is called the *j*th *main effect sensitivity index* and $S_{j_1 j_2} = v_{j_1 j_2}/v$ is a *two-factor sensitivity index*. By (2.8),

$$\sum_{j=1}^d S_j + \sum_{1 \leq j_1 < j_2 \leq d} S_{j_1 j_2} + \dots + S_{1,2,\dots,d} = 1.$$

The *total effect sensitivity index* of input x_j was defined by Homma and Saltelli (1996) to be the sum of *all* sensitivity indices involving input x_j ,

$$T_j = S_j + \sum_{k \neq j} S_{kj} + \dots + S_{1,2,\dots,d}. \quad (2.9)$$

For example, when there are $d = 3$ inputs, then $T_1 = S_1 + S_{12} + S_{13} + S_{123}$. Notice that by construction, $S_j \leq T_j$ for all $j \in \{1, \dots, d\}$. The difference between T_j and S_j will be large if interactions involving x_j account for a large proportion of the variance v .

The main effect sensitivity indices $S_j = v_j/v$, $j = 1, \dots, d$, can also be computed easily in terms of the variances, $v_j^u = \text{Var}_g[u_j(x_j)]$, of the uncorrected main effect functions, since

$$v_j = \text{Var}_g[y_j(X_j)] = \text{Var}_g[u_j(X_j) - y_0] = \text{Var}_g[u_j(X_j)] = v_j^u. \quad (2.10)$$

The total effect sensitivity index, T_j , can also be computed efficiently in terms of the variances of uncorrected mean functions (Homma and Saltelli 1996) as follows. For given nonempty $Q \subseteq \{1, \dots, d\}$, the variance v_Q^u of the uncorrected mean effect function $u_Q(\mathbf{x}_Q)$ is

$$v_Q^u = \text{Var}_g[u_Q(\mathbf{X}_Q)] = \text{Var}_g[E_g[y(\mathbf{X})|\mathbf{X}_Q]] = \text{Var}_g[y(\mathbf{X})] - E_g[\text{Var}_g(y(\mathbf{X})|\mathbf{X}_Q)].$$

So v_Q^u can be interpreted as the expected reduction in uncertainty in $y(\mathbf{X})$ due to observing \mathbf{x}_Q . Denoting the set of indices $\{1, \dots, j-1, j+1, \dots, d\}$ by “ $-j$ ”, (2.4)–(2.6) imply that

$$v_{-j}^u = \text{Var}_g[u_{-j}(\mathbf{X}_{-j})] = \text{Var}_g\left[\sum y_Q(\mathbf{X}_Q)\right] = \sum v_Q, \quad (2.11)$$

where the sum is over nonempty sets Q contained in $\{1, \dots, d\} \setminus \{j\}$. In words, v_{-j}^u is the sum of all v_Q components that do *not involve* the subscript j in the variance decomposition (2.8). Thus, $v - v_{-j}^u$ is the sum of all v_Q components for which $j \in Q$, and so the total effect sensitivity index T_j in (2.9) can be expressed as

$$T_j = (v - v_{-j}^u)/v. \quad (2.12)$$

Consequently, if only the *main effect* and *total effect* sensitivity indices $\{S_j\}_{j=1}^d$ and $\{T_j\}_{j=1}^d$ are to be estimated, then one need only estimate the variances of $2d$ uncorrected effect functions rather than the variances of $2^d - 1$ corrected effect functions required by (2.9); see Homma and Saltelli (1996).

Sections 3 and 4 describe two general methods of estimating the variance v_Q^u , each using a Gaussian process underlying model (1.1). The first uses quadrature-based estimation, while the second uses Bayesian or empirical Bayesian process-based estimation. The estimates are compared via examples in Section 6.

3 Quadrature-based Estimators of Global Sensitivity Indices

This section describes the calculation of quadrature-based estimators of global sensitivity indices which rely on predictors based on the Gaussian process model (1.1) which has a regression mean. Quadrature-based estimation replaces $y(\mathbf{x})$ in the variance expressions such as v in (2.7) and v_j^u , $j = 1, \dots, d$, in (2.10) by a predictor $\hat{y}(\mathbf{x})$ and integrates the associated expectations. We illustrate the calculations for the special case of predictors based on a

stationary process $Z(\mathbf{x})$ with separable correlation function

$$Cor(Z(\mathbf{x}_i), Z(\mathbf{x}_k)) = \prod_{j=1}^d R(x_{ij} - x_{kj} | \psi_j), \quad (3.1)$$

for $\mathbf{x}_\ell = (x_{\ell 1}, \dots, x_{\ell d})$ with $R(\cdot | \psi_j)$ known up to an unknown (vector of) parameter(s), ψ_j , associated with the j^{th} input, and each input x_j scaled to $[0, 1]$. When $Z(\mathbf{x})$ has a correlation function of the form (3.1) and the weight function $g(\mathbf{x})$ consists of independent components, $g(\mathbf{x}) = \prod_{j=1}^d g_j(x_j)$, then variances such as v and v_j^u can be calculated as a product of one-dimensional integrals. For some correlation functions and choices of independent components for $g(\mathbf{x})$, these one-dimensional integrals can be integrated explicitly.

The most widely-applied version of (3.1) is the separable Gaussian correlation function

$$R_G(h_j | \psi_j) = \exp[-\psi_j h_j^2] \quad \psi_j > 0. \quad (3.2)$$

The Bohman and cubic correlation functions are other useful examples of (3.1) for which quadrature-based estimators can be derived explicitly. These are compactly supported correlation functions which allow large data sets to be handled (see Kaufman et al. 2011). The Bohman correlation function has the form

$$R_B(h_j | \psi_j) = \begin{cases} \left(1 - \frac{|h_j|}{\psi_j}\right) \cos\left(\frac{\pi|h_j|}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi|h_j|}{\psi_j}\right), & |h_j| < \psi_j; \\ 0, & \psi_j \leq |h_j|; \end{cases} \quad (3.3)$$

where $\psi_j > 0$, while the cubic correlation function has the form

$$R_C(h_j | \psi_j) = \begin{cases} 1 - 6\left(\frac{h_j}{\psi_j}\right)^2 + 6\left(\frac{|h_j|}{\psi_j}\right)^3, & |h_j| \leq \frac{\psi_j}{2}; \\ 2\left(1 - \frac{|h_j|}{\psi_j}\right)^3, & \frac{\psi_j}{2} \leq |h_j| \leq \psi_j; \\ 0, & \psi_j < |h_j|, \end{cases} \quad (3.4)$$

where $\psi_j > 0$.

General formulae for quadrature-based estimators of global sensitivity indices are described next under a regression mean and correlation functions of the form (3.1) and $g(\mathbf{x}) = \prod_{j=1}^d g_j(x_j)$. Then explicit formulae are given under Gaussian and Bohman correlation functions with $g_j(x_j)$ being $U(0, 1)$. The corresponding formulae for the cubic case are provided in the Supplementary Materials. These formulae provide extensions to the cases studied by

Chen et al. (2005) who gave explicit integrals under constant mean and Gaussian correlation function for normal and uniform weight functions, as well as other types of emulator.

As shown, for example, by Santner et al. (2003), an empirical best linear unbiased predictor (EBLUP) of $y(\mathbf{x}^*)$ at input \mathbf{x}^* based on (1.1) with training data $\{\mathbf{x}_i = (x_{i1}, \dots, x_{id}), y(\mathbf{x}_i)\}_{i=1}^n$ has the form

$$\hat{y}(\mathbf{x}^*) = d_0(\mathbf{x}^*) + \sum_{i=1}^n d_i \prod_{j=1}^d R(x_j^* - x_{ij} \mid \hat{\psi}_j) \quad (3.5)$$

where $\hat{\psi}_j$ is a REML (or other) estimate of the unknown correlation parameter (vector) ψ_j , and where $d_0(\mathbf{x}^*) = \mathbf{f}^\top(\mathbf{x}^*)\hat{\boldsymbol{\beta}}$, with $\hat{\boldsymbol{\beta}} = (\mathbf{F}^\top \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^\top \mathbf{R}^{-1} \mathbf{y}(\mathbf{x})$ being the weighted least squares estimator of $\boldsymbol{\beta}$, $\mathbf{F} = [\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_n)]^\top$ is the matrix of regression functions for the training data, \mathbf{R} is the matrix with $(i, k)^{th}$ element $\prod_{j=1}^d R(x_{ij} - x_{kj} \mid \hat{\psi}_j)$, and $\mathbf{y}(\mathbf{x})$ is the $n \times 1$ vector of output training data; d_i is the i^{th} element of the vector $\mathbf{R}^{-1}(\mathbf{y}(\mathbf{x}) - \mathbf{F}\hat{\boldsymbol{\beta}})$. The expression (3.5) with parameter values replacing point estimators is the conditional predictor of $y(\mathbf{x}^*)$ given $(\boldsymbol{\beta}, \boldsymbol{\psi})$.

As an illustration of the method of calculation, consider the quadrature-based estimator of the total variance v in (2.7), when $Y(\mathbf{x})$ has regression mean

$$\mathbf{f}^\top(\mathbf{x})\boldsymbol{\beta} = \sum_{k_1=0}^{m_{k_1}} \dots \sum_{k_d=0}^{m_{k_d}} \beta_{k_1 \dots k_d} \prod_{j=1}^d x_j^{k_j} \quad (3.6)$$

for integers $m_{k_j} \geq 0$, $j = 1, \dots, d$, and has arbitrary but separable correlation function of the form (3.1).

Using (3.5), the first term of $v = E_g [y^2(\mathbf{X})] - (y_0)^2$ in (2.7) is estimated by

$$\begin{aligned} E_g [\hat{y}^2(\mathbf{X})] &= \int_0^1 \dots \int_0^1 \left[d_0(\mathbf{x}^*) + \sum_{i=1}^n d_i \prod_{j=1}^d R(x_j^* - x_{ij} \mid \hat{\psi}_j) \right]^2 \prod_{j=1}^d g_j(x_j^*) dx_j^* \\ &= \int_0^1 \dots \int_0^1 \left[d_0^2(\mathbf{x}^*) + 2 d_0(\mathbf{x}^*) \sum_{i=1}^n d_i \prod_{j=1}^d R(x_j^* - x_{ij} \mid \hat{\psi}_j) \right. \\ &\quad + \sum_{i=1}^n d_i^2 \prod_{j=1}^d R^2(x_j^* - x_{ij} \mid \hat{\psi}_j) \\ &\quad \left. + 2 \sum_{1 \leq i < k \leq n} d_i d_k \prod_{j=1}^d R(x_j^* - x_{ij} \mid \hat{\psi}_j) R(x_j^* - x_{kj} \mid \hat{\psi}_j) \right] \prod_{j=1}^d g_j(x_j^*) dx_j^*, \end{aligned} \quad (3.7)$$

and each component in (3.7) can be expressed as a product of one-dimensional integrals, as

follows. First, for $Y(\mathbf{x})$ having mean (3.6), the first term of (3.7) is

$$\int_0^1 \cdots \int_0^1 d_0^2(\mathbf{x}^*) \prod_{j=1}^d g_j(x_j^*) dx_j^* = \int_0^1 \cdots \int_0^1 \left[\sum_{k_1=0}^{m_{k_1}} \cdots \sum_{k_d=0}^{m_{k_d}} \widehat{\beta}_{k_1 \dots k_d} \prod_{j=1}^d (x_j^*)^{k_j} \right]^2 \prod_{j=1}^d g_j(x_j^*) dx_j^*, \quad (3.8)$$

and if $g_j(\cdot)$ is uniform, (3.8) reduces to

$$\sum_{k_1=0}^{m_{k_1}} \cdots \sum_{k_d=0}^{m_{k_d}} \sum_{k'_1=0}^{m_{k_1}} \cdots \sum_{k'_d=0}^{m_{k_d}} \widehat{\beta}_{k_1 \dots k_d} \widehat{\beta}_{k'_1 \dots k'_d} \prod_{j=1}^d (k_j + k'_j + 1)^{-1} = \mathbf{m}2_{\{1, \dots, d\}}(\widehat{\beta}). \quad (3.9)$$

The i th component of the second term in (3.7) involves the integral

$$\begin{aligned} & \int_0^1 \cdots \int_0^1 d_0(\mathbf{x}^*) \prod_{j=1}^d R(x_j^* - x_{ij} \mid \widehat{\psi}_j) \prod_{j=1}^d g_j(x_j^*) dx_j^* \\ &= \int_0^1 \cdots \int_0^1 \sum_{k_1=0}^{m_{k_1}} \cdots \sum_{k_d=0}^{m_{k_d}} \widehat{\beta}_{k_1 \dots k_d} \prod_{j=1}^d (x_j^*)^{k_j} R(x_j^* - x_{ij} \mid \widehat{\psi}_j) g_j(x_j^*) dx_j^* \\ &= \sum_{k_1=0}^{m_{k_1}} \cdots \sum_{k_d=0}^{m_{k_d}} \widehat{\beta}_{k_1 \dots k_d} \prod_{j=1}^d \mathbf{S1}_{k_j}(x_{ij}; \widehat{\psi}_j) \end{aligned} \quad (3.10)$$

where

$$\mathbf{S1}_{k_j}(x_{ij}; \widehat{\psi}_j) = \int_0^1 (x_j^*)^{k_j} R(x_j^* - x_{ij} \mid \widehat{\psi}_j) g_j(x_j^*) dx_j^*, \quad (3.11)$$

where $\mathbf{S1}_{k_j}(\cdot; \cdot)$ denotes an integral over a single variable with integrand involving *one* $R(\cdot \mid \cdot)$ term. The third and fourth terms in (3.7) can be expressed as a product of one-dimensional integrals involving

$$\begin{aligned} & \int_0^1 \cdots \int_0^1 \prod_{j=1}^d R(x_j^* - x_{ij} \mid \widehat{\psi}_j) R(x_j^* - x_{kj} \mid \widehat{\psi}_j) \prod_{j=1}^d g_j(x_j^*) dx_j^* \\ &= \prod_{j=1}^d \int_0^1 R(x_j^* - x_{ij} \mid \widehat{\psi}_j) R(x_j^* - x_{kj} \mid \widehat{\psi}_j) g_j(x_j^*) dx_j^* = \prod_{j=1}^d \mathbf{S2}(x_{ij}, x_{kj}; \widehat{\psi}_j), \end{aligned} \quad (3.12)$$

for $1 \leq i \leq k \leq d$. Here $\mathbf{S2}(\cdot, \cdot; \cdot)$ denotes an integral over a single variable with integrand involving *two* $R(\cdot \mid \cdot)$ terms.

Using (2.3) and (3.5), an estimate of the overall mean y_0 is

$$\widehat{y}_0 = \mathbf{m1}(\widehat{\beta}) + \sum_{i=1}^n d_i \prod_{j=1}^d \mathbf{S1}_0(x_{ij}; \widehat{\psi}_j) \quad (3.13)$$

where

$$\mathbf{S1}_0(x_{ij}; \hat{\psi}_j) = \int_0^1 R(x_j^* - x_{ij} \mid \hat{\psi}_j) g_j(x_j^*) dx_j^*, \quad (3.14)$$

and

$$\mathbf{m1}(\hat{\beta}) = \int_0^1 \cdots \int_0^1 \sum_{k_1=0}^{m_{k_1}} \cdots \sum_{k_d=0}^{m_{k_d}} \hat{\beta}_{k_1 \dots k_d} \prod_{j=1}^d (x_j^*)^{k_j} g_j(x_j^*) dx_j^*.$$

If $g_j(\cdot)$ is uniform, $j = 1, \dots, d$, then d_0 becomes

$$\mathbf{m1}(\hat{\beta}) = \sum_{k_1=0}^{m_{k_1}} \cdots \sum_{k_d=0}^{m_{k_d}} \hat{\beta}_{k_1 \dots k_d} \prod_{j=1}^d (k_j + 1)^{-1}. \quad (3.15)$$

Combining $E_g[\hat{y}^2(\mathbf{X})]$ from (3.7) and \hat{y}_0 from (3.13) gives an estimate for v . Corresponding formulae for calculation of estimates for v_j and v_{-j} are given in the Supplementary Material.

3.1 Formulae using the Gaussian correlation function

For some correlation functions and the choice of $g_j(x_j^*)$ being uniform, $j = 1, \dots, d$, the integrals $\mathbf{S1}_{k_j}(\cdot; \cdot)$, $\mathbf{S2}(\cdot, \cdot; \cdot)$ and $\mathbf{S1}_0(\cdot; \cdot)$ in (3.11), (3.12), and (3.14), respectively, can be expressed in closed form. For example, for the Gaussian correlation (3.2), $\mathbf{S2}(x_{ij}, x_{kj}; \hat{\psi}_j)$ is

$$\exp\{-\hat{\psi}_j(x_{ij} - x_{kj})/2\} \mathbf{S1}_0((x_{ij} + x_{kj})/2, \hat{\psi}_j) \quad (3.16)$$

where $\mathbf{S1}_0(x_{ij}, \hat{\psi}_j)$ is

$$\frac{\sqrt{\pi}}{\sqrt{\hat{\psi}_j}} \left\{ \Phi\left(\sqrt{2\hat{\psi}_j}(1 - x_{ij})\right) - \Phi\left(\sqrt{2\hat{\psi}_j}(0 - x_{ij})\right) \right\}, \quad (3.17)$$

cf. Chen et al. (2005). The integral $\mathbf{S1}_{k_j}(x_{ij}; \hat{\psi}_j)$ is

$$\begin{aligned} \frac{\sqrt{\pi}}{\sqrt{\hat{\psi}_j}} \left\{ \Phi\left(\frac{1 - x_{ij}}{\sqrt{1/(2\hat{\psi}_j)}}\right) \sum_{r=0}^{k_j} \binom{k_j}{r} x_{ij}^{k_j-r} (2\hat{\psi}_j)^{-r/2} I_r^{h_1} \right. \\ \left. - \Phi\left(\frac{-x_{ij}}{\sqrt{1/(2\hat{\psi}_j)}}\right) \sum_{r=0}^{k_j} \binom{k_j}{r} x_{ij}^{k_j-r} (2\hat{\psi}_j)^{-r/2} I_r^{h_0} \right\} \end{aligned} \quad (3.18)$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, $h_0 = -x_{ij}/\sqrt{1/(2\hat{\psi}_j)}$, $h_1 = (1 - x_{ij})/\sqrt{1/(2\hat{\psi}_j)}$, and I_r^h is defined recursively by $I_0^h = 1$, $I_1^h = -\phi(h)/\Phi(h)$, and

$$I_r^h = \frac{1}{\Phi(h)} \left[-h^{r-1} \phi(h) + (r-1) I_{r-2}^h \right], \quad r \in \{2, 3, \dots\},$$

where $\phi(\cdot)$ denotes the probability density function of the standard normal distribution.

3.2 Formulae using the Bohman correlation function

For the Bohman correlation function (3.3), $\mathbf{S2}(x_{ij}, x_{kj}; \hat{\psi}_j)$ depends on the relationships among x_{ij} , x_{kj} , and $\hat{\psi}_j$ and can be written as shown in (3.19) with ψ_j replaced by $\hat{\psi}_j$, and with $m_l = \text{med}(x_{ij}, 0, x_{kj} - \hat{\psi}_j)$ and $m_u = \text{med}(x_{kj}, 1, x_{ij} + \hat{\psi}_j)$:

$$\begin{aligned}
& \mathbf{1}_{[|x_{ij}-x_{kj}|<2\psi_j]} \times \left\{ \int_{m_l}^{x_{ij}} \left[\left(1 - \frac{x_{ij} - x_j^*}{\psi_j}\right) \cos\left(\frac{\pi(x_{ij} - x_j^*)}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_{ij} - x_j^*)}{\psi_j}\right) \right] \right. \\
& \quad \left[\left(1 - \frac{x_{kj} - x_j^*}{\psi_j}\right) \cos\left(\frac{\pi(x_{kj} - x_j^*)}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_{kj} - x_j^*)}{\psi_j}\right) \right] dx_j^* \\
& \quad + \int_{x_{ij}}^{x_{kj}} \left[\left(1 - \frac{x_j^* - x_{ij}}{\psi_j}\right) \cos\left(\frac{\pi(x_j^* - x_{ij})}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_j^* - x_{ij})}{\psi_j}\right) \right] \\
& \quad \left[\left(1 - \frac{x_{kj} - x_j^*}{\psi_j}\right) \cos\left(\frac{\pi(x_{kj} - x_j^*)}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_{kj} - x_j^*)}{\psi_j}\right) \right] dx_j^* \\
& \quad + \int_{x_{kj}}^{m_u} \left[\left(1 - \frac{x_j^* - x_{ij}}{\psi_j}\right) \cos\left(\frac{\pi(x_j^* - x_{ij})}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_j^* - x_{ij})}{\psi_j}\right) \right] \\
& \quad \left. \left[\left(1 - \frac{x_j^* - x_{kj}}{\psi_j}\right) \cos\left(\frac{\pi(x_j^* - x_{kj})}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_j^* - x_{kj})}{\psi_j}\right) \right] dx_j^* \right\} \quad (3.19)
\end{aligned}$$

where $\mathbf{1}_E$ is the function which is 1 or 0 as E occurs or not. Further simplification of these integrals can be made but the expressions are lengthy and omitted here. For purposes of computer code implementation, the symbolic tool box in MATLAB, for example, can be used to provide code. The integral $\mathbf{S1}_{k_j}$ in (3.11) is

$$\begin{aligned}
& \int_0^1 (x_j^*)^{k_j} R_B(x_j^* - x_{ij} | \psi_j) dx_j^* = \\
& \quad \int_{l^*}^{x_{ij}} (x_j^*)^{k_j} \left\{ \left(1 - \frac{x_{ij} - x_j^*}{\psi_j}\right) \cos\left(\frac{\pi(x_{ij} - x_j^*)}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_{ij} - x_j^*)}{\psi_j}\right) \right\} dx_j^* \\
& \quad + \int_{x_{ij}}^{u^*} (x_j^*)^{k_j} \left\{ \left(1 - \frac{x_j^* - x_{ij}}{\psi_j}\right) \cos\left(\frac{\pi(x_j^* - x_{ij})}{\psi_j}\right) + \frac{1}{\pi} \sin\left(\frac{\pi(x_j^* - x_{ij})}{\psi_j}\right) \right\} dx_j^* \quad (3.20)
\end{aligned}$$

where $l^* = \max(0, x_{ij} - \psi_j)$ and $u^* = \min(1, x_{ij} + \psi_j)$, and ψ_j is replaced by $\hat{\psi}_j$. Equation (3.20) can be simplified using formulas for the sine and cosine of the difference of two angles.

The Supplementary Material supplies further (but lengthy) simplifications of $\mathbf{S}\mathbf{1}_{kj}$.

Lastly, $\mathbf{S}\mathbf{1}_0(x_{ij}; \hat{\psi}_j)$ can be written as

$$\left\{ \frac{4\psi_j}{\pi^2} - \frac{2\psi_j}{\pi^2} \cos(l^{**}(x_{ij})) - \frac{2\psi_j}{\pi^2} \cos(u^{**}(x_{ij})) \right\} + \left\{ \left(\frac{\psi_j}{\pi} - \frac{\psi_j l^{**}(x_{ij})}{\pi^2} \right) \sin(l^{**}(x_{ij})) + \left(\frac{\psi_j}{\pi} - \frac{\psi_j u^{**}(x_{ij})}{\pi^2} \right) \sin(u^{**}(x_{ij})) \right\}, \quad (3.21)$$

with $l^{**}(x_{ij}) = ((x_{ij} - \max(0, x_{ij} - \psi_j))\pi) / \psi_j$, and $u^{**}(x_{ij}) = ((\min(1, x_{ij} + \psi_j) - x_{ij})\pi) / \psi_j$, and ψ_j replaced by $\hat{\psi}_j$. The Supplementary Material gives additional details of these calculations and the corresponding expressions for the cubic correlation function. It also shows how the terms v_j^u and v_{-j}^u in (2.10) and (2.11) can be calculated.

4 Process-based Estimation of Global Sensitivity Indices

This section presents Bayesian and empirical (plug-in) Bayesian estimates of main effect and total sensitivity indices when the true simulator output $y(\mathbf{x})$ can be modeled as a draw from a (smooth) Gaussian stochastic process, $Y(\mathbf{x})$, that has polynomial mean (3.6) and separable covariance function (3.1). To allow a greater breadth of applications, this section allows the *observed output* from the simulator at \mathbf{x} , say $z_{sim}(\mathbf{x})$, be the true simulator value $y(\mathbf{x})$ plus noise, for example, numerical noise. The model for $z_{sim}(\mathbf{x})$ used throughout is

$$Z_{sim}(\mathbf{x}) = Y(\mathbf{x}) + \epsilon_{sim}(\mathbf{x}), \quad (4.1)$$

where $\epsilon_{sim}(\mathbf{x})$ is an independent white noise process with mean zero and variance σ_ϵ . The term $\epsilon_{sim}(\mathbf{x})$ can be thought of as modeling non-deterministic computer output or of enhancing numerical stability in the estimation of the correlation parameters. For truly deterministic outputs, σ_ϵ can be set to zero in the formulae below. Here, and below, $E_p[\cdot]$ and $Cov_p[\cdot, \cdot]$ denote expectation and covariance with respect to the process to distinguish them from expectations with respect \mathbf{X} which are denoted by $E_g[\cdot]$.

Assuming simulator evaluations are made at input sites $\mathbf{x}_1, \dots, \mathbf{x}_n$, the $n \times 1$ vector of observed outputs is viewed as a realization of the stochastic process

$$\mathbf{Z}_{sim} = (Z_{sim}(\mathbf{x}_1), \dots, Z_{sim}(\mathbf{x}_n))^T$$

which has mean vector $\mathbf{F}\boldsymbol{\beta}$ with $\mathbf{F} = [\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_n)]^\top$ and polynomial $\mathbf{f}^\top(\mathbf{x})\boldsymbol{\beta}$ as in (3.6), and covariance matrix

$$\boldsymbol{\Sigma}_{sim}^Z = \sigma^2 \mathbf{R} + \sigma_\epsilon^2 \mathbf{I}_n = \sigma^2 (\mathbf{R} + a \mathbf{I}_n)$$

with $a = \sigma_\epsilon^2/\sigma^2$, where the $(i, k)^{th}$ element of the $n \times n$ matrix \mathbf{R} is of the form (3.1) and \mathbf{I} is the $n \times n$ identity matrix.

To simplify the expressions derived below, the following additional assumption is made that the weight function $g(\cdot)$ is uniform on $[0, 1]^d$. However, as in Section 3, any weight function with independent components can be used.

4.1 Bayesian and Empirical Bayesian estimators of sensitivity indices

By (2.5), (2.10), and (2.12), the j^{th} main effect sensitivity index, S_j , and total effect sensitivity index, T_j , can be expressed in terms of v_Q^u for $Q = \{j\}$, and $\{1, \dots, d\}$, and $\{-j\} = \{1, \dots, j-1, j+1, \dots, d\}$. Bayesian and empirical Bayesian estimation of v_Q^u firstly replaces $y(\cdot)$ by the process $Y(\cdot)$ yielding a random variable

$$V_Q^u = Var_g [E_g[Y(\mathbf{X})|\mathbf{X}_Q]] .$$

Then the calculations below give the $E_P\{\cdot\}$ expectation of V_Q^u given the observed data $\mathbf{Z}_{sim} = \mathbf{z}_{sim}$ and the GP model parameters, say $\boldsymbol{\xi} = (\boldsymbol{\beta}, \sigma^2, a, \boldsymbol{\psi})$, i.e.,

$$E_P \left\{ V_Q^u \mid \mathbf{Z}_{sim} = \mathbf{z}_{sim}, \boldsymbol{\xi} \right\} . \quad (4.2)$$

Empirical Bayesian estimators of S_j and T_j are obtained by plugging an estimate (for example, MLE or REML) of $\boldsymbol{\xi}$ into (4.2). If prior information about the values of $\boldsymbol{\xi}$ is available in the form of a distribution $[\boldsymbol{\xi}]$, then fully Bayesian estimators can be obtained as

$$E_P \left\{ V_Q^u \mid \mathbf{Z}_{sim} = \mathbf{z}_{sim} \right\} = E_{[\boldsymbol{\xi}|\mathbf{Z}_{sim}]} \left\{ E_P \left\{ V_Q^u \mid \mathbf{Z}_{sim} = \mathbf{z}_{sim}, \boldsymbol{\xi} \right\} \right\} , \quad (4.3)$$

which is (4.2) weighted by draws from the posterior of the parameters given the data.

A formula for (4.2) is presented in the following theorem which uses $\mathbf{S1}_{k_j}(x_{ij}; \psi_j)$, $\mathbf{S2}(x_{ij}, x_{kj}; \psi_j)$, $\mathbf{S1}_0(x_{kj}; \psi_j)$, and $\mathbf{m1}(\boldsymbol{\beta})$ in (3.11), (3.12), (3.14), and (3.15), respectively, with $\hat{\psi}_j$ and $\hat{\boldsymbol{\beta}}$

replaced by ψ_j and $\boldsymbol{\beta}$, and

$$\begin{aligned} \mathbf{D}(\psi_j) &= \int_0^1 \int_0^1 R(w-x | \psi_j) dx dw, \\ \mathbf{m2}_Q(\boldsymbol{\beta}) &= \sum_{k_1, \dots, k_d} \sum_{k'_1, \dots, k'_d} \beta_{k_1, \dots, k_d} \beta_{k'_1, \dots, k'_d} \\ &\quad \times \left[\prod_{j \notin Q} (k_j + 1) (k'_j + 1) \right]^{-1} \left[\prod_{\ell \in Q} (k_\ell + k'_\ell + 1) \right]^{-1}. \end{aligned}$$

Theorem 1 Assume that the true simulator output, $y(\mathbf{x})$ can be modeled by a Gaussian process $Y(\cdot)$ with mean and covariance function of the form (3.6) and (3.1), respectively. Also assume that the observed output \mathbf{z}_{sim} at the training data sites, is modeled by a process $Z_{sim}(\mathbf{x})$ satisfying (4.1). For a fixed $Q \subseteq \{1, \dots, d\}$,

$$\begin{aligned} \widehat{v}_Q^u(\boldsymbol{\xi}) &= E_P \left\{ V_Q^u | \mathbf{Z}_{sim} = \mathbf{z}_{sim}, \boldsymbol{\xi} \right\} \\ &= \left\{ \sigma_Y^2 \prod_{j \notin Q} \mathbf{D}(\psi_j) - \text{trace} \left[\left(\boldsymbol{\Sigma}_{sim}^Z \right)^{-1} \mathbf{C} \right] \right\} \\ &\quad + \left\{ \mathbf{m2}_Q(\boldsymbol{\beta}) - \mathbf{m1}^2(\boldsymbol{\beta}) + 2 \left(\mathbf{t}^\top - \mathbf{m1}(\boldsymbol{\beta}) \mathbf{q}^\top \right) \left(\boldsymbol{\Sigma}_{sim}^Z \right)^{-1} \left(\mathbf{z}_{sim} - \mathbf{F}^\top \boldsymbol{\beta} \right) \right. \\ &\quad + \left. \left(\mathbf{z}_{sim} - \mathbf{F}^\top \boldsymbol{\beta} \right)^\top \left(\boldsymbol{\Sigma}_{sim}^Z \right)^{-1} \left(\mathbf{C} - \mathbf{q} \mathbf{q}^\top \right) \left(\boldsymbol{\Sigma}_{sim}^Z \right)^{-1} \left(\mathbf{z}_{sim} - \mathbf{F}^\top \boldsymbol{\beta} \right) \right\} \\ &\quad - \left\{ \sigma_Y^2 \prod_{j=1}^d \mathbf{D}(\psi_j) - \text{trace} \left[\left(\boldsymbol{\Sigma}_{sim}^Z \right)^{-1} \mathbf{q} \mathbf{q}^\top \right] \right\}, \end{aligned} \tag{4.4}$$

where \mathbf{q} is the $n \times 1$ vector with i^{th} element

$$q_i = q(\mathbf{x}_i, \boldsymbol{\psi}) = \sigma^2 \prod_{j=1}^d \mathbf{S1}_0(x_{ij}; \psi_j), \quad 1 \leq i \leq n,$$

\mathbf{C} is the $n \times n$ matrix with $(i, k)^{\text{th}}$ element

$$C_{ik} = \sigma^4 \prod_{j \notin Q} \mathbf{S1}_0(x_{ij}; \psi_j) \mathbf{S1}_0(x_{kj}; \psi_j) \prod_{j \in Q} \mathbf{S2}(x_{ij}, x_{kj}; \psi_j), \quad 1 \leq i, k \leq n,$$

\mathbf{t} is the $n \times 1$ vector with i^{th} element

$$\begin{aligned} \mathbf{t}(\mathbf{x}_i, \boldsymbol{\beta}, \psi, \boldsymbol{\beta}) &= \left[\sigma^2 \prod_{j \notin Q} \mathbf{S1}_0(x_{ij}; \psi_j) \right] \\ &\times \sum_{k_1=0}^{m_{k_1}} \dots \sum_{k_d=0}^{m_{k_d}} \left\{ \beta_{k_1 \dots k_d} \prod_{j \notin Q} (k_j + 1)^{-1} \prod_{\ell \in Q} \mathbf{S1}_{k_\ell}(x_{h\ell}; \boldsymbol{\psi}_\ell) \right\}, \quad 1 \leq i \leq n, \end{aligned}$$

Proof. The proof of Theorem 1 involves three steps: (i) the derivation of the conditional distribution of the process $U_Q(\mathbf{x}_Q) \equiv E_g[Y(\mathbf{X}) | \mathbf{X}_Q = \mathbf{x}_Q]$ given $\boldsymbol{\xi}$; (ii) the determination of the conditional distribution of $[U_Q(\mathbf{x}_Q) | \mathbf{Z}_{sim}, \boldsymbol{\xi}]$; and (iii) the derivation of a formula for $E_P[Var_g(U_Q(\mathbf{x}_Q)) | \mathbf{Z}_{sim}, \boldsymbol{\xi}]$. The details are given in the Supplementary Material.

An example of a prior used for fully Bayes estimation of sensitivity indices for the zero mean, Gaussian correlation model are given by Higdon, Gattiker, Williams, and Rightley (2008). As an example of Empirical Bayes estimation, S_j and T_j are computed using the facts that the estimate $\hat{v}(\hat{\boldsymbol{\xi}})$ of the total variance v is given by (4.4) for $Q = \{1, \dots, d\}$. The main effect sensitivity index S_j in (2.5) with $Q = \{j\}$ for the individual input x_j is estimated by $\hat{S}_j = \hat{v}_j^u(\hat{\boldsymbol{\xi}}) / \hat{v}(\hat{\boldsymbol{\xi}})$, where $\hat{v}_j^u(\hat{\boldsymbol{\xi}})$ is obtained from (4.4). The total effect sensitivity index is estimated by $\hat{T}_j = (\hat{v}(\hat{\boldsymbol{\xi}}) - \hat{v}_{-j}^u(\hat{\boldsymbol{\xi}})) / \hat{v}(\hat{\boldsymbol{\xi}})$, where $\hat{v}_{-j}^u(\hat{\boldsymbol{\xi}})$ is obtained from (4.4) with $Q = \{1, \dots, i-1, i+1, \dots, d\}$.

Given the model parameters, all components of \hat{v}_Q^u are specified above except the integrals $\mathbf{S1}_k$, \mathbf{D} , $\mathbf{S2}$, which depend on the user-selected correlation function. Formulas for these integrals are stated next for the Gaussian and Bohman correlation functions and, in the Supplementary Material, for the cubic correlation function $R_C(w-x | \psi)$ in (3.4) for $\psi > 0$.

4.2 Formulae for the Gaussian correlation function

The integrals $\mathbf{S2}(x_{ij}, x_{kj}; \psi_j)$, $\mathbf{S1}_0(x_{kj}; \psi_j)$, and $\mathbf{S1}_{k_j}(x_{ij}; \psi_j)$, were given for the Gaussian correlation function (3.2) in (3.16), (3.17), and (3.18) respectively, with $\hat{\psi}_j$ and $\hat{\beta}$ replaced by ψ_j and β . In addition, the integral $\mathbf{D}(\psi_j)$ is

$$\begin{aligned} \mathbf{D}(\psi_j) &= \int_0^1 \int_0^1 \exp[-\psi_j(w-x)^2] dx dw \\ &= \frac{1}{\psi_j} \left[\sqrt{2\pi} \phi(\sqrt{2\psi_j}) - 1 \right] + \sqrt{\frac{\pi}{\psi_j}} \left[2\Phi(\sqrt{2\psi_j}) - 1 \right]. \end{aligned}$$

4.3 Formulae using the Bohman correlation function

For the Bohman correlation function, $R_B(w - x | \psi_j)$ in (3.3) the integrals $\mathbf{S2}$, $\mathbf{S1}_{k_j}$, $\mathbf{S1}_0$, are as given in (3.19), (3.20), and (3.21), respectively. For $\psi_j > 0$, the integral for $\mathbf{D}(\psi_j)$ is defined piecewise by

$$\mathbf{D}(\psi_j) = \begin{cases} \frac{4\psi_j}{\pi^2} + \frac{2\psi_j^2}{\pi^2} - \frac{4\psi_j}{\pi^2} (\psi_j - 1.0), & 0 < \psi_j < 1.0 \\ \frac{4\psi_j}{\pi^2} + \frac{2\psi_j^2}{\pi^2} \left\{ 1 + \left(\frac{1.0 - \psi_j}{\psi_j} \right) \cos\left(\frac{\pi}{\psi_j}\right) - \frac{3}{\pi} \sin\left(\frac{\pi}{\psi_j}\right) \right\}, & 1.0 \leq \psi_j. \end{cases}$$

5 Sparse correlation matrices

The motivation for providing sensitivity index estimators for compactly supported correlation functions is that they can provide sparse correlation matrices which makes their inversion numerically more stable (see Barry and Pace (1997) and, in MATLAB, Gilbert, Moler, and Schreiber (1991)) and allows the analysis of larger training data sets than using the widely-used Gaussian correlation function. Indeed, Kaufman, Bingham, Habib, Heitmann, and Frieman (2010) demonstrated that, using a suitably rich regression mean with such a sparse correlation matrix, the predictive ability of the stochastic model is comparable to that based on a model with Gaussian correlation function when both correlation functions can be implemented.

Let $\boldsymbol{\psi} = (\psi_1, \dots, \psi_d)$ denote the parameter vector for the Bohman or cubic correlation function. Kaufman et al. (2010) proposed enforcing sparsity in the matrix of correlations by restricting attention to a parameter space of the form

$$\Omega(K) = \left\{ \boldsymbol{\psi} \in \mathbb{R}^d : \psi_j \geq 0 \forall j \in \{1, \dots, d\}; \sum_{j=1}^d \psi_j \leq K \right\}, \quad (5.1)$$

where $K > 0$ is chosen so that at least a given proportion α of the $n(n-1)/2$ off diagonal elements of $\boldsymbol{\Sigma}_{sim}^Z$ are zero. We note that one method of selecting K to force at least a proportion α of zeroes among the off-diagonal elements of $\boldsymbol{\Sigma}_{sim}^Z$ is as follows. Calculate $d_{i,k}^1 \equiv \sum_{j=1}^d |x_{ij} - x_{kj}|$ for each of the $(\mathbf{x}_i, \mathbf{x}_k)$ pairs with $1 \leq i < k \leq n$. Then, set K to be the $\lfloor \binom{n}{2} \times \alpha \rfloor^{th}$ smallest value among the $d_{i,k}^1$'s where $\lfloor \cdot \rfloor$ denotes the integer part of $\binom{n}{2} \times \alpha$. It follows that, for any $\boldsymbol{\psi} \in \Omega(K)$, at most $\alpha 100\%$ of the off-diagonal elements of $\boldsymbol{\Sigma}_{sim}^Z$ are

nonzero. To see that this is true for the Bohman and cubic correlation functions, first note that $R(\mathbf{x}_i, \mathbf{x}_k | \psi_j) = 0$ for each as long as $|x_{ij} - x_{kj}| \geq \psi_j$ for some $j \in \{1, \dots, d\}$. Now select any $(\mathbf{x}_i, \mathbf{x}_k)$ with $d_{i,k}^1 \geq K$; there are at least $(1 - \alpha) \times 100\%$ such pairs among the $\binom{n}{2}$ pairings of rows. To see that

$$\prod_{j=1}^d R(x_{ij}, x_{kj} | \psi_j) = 0 \quad (5.2)$$

for any $\boldsymbol{\psi} \in \Omega(K)$, assume instead that (5.2) is positive, for some $\boldsymbol{\psi} \in \Omega(K)$. Then $|x_{ij} - x_{kj}| < \psi_j$ for all $j \in \{1, \dots, d\}$. Hence $d_{i,k}^1 \equiv \sum_{j=1}^d |x_{ij} - x_{kj}| < \sum_{j=1}^d \psi_j \leq K$ where the last inequality holds because $\boldsymbol{\psi} \in \Omega(K)$. But this contradicts the assumption that $d_{i,k}^1 \geq K$ and hence $|x_{ij} - x_{kj}| \geq \psi_j$ for some j and hence (5.2) holds.

6 Two Examples

Two examples are given below to compare the results of applying the estimation methods described in this paper. The first example, which uses a relatively small sample size, is the Sobol'-Levitan function introduced in Sobol' and Levitan (1999). The second example uses a closed-form "synthetic" function which Oakley and OHagan (2004) present with $n = 250$ function evaluations to illustrate their fully Bayesian sensitivity index (SI) calculations.

6.1 Sensitivity Indices for the Sobol'-Levitan function

This example uses a scaled version of the function

$$y(x_1, \dots, x_d) = \exp\left(\sum_{j=1}^d b_j x_j\right) - I_d, \quad \mathbf{x} \in [0, 1]^d, \quad (6.1)$$

introduced in Sobol' and Levitan (1999) where $I_d = \prod_{j=1}^d (e^{b_j} - 1) / b_j$. Analytical formulas for the main effect and total effect SIs are known for $y(\mathbf{x})$ for any d , any $\mathbf{b} = (b_1, \dots, b_d)$, and uniform weight on each input; the SIs are the same for any scaled version of $y(\mathbf{x})$. The $d = 8$ input Sobol'-Levitan function, scaled to have variance 100, is used as the response in the following calculations where \mathbf{b} is selected so that $y(\mathbf{x})$ has the $\{S_j\}_{j=1}^d$ and $\{T_j\}_{j=1}^d$ values shown in Table 6.1. This choice of \mathbf{b} produces an output function with substantial interactions because the sum of the main effect sensitivity indices is only 50% of the total $y(\mathbf{x})$ variance.

Table 6.1

Parameters, true main effect and total effect sensitivity indices for the Sobol'-Levitani $y(\mathbf{x})$ (6.1).

j	1	2	3	4	5	6	7	8
b_j	3.2797	2.6467	2.0606	1.5258	1.0476	0.6340	0.2973	0.0638
S_j	0.2003	0.1370	0.0863	0.0487	0.0234	0.0087	0.0019	0.0001
T_j	0.5477	0.4342	0.3136	0.1985	0.1039	0.0406	0.0092	0.0004

First, $n = 64$ function runs were used to estimate the SIs using seven estimators (permuted column sampling, a quadrature-based method, four different empirical Bayesian methods, and a fully Bayesian method). The permuted column sampling method was implemented using the permutations specified by the OA(64,9,8,2) design (from <http://www.research.att.com/~njas/oadir>) applied to the maximin Latin Hypercube (LH) base design scaled to its midpoints (<http://www.spacefillingdesigns.nl/>). The remaining estimators use $y(\mathbf{x})$ evaluations from \mathbf{x} that form a 64×8 maximin LH design obtained from <http://www.spacefillingdesigns.nl/> and scaled to include endpoints.

The quadrature-based method and the EB_G empirical Bayesian estimator assume a GP process with constant mean and the Gaussian correlation function. The remaining three empirical Bayesian estimators (EB_C , EB_{cC} , EB_B) are based on GP models having cubic means and either a Bohman and cubic correlation function. The EB_C and EB_B estimators are based on the REML estimate of ψ in the unconstrained parameter space. The estimator EB_{cC} is based on cubic correlation function with parameter space restricted by (5.1) to provide a sparse correlation function with no more than 15% nonzero off-diagonal correlations and REML estimation of the parameter. The quadrature-based estimates were computed using the Gaussian process modeling tool of the JMP software (JMP 2011). All the EB estimators were computed using MATLAB code written by the first and fourth authors of this paper. The fully Bayesian method assumes a GP with constant mean, the Gaussian correlation function, and the default prior specified by `gpmsa` (see Gattiker 2008).

The absolute error of estimation of the main effect SIs, S_j , $j = 1, \dots, 8$, are listed in Table 6.2 for the seven methods. The quadrature, EB, and fully Bayesian estimators provide comparable estimates of all main effect sensitivity indices with respect to the absolute error of estimation and appear to be the preferable methods. For all but S_2 , the permuted column sampling estimator is inferior to the other six estimators. This is in line with the findings of, for example, Oakley and OHagan (2004) and Chen et al. (2005) for small sample sizes.

Table 6.2

Errors estimating the *main effect* SIs of the Sobol'-Levitán function (6.1) with \mathbf{b} listed in Table 6.1 based on $n = 64$ function runs using (1) Permuted column sampling (PermS); (2) quadrature-based estimation (Quad), and (3) four empirical Bayes Estimators, and (4) a fully Bayesian method based on the Gaussian correlation function with prior described in `gpmsa` (see Gattiker (2008))

S_j	0.2003	0.1370	0.0863	0.0487	0.0234	0.0087	0.0019	0.0001
$ \widehat{S}_j^{PermS} - S_j $	0.0885	0.0368	0.0349	0.0175	0.0187	0.0253	0.0366	0.0043
$ \widehat{S}_j^{Quad} - S_j $	0.0525	0.0675	0.0235	0.0031	0.0028	0.0066	0.0051	0.0025
$ \widehat{S}_j^{EB_G} - S_j $	0.0474	0.0634	0.0213	0.0023	0.0032	0.0066	0.0051	0.0025
$ \widehat{S}_j^{EB_C} - S_j $	0.0505	0.0635	0.0262	0.0063	0.0014	0.0065	0.0063	0.0028
$ \widehat{S}_j^{EB_{cB}} - S_j $	0.0179	0.0140	0.0657	0.0242	0.0559	0.0649	0.0293	0.0411
$ \widehat{S}_j^{EB_B} - S_j $	0.0506	0.0633	0.0204	0.0073	0.0016	0.0074	0.0053	0.0022
$ \widehat{S}_j^{gpmsa} - S_j $	0.0447	0.0613	0.0249	0.0008	0.0036	0.0064	0.0056	0.0027

Comparing the six superior methods further, we see in Table 6.3 that, for inputs with large total effect, i.e., those with $T_j > 0.10$, aside from the parameter constrained EB_{cB} , the EB estimators and the fully Bayesian estimator have slightly smaller absolute error than the quadrature-based estimator. For T_2 , the fully Bayesian estimator appears to have superior performance. We increased the percentage of non-zero non-diagonal correlation elements to allow more data in the prediction process up to 50%. The errors of estimation of the total effect SIs remained substantially larger than those of the other EB estimators. Of course allowing all the data to be used in the estimation of ψ as for EB_B or EB_C does produce reasonable SI estimates. EB estimation based on a constrained parameter space suffers from the same type of estimation errors as does EB_{cB} . In sum, when the amount of data used to estimate the model parameters is too “small”, the EB estimators of the SIs can be severely negatively impacted.

We increased the number of function evaluations to 81, using the orthogonal array $OA(81, 10, 9, 2)$ from the website above for the permuted column sampling. All methods provided better estimates, but the relative performance remained the same. Consequently, it appears that, on the whole, the Bayesian methodology preforms slightly better than the other two methods for estimation of main effect and total effect sensitivity indices.

Table 6.3

Errors estimating the *total effect* SIs of the Sobol'-Levitan function (6.1) with \mathbf{b} listed in Table 6.1 based on $n = 64$ function runs using (1) quadrature-based estimation (Quad), and (2) three (empirical) Bayesian Estimation based on the Gaussian, Bohman, and cubic correlation functions, and (3) a fully Bayesian method based on the Gaussian correlation function with prior described in `gpmsa` (see Gattiker (2008))

T_j	0.5477	0.4342	0.3136	0.1985	0.1039	0.0406	0.0092	0.0004
$ \hat{T}_j^{Quad} - T_j $	0.0733	0.0342	0.1165	0.0798	0.0660	0.0020	0.0028	0.0036
$ \hat{T}_j^{EBG} - T_j $	0.0479	0.0080	0.0974	0.0656	0.0614	0.0068	0.0053	0.0045
$ \hat{T}_j^{EBc} - T_j $	0.0605	0.0096	0.0905	0.0672	0.0640	0.0065	0.0074	0.0040
$ \hat{T}_j^{EBcC} - T_j $	0.1771	0.0845	0.0204	0.0195	0.1670	0.1654	0.2361	0.2467
$ \hat{T}_j^{EBB} - T_j $	0.0491	0.0147	0.1044	0.0646	0.0621	0.0066	0.0055	0.0035
$ \hat{T}_j^{gpmsa} - T_j $	0.0585	0.0064	0.0780	0.0694	0.0624	0.0059	0.0122	0.0066

6.2 Estimating Sensitivity Indices in a Synthetic Example

Oakley and OHagan (2004) presented the $d = 15$ input example

$$y(\mathbf{x}) = \mathbf{a}_1^\top \mathbf{x} + \mathbf{a}_2^\top \sin(\mathbf{x}) + \mathbf{a}_3^\top \cos(\mathbf{x}) + \mathbf{x}^\top \mathbf{M} \mathbf{x} \quad (6.2)$$

with input space \mathbb{R}^{15} and independent $N(0, 1)$ weight functions $g_j(x_j)$ on each component. We use their function $y(\mathbf{x})$ but with input space $[0, 1]^{15}$ and independent uniform weight functions on each component. This function is smooth and was constructed to have five relatively inactive inputs (x_1 - x_5), five moderately active inputs (x_6 - x_{10}), and five very active inputs (x_{11} - x_{15}); $y(\mathbf{x})$ has numerous local maxima and minima.

As in Oakley and OHagan (2004), the estimates of main effect and total effect sensitivity indices were based on $n = 250$ evaluations of $y(\mathbf{x})$. The 250×15 input design was selected to be an (approximate) maximin Latin Hypercube which was computed using the genetic algorithm implemented in the software `bestlh.m` (Forrester, Sobester, and Keane (2008)).

Table 6.4 lists the estimated main effect and total effect sensitivity indices based on this data for two different empirical Bayesian (EB) methods and one fully Bayesian method. It also lists the total cpu time required to estimate parameters or make draws from the

posterior distribution and compute both main effect and total effect sensitivity indices for all 15 inputs (all computations performed on a Linux machine with two eight core Xeon E5-2680 processors running at 2.7GHz). The first EB method, denoted EB_G , is based on a GP prior for $y(\mathbf{x})$ with Gaussian correlation function and (the frequently-used choice of) constant mean. The second EB method, denoted EB_{cB} , is based on a GP with Bohman correlation function (3.3) and cubic mean

$$E_P\{Y(\mathbf{x})\} = \beta_0 + \sum_{j=1}^{15} x_j + \sum_{j=1}^{15} x_j^2 + \sum_{j=1}^{15} x_j^3. \quad (6.3)$$

For EB_{cB} , ψ parameters were restricted to the parameter space (5.1) where K was selected so that at most 15% of the off-diagonal entries of the correlation matrix were zero. Both EB methods use REML estimation of ψ . The EB SI estimates were computed using the same MATLAB code as for Sobol'-Levitin example and the fully Bayesian estimate was computed using `gpmsa`.

For this smooth function, $n = 250$ observations are more than sufficient to provide a good overall estimate (see Loepky, Sacks, and Welch (2009)). The results listed in Table 6.4 show that all three methods provide comparable estimates of both the main effect and total effect sensitivity indices. However the methods are distinguished by the times required to produce these estimates. The EB_{cB} method with sparse correlation matrix required only 1,167 seconds of cpu time to produce estimates similar to those given by the EB_G estimator while the latter required over 9,000 seconds of cpu time. This example provides another illustration suggesting that, even with a more complicated fitted mean, the GP model with a sparse correlation matrix induced by the Bohman correlation, can be used to estimate and predict in cases of larger n than is possible with the Gaussian correlation function with constant mean (see Kaufman et al. (2011)). In fact, an additional SI estimation based on GP with Gaussian correlation function model but with a cubic mean (6.3) (not shown) required 12,926 seconds of cpu time (and resulted in essentially equivalent numerical SI estimates). The fully Bayesian FB method required much longer, 46,380 seconds of cpu time, than either EB method to estimate the SIs; of this computational time, 41,167 seconds were used to construct 10,000 posterior draws of the parameters and an additional 5,213 seconds to evaluate the SI formulas for the 500 equally-spaced of the posterior draws selected from the 10,000 generated. We conclude by noting that estimates of the SIs based on a GP with cubic correlation and cubic mean produced similar numerical estimates to those shown in Table 6.4 and required 1,139 seconds of cpu time, a value comparable to the Bohman-based

Table 6.4

Estimated ME and TE Sensitivity Indices Using Empirical Bayesian and Full Bayesian Methods (EB_G denotes the Empirical Bayesian with Constant Mean and Gaussian Correlation), (EB_B denotes the Empirical Bayesian with Cubic Mean and Bohmans Correlation), (FB denotes the Fully Bayesian Estimator with Constant Mean and Gaussian Correlation as computed using `gpmsa`)

	S_j			T_j		
	EB_G	EB_{cB}	FB	EB_G	EB_{cB}	FB
x_1	0.0111	0.0114	0.0112	0.0146	0.0179	0.0130
x_2	0.0040	0.0044	0.0047	0.0065	0.0105	0.0066
x_3	0.0212	0.0202	0.0221	0.0228	0.0250	0.0230
x_4	0.0078	0.0078	0.0086	0.0103	0.0136	0.0102
x_5	0.0017	0.0014	0.0017	0.0029	0.0087	0.0023
x_6	0.0181	0.0196	0.0171	0.0199	0.0274	0.0176
x_7	0.0244	0.0263	0.0260	0.0261	0.0334	0.0270
x_8	0.0525	0.0559	0.0529	0.0554	0.0639	0.0547
x_9	0.0455	0.0467	0.0439	0.0479	0.0523	0.0454
x_{10}	0.0165	0.0140	0.0156	0.0179	0.0204	0.0162
x_{11}	0.1847	0.1942	0.1889	0.1883	0.1999	0.1904
x_{12}	0.1764	0.1759	0.1736	0.1792	0.1834	0.1740
x_{13}	0.1782	0.1741	0.1791	0.1815	0.1820	0.1803
x_{14}	0.1207	0.1163	0.1218	0.1237	0.1239	0.1229
x_{15}	0.1205	0.1236	0.1243	0.1238	0.1298	0.1253
cputime (s)	9,308	1,167	46,380	—	—	—

estimates. In sum, this example provides additional evidence of the value of using compactly supported correlation functions when the amount of the data is “large.”

7 Summary and Discussion

This paper presents estimation formulae for quadrature-based, Bayesian, and empirical Bayesian estimators of main effect and total effect sensitivity indices. The Bayesian estimator is a posterior mean of the variance of an averaged output value using a Gaussian process prior for the output function. Specific estimation formulae are given for a broad class of regression plus stationary Gaussian process priors models that are based on the Gaussian, Bohman or cubic correlation functions. For small n/d , all methods yield numer-

ically similar estimates. For larger n/d , estimators based on compactly supported Bohman and cubic correlation functions (when used with non-constant regression means) can provide significant cpu savings when the correlation matrix is constrained to be sparse and they are numerically similar to sensitivity index estimates based on non-compactly supported correlation functions. To facilitate the use of sparse correlation structures, we provide a method determining the parameter space that controls the degree of sparsity for the Bohman and cubic correlation functions.

Acknowledgments

This research was sponsored, in part, by the National Science Foundation under Agreement DMS-0806134 (The Ohio State University). This paper was completed at the Isaac Newton Institute for the Mathematical Sciences during the 2011 Programme on the Design and Analysis of Experiments. The authors would like to thank the referees for their helpful comments which have led to a considerably improved paper, and to thank Fangfang Sun for help with the some of the computations used in the paper.

References

- Barry, R. P. and Pace, R. K. (1997), “Kriging with large data sets using sparse matrix techniques,” *Communications in Statistics - Simulation and Computation*, 26, 619–629.
- Campolongo, F., Cariboni, J., and Saltelli, A. (2007), “An effective screening design for sensitivity analysis of large models,” *Environmental Modelling & Software*, 22, 1509–1518.
- Chen, W., Jin, R., and Sudjianto, A. (2005), “Analytical Variance-Based Global Sensitivity Analysis in Simulation-Based Design Under Uncertainty,” *Journal of Mechanical Design*, 127, 875–876.
- (2006), “Analytical Global Sensitivity Analysis and Uncertainty Propagation for Robust Design,” *Journal of Quality Technology*, 38, 333–348.
- Currin, C., Mitchell, T. J., Morris, M. D., and Ylvisaker, D. (1991), “Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments,” *Journal of the American Statistical Association*, 86, 953–963.
- Da Viegua, S., Wahl, F., and Gamboa, F. (2009), “Local polynomial estimation for sensitivity analysis on models with correlated inputs,” *Technometrics*, 51, 452–463.
- Forrester, A., Sobester, A., and Keane, A. (2008), *Engineering design via surrogate modelling: a practical guide*, Chichester, UK: Wiley.

- Gattiker, J. R. (2008), “Gaussian Process Models for Simulation Analysis (GPM/SA) Command, Function, and Data Structure Reference,” Tech. Rep. LA-UR-08-08057, Los Alamos National Laboratory.
- Gilbert, J. R., Moler, C., and Schreiber, R. (1991), “Sparse Matrices in MATLAB: Design and Implementation,” *SIAM J. Matrix Anal. Appl.*, 13, 333–356.
- Helton, J. C., Johnson, J. D. Sallaberry, C. J., and Storlie, C. B. (2006), “Survey of sampling-based methods for uncertainty and sensitivity analysis,” *Reliability Engineering and System Safety*, 91, 1175–1209.
- Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008), “Computer Model Calibration using High Dimensional Output,” *Journal of the American Statistical Association*, 103, 570–583.
- Homma, T. and Saltelli, A. (1996), “Importance measures in global sensitivity analysis of model output,” *Reliability Engineering and System Safety*, 52, 1–17.
- JMP (2011), “Version 9,” .
- Jones, D. R., Schonlau, M., and Welch, W. J. (1998), “Efficient global optimization of expensive black-box functions,” *Journal of Global Optimization*, 13, 455–492.
- Kaufman, C., Bingham, D., Habib, S., Heitmann, K., and Frieman, J. (2010), “Efficient Emulators of Computer Experiments Using Compactly Supported Correlation Functions,” *Submitted*.
- (2011), “Efficient Emulators of Computer Experiments Using Compactly Supported Correlation Functions, with an Application to Cosmology,” *The Annals of Applied Statistics*, 5, 24702492.
- Linkletter, C., Bingham, D., Hengartner, N., Higdon, D., and Ye, K. Q. (2006), “Variable Selection for Gaussian Process Models in Computer Experiments,” *Technometrics*, 48, 478–490.
- Loeppky, J. L., Sacks, J., and Welch, W. J. (2009), “Choosing the Sample Size of a Computer Experiment: A Practical Guide,” *Technometrics*, 51, 366–376.
- Marrel, A., Iooss, B., Lauren, B., and Roustant, O. (2009), “Calculations of Sobol indices for the Gaussian process metamodel,” *Reliability Engineering and System Safety*, 94, 742–751.
- Moon, H., Santner, T. J., and Dean, A. M. (2012), “Two-stage Sensitivity-based Group Screening in Computer Experiments,” *Technometrics*, 54, 376–387.
- Morris, M. D. (1991), “Factorial sampling plans for preliminary computational experiments,” *Technometrics*, 33, 161–174.
- Morris, M. D., Moore, L. M., and McKay, M. D. (2008), “Using Orthogonal Arrays in the Sensitivity Analysis of Computer Models,” *Technometrics*, 50, 205–215.

- Oakley, J. E. and O'Hagan, A. (2004), "Probabilistic sensitivity analysis of complex models: a Bayesian approach," *J. Royal Stat. Society B*, 66, 751–769.
- O'Hagan, A. (1992), "Some Bayesian numerical analysis," in *Bayesian Statistics*, eds. Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., Oxford University Press, vol. 4, pp. 345–363.
- Sacks, J., Schiller, S. B., and Welch, W. J. (1989a), "Design for computer experiments," *Technometrics*, 31, 41–47.
- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989b), "Design and analysis of computer experiments," *Statistical Science*, 4, 409–423.
- Saltelli, A. (2002), "Making best use of model evaluations to compute sensitivity indices," *Computer Physics Communications*, 145 (2), 280–297.
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010), "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," *Computer Physics Communications*, 181, 259–270.
- Saltelli, A., Tarantola, S., and Campolongo, F. (2000), "Sensitivity Analysis As an Ingredient of Modeling," *Statistical Science*, 15, 377–395.
- Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, New York: Springer Verlag.
- Sobol', I. and Levitan, Y. (1999), "On the Use of Variance Reducing Multipliers in Monte Carlo Computations of a Global Sensitivity Index," *Computer Physics Communication*, 117, 52–61.
- Sobol', I. M. (1990), "Sensitivity estimates for non-linear mathematical models," *Matematicheskoe Modelirovanie*, 2, 112–118.
- (1993), "Sensitivity analysis for non-linear mathematical models," *Mathematical Model. Comput. Exp.*, 1, 407–414.
- Storlie, C. B., Reich, B. J., Helton, J. C., Swiler, L. P., and Sallaberry, C. J. (2013), "Analysis of computationally demanding models with continuous and categorical inputs," *Reliability Engineering and System Safety*, 113, 30–41.
- Van Der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge University Press, Cambridge, U.K.
- Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell, T. J., and Morris, M. D. (1992), "Screening, predicting, and computer experiments," *Technometrics*, 34, 15–25.