Generalized Sobol' Indices for Multi-Output Regression Models*

Robert A. Milton[†], Solomon F. Brown[‡], and Aaron S. Yeardley[§]

Abstract. Variance based global sensitivity usually measures the relevance of inputs to a single output using Sobol' indices. This paper extends the definition in a natural way to multiple outputs, directly measuring the relevance of inputs to the linkages between outputs in a correlation-like matrix of indices. The usual Sobol' indices constitute the diagonal of this matrix. Existence, uniqueness and uncertainty quantification are established by developing the indices from a putative regression model. Analytic expressions for generalized Sobol' indices and their standard deviations are given, and tested computationally against test functions whose ANOVA can be performed.

Key words. Global Sensitivity Analysis, Sobol' Index, Surrogate Model, Multi-Output, Gaussian Process, Uncertainty Quantification

MSC codes. 60G07,60G15,62J10, 62H99

1. Introduction.

This paper is concerned with analysing the results of experiments or computer simulations in a design matrix of $M \geq 1$ input and $L \geq 1$ output columns, over N rows (datapoints). Global Sensitivity Analysis (GSA) [20] examines the relevance of the various inputs to the various outputs. When pursued via ANOVA decomposition of a single output $L = 1$, this leads naturally to the well known Sobol' indices, which have by now been applied across most fields of science and engineering [23, 10]. This paper extends the definition in a natural way to multiple outputs $L \geq 1$.

The Sobol' decomposition apportions the variance of a single output to sets of one or more inputs [24]. We shall use ordinals of inputs, tuples which are conveniently also naive sets.

$$\mathbf{m} := (0, \ldots, m-1) \subseteq (0, \ldots, M-1) =: \mathbf{M} \tag{1.1}$$

Obviously this restricts the subsets of inputs being studied, but this is a loss of convenience not generality, as any desired subset may be studied by ordering the inputs appropriately in advance. The maximal ordinal $\mathbf{M}$ of all M inputs explains everything explicable, so its Sobol' index is 1 by definition. The void ordinal $\mathbf{0}$ explains nothing, so its Sobol' index is 0 by definition. The influence of an isolated ordinal of inputs $\mathbf{m}$ is measured by its closed Sobol' index $S_{\mathbf{m}} \in [0, 1]$. A first-order Sobol' index $S_{m'}$ is simply the closed Sobol' index of a single input m'. Because inputs in an isolated ordinal may act in concert with each other, the influence of the ordinal often exceeds the sum of first-order contributions from its members,

*Submitted to the editors DATE.

Funding: This work was funded by the Fog Research Institute under contract no. FRI-454.

[†]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom (r.a.milton@sheffield.ac.uk, https://www.browngroupsheffield.com/).

[‡]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom (s.f.brown@sheffield.ac.uk, https://www.browngroupsheffield.com/).

[§]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom (asyeardley1@sheffield.ac.uk, https://www.browngroupsheffield.com/).

always obeying

$$S_{\mathbf{m}} \geq \sum_{m' \in \mathbf{m}} S_{m'}$$

Using the set-theoretic difference

$$\mathbf{M} - \mathbf{m} := \left\{ m' \in \mathbf{M} \,\middle|\, m' \notin \mathbf{m} \right\} \tag{1.2}$$

the complement of a closed index $S_{\mathbf{m}}$ is the total Sobol index of its complement

$$S^{T}_{\mathbf{M}-\mathbf{m}} := 1 - S_{\mathbf{m}} \in [0, 1] \tag{1.3}$$

This expresses the influence of non-isolated inputs $\mathbf{M} - \mathbf{m}$ allowed to act in concert with each other *and* isolated inputs $\mathbf{m}$. When speaking of irrelevant inputs $\mathbf{M} - \mathbf{m}$, we mean that $S^{T}_{\mathbf{M}-\mathbf{m}} \approx 0$. This is synonymous with the isolated ordinal of inputs $\mathbf{m}$ explaining everything explicable $S_{\mathbf{m}} \approx 1$.

Perhaps the most significant use of closed Sobol' indices is to identify a representative reduced model of $m \leq M$ inputs within the full model $\mathbf{M}$. Apportioning variance is mathematically equivalent to squaring a correlation coefficient to produce a coefficient of determination R^2 [5]. A closed Sobol' index is thus a coefficient of determination between the predictions from the reduced model $\mathbf{m}$ and predictions from the full model $\mathbf{M}$. A closed Sobol' index close to 1 confirms that the two models make nearly identical predictions. Simplicity and economy (not least of calculation) motivate the adoption of a reduced model, a closed Sobol' index close to 1 is what permits it.

There is no generally accepted extension of Sobol' indices to multiple outputs $L > 1$. The discussion thus far, and almost all prior GSA, has dealt with a single (i.e scalar) output. With multiple (i.e vector) outputs, the Sobol' decomposition apportions the covariance matrix of outputs rather than the variance of a single output. With L outputs, the closed Sobol' index $S_{\mathbf{m}}$ is generally a symmetric $\mathbf{L} \times \mathbf{L}$ matrix. The diagonal elements express the relevance of inputs to the output variables themselves. The off-diagonal elements express relevance to the linkages between outputs. This may be of considerable interest when outputs are, for example, yield and purity of a product, or perhaps a single output measured at various times. The Sobol indices reveal (amongst other things) which inputs it is worthwhile varying in an effort to alter the linkages between outputs. Prior work on Sobol' indices with multiple outputs [8, 30, 9] has settled ultimately on just the diagonal elements of the covariance matrix, so this linkage remains unexamined. Although output covariance has been incoroporated indirectly in prior studies by performing principal component analysis (PCA) on ouputs prior to GSA on the (diagonal) variances of the resulting output basis [2]. This has been used in particular to study synthetic "multi-outputs" which are actually the dynamic response of a single output over time [14, 33].

Accurate calculation of Sobol' indices even for a single output is computationally expensive and requires 10,000+ datapoints [15]. A (sometimes) more efficient approach is calculation via a surrogate model, such as Polynomial Chaos Expansion (PCE) [11, 32, 31], low-rank tensor approximation [4, 13], and support vector regression [7]. As well as being efficient, surrogate models also smooth out noise in the output, which is often highly desirable in practice. This paper employs one of the most popular surrogates, the Gaussian Processes (GP) [22, 19] as it is highly tractable. We shall follow the multi-output form (MOGP) described in [1], in order to examine the linkages between outputs. This paper deals exclusively with the anisotropic Radial Basis Function kernel, known as RBF/ARD, which is widely accepted as the kernel of choice for smooth outputs [18]. This uses the classic Gaussian bell curve to express the proximity of two input points, described in detail in **????**.

Semi-analytic expressions for Sobol' indices are available for scalar PCEs [26], and the diagonal elements of multi-output PCEs [9]. Semi-analytic expressions for Sobol' indices of GPs have been provided in integral form by [17] and alternatively by [3]. These approaches are implemented, examined and compared in [16, 25]. Both [17, 16] estimate the errors on Sobol' indices in semi-analytic, integral form. Fully analytic, closed form expressions have been derived without error estimates for uniformly distributed inputs [29] with an RBF kernel. There are currently no closed form expressions for MOGPs, or the errors on Sobol' indices, or any GPs for which inputs are not uniformly distributed.

In this paper we provide explicit, closed-form analytic formulae for the $\mathbf{L} \times \mathbf{L}$ matrices of closed Sobol' indices and their errors, for a class of MOGP with an RBF/ARD kernel applicable to smoothly varying outputs. We transform uniformly distributed inputs u to normally distributed inputs z prior to fitting a GP and performing analytic calculation of closed Sobol' indices. This leads to relatively concise expressions in terms of exponentials, and enables ready calculation of the errors (variances) of these expressions. It also allows for an arbitrary rotation Θ of inputs, as normal variables are additive, whereas summing uniform inputs does not produce uniform inputs. If the goal is reducing inputs, rotating their basis first boosts the possibilities immensely [6]. It presents the possibilty of choosing Θ to maximise the closed Sobol' index of the first few inputs.

The quantities to be calculated and their formal context are introduced in section 2, assuming only that the output is an integrable function of the input. Our approach effectively regards a regression model which quantifies uncertainty with each prediction as just another name for a stochastic process. A great deal of progress is made in section 3 using general stochastic (not necessarily Gaussian) processes. This approach is analytically cleaner, as it is not obfuscated by the GP details. Furthermore, it turns out that the desirable properties of the Gaussian (lack of skew, simple kurtosis) are not actually helpful, as these terms cancel of their own accord. This development leaves just two terms to be calculated, which require the stochastic process to be specified. MOGPs with an RBF/ARD kernel are tersely developed and described in **??**, then used to calculate the two unknown terms in **????**. Methods to reduce computational complexity are discussed in **??**.Conclusions are drawn in **??**.

2. Generalized Sobol' indices.

Apply a constant offset to a Lebesgue integrable model so that

$$y\colon [0,1]^{M+1} \to \mathbb{R}^L \quad \text{obeys} \quad \int y(u)\, \mathrm{d}u = [0]_{\mathbf{L}} \tag{2.1}$$

taking as input a uniformly distributed random variable (RV)

$$\mathsf{u} \sim \mathsf{U}\big([0]_{\mathbf{M+1}}, [1]_{\mathbf{M+1}}\big) := \mathsf{U}(0,1)^{M+1} \tag{2.2}$$

Throughout this paper exponentiation is categorical – repeated cartesian $\times$ or tensor $\otimes$ – unless otherwise specified. Square bracketed quantities are tensors, carrying their axes as a subscript tuple. In this case the subscript tuple is the von Neumann ordinal

$$\mathbf{M}+\mathbf{1} := (0,\ldots,M) \supset \mathbf{m} := (0,\ldots,m-1 \leq M-1)$$

with void $\mathbf{0} := ()$ voiding any tensor it subscripts. Ordinals are concatenated into tuples by Cartesian $\times$ and will be subtracted like sets, as in $\mathbf{M}-\mathbf{m} := (m,\ldots,M-1)$. Subscripts label the tensor prior to any superscript operation, so $[y(\mathsf{u})]^2_{\mathbf{L}}$ is an $\mathbf{L}^2 := \mathbf{L}\times\mathbf{L}$ tensor, for example. The preference throughout this work is for uppercase constants and lowercase variables, in case of ordinals the lowercase ranging over the uppercase. We prefer o for an unbounded positive integer, avoiding O.

Expectations and variances will be subscripted by the dimensions of u marginalized. Conditioning on the remaining inputs is left implicit after (2.3), to lighten notation. Now, construct $M+1$ stochastic processes (SPs)

$$[\mathsf{y}_{\mathbf{m}}]_{\mathbf{L}} := \mathbb{E}_{\mathbf{M}-\mathbf{m}}[y(\mathsf{u})] := \mathbb{E}_{\mathbf{M}-\mathbf{m}}\left[y(\mathsf{u})\middle|\, [u]_{\mathbf{m}}\right] \tag{2.3}$$

ranging from $[\mathsf{y}_{\mathbf{0}}]_{\mathbf{L}}$ to $[\mathsf{y}_{\mathbf{M}}]_{\mathbf{L}}$. Every SP depends stochastically on the ungoverned noise dimension $[\mathsf{u}]_M \perp [\mathsf{u}]_{\mathbf{M}}$ and deterministically on the first m governed inputs $[u]_{\mathbf{m}}$, marginalizing the remaining inputs $[\mathsf{u}]_{\mathbf{M}-\mathbf{m}}$. Sans serif symbols such as u,y generally refer to RVs and SPs, italic u, y being reserved for (tensor) functions and variables. Each SP is simply a regression model for y on the first m dimensions of u.

Following the Kolmogorov extension theorem [21, pp.124] we may regard an SP as a random function, from which we shall freely extract finite dimensional distributions generated by a design matrix $[u]_{\mathbf{M}\times\mathbf{o}}$ of $o \in \mathbb{Z}^+$ input samples. The Kolmogorov extension theorem incidentally secures u. Because y is (Lebesgue) integrable it must be measurable, guaranteeing $[\mathsf{y}_{\mathbf{0}}]_{\mathbf{L}}$. Because all probability measures are finite, integrability of y implies integrability of y^n for all $n \in \mathbb{Z}^+$ [27]. So Fubini's theorem [28, pp.77] allows all expectations to be taken in any order. These observations suffice to secure every object appearing in this paper. Changing the order of expectations, as permitted by Fubini's theorem, is the vital tool used throughout to construct this work.

Our aim is to compare predictions from a reduced regression model $\mathsf{y}_{\mathbf{m}}$ with those from the full regression model $\mathsf{y}_{\mathbf{M}}$. Correlation between these predictions is squared – using element-wise (Hadamard) multiplication $\circ$ and division $/$ – to form an RV called the coefficient of determination

$$\left[\mathsf{R}^2_{\mathbf{mM}}\right]_{\mathbf{L}^2} := \frac{\mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{m}},\mathsf{y}_{\mathbf{M}}] \circ \mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{m}},\mathsf{y}_{\mathbf{M}}]}{\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}] \circ \mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{M}}]} = \frac{\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}]}{\mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{M}}]} \tag{2.4}$$

However, this is undefined whenever $\mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{M}}]_{l\times l'} = 0$, obscuring potentially useful information about $\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}]_{l\times l'}$. Introducing 1-tensors representing the square root diagonal of a covariance matrix

$$\left[\sqrt{\mathbb{V}[\cdot,\cdot]_{\mathbf{L}^2}}\right]_l := \mathbb{V}[\cdot,\cdot]^{1/2}_{l^2} \tag{2.5}$$

the correlation coefficient between output dimensions is

$$[\mathsf{R}_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} := \frac{\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'}}{\sqrt{\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}]_{\mathbf{L}^2}} \otimes \sqrt{\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}]_{\mathbf{L}'^2}}} \quad \forall \mathbf{m} \subseteq \mathbf{M} \tag{2.6}$$

Let us define the multi-output closed Sobol' index as the product of the full correlation between output dimensions and the coefficient of determination

$$[\mathsf{S}_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} := [\mathsf{R}_{\mathbf{M}}]_{\mathbf{L}\times\mathbf{L}'} \circ \left[\mathsf{R}^2_{\mathbf{mM}}\right]_{\mathbf{L}\times\mathbf{L}'} \tag{2.7}$$

and the multi-output total Sobol' index as its complement

$$\left[\mathsf{S}^{\mathsf{T}}_{\mathsf{M}-\mathsf{m}}\right]_{\mathbf{L}\times\mathbf{L}'} := [\mathsf{S}_{\mathbf{M}}]_{\mathbf{L}\times\mathbf{L}'} - [\mathsf{S}_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} \tag{2.8}$$

These definitions coincide precisely with the traditional Sobol' index along the diagonal $\sqrt{[\mathsf{S}_{\mathbf{m}}]_{\mathbf{L}^2}} \circ \sqrt{[\mathsf{S}_{\mathbf{m}}]_{\mathbf{L}^2}}$, which has been very much the focus of prior literature [8, 30, 9]. The off-diagonal elements are bound by the diagonal as

$$-\left[\mathsf{S}_{\mathbf{m}}\right]_{l^2}^{1/2}\left[\mathsf{S}_{\mathbf{m}}\right]_{l'^2}^{1/2} \leq [\mathsf{S}_{\mathbf{m}}]_{l\times l'} = [\mathsf{R}_{\mathbf{m}}]_{l\times l'}\left[\mathsf{S}_{\mathbf{m}}\right]_{l^2}^{1/2}\left[\mathsf{S}_{\mathbf{m}}\right]_{l'^2}^{1/2} \leq \left[\mathsf{S}_{\mathbf{m}}\right]_{l^2}^{1/2}\left[\mathsf{S}_{\mathbf{m}}\right]_{l'^2}^{1/2} \tag{2.9}$$

To calculate moments over ungoverned noise we use the Taylor series method [12, pp.353], which is valid provided $\mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{M}}]_{l^2}$ is well approximated by its mean

$$[V_{\mathbf{M}}]_{l^2} := \mathbb{E}_M \mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{M}}]_{l^2} \gg \left|\mathbb{V}_{\mathbf{M}}[\mathsf{y}_{\mathbf{M}}]_{l^2} - [V_{\mathbf{M}}]_{l^2}\right| \tag{2.10}$$

This provides the mean Sobol' index

$$[S_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} := \mathbb{E}_M[\mathsf{S}_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} = \frac{[V_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'}}{\sqrt{[V_{\mathbf{M}}]_{\mathbf{L}^2}} \otimes \sqrt{[V_{\mathbf{M}}]_{\mathbf{L}'^2}}} \tag{2.11}$$

$$\text{where } [V_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} := \mathbb{E}_M \mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} \quad \forall \mathbf{m} \subseteq \mathbf{M} \tag{2.12}$$

with standard deviation due to ungoverned noise of

$$[T_{\mathbf{m}}]^2_{\mathbf{L}\times\mathbf{L}'} := \mathbb{V}_M[\mathsf{S}_{\mathbf{m}}]_{(\mathbf{L}\times\mathbf{L}')^2} = \frac{[Q_{\mathbf{m}}]_{(\mathbf{L}\times\mathbf{L}')^2}}{[V_{\mathbf{M}}]^{2/2}_{\mathbf{L}^2} \otimes [V_{\mathbf{M}}]^{2/2}_{\mathbf{L}'^2}} \tag{2.13}$$

where improper fractions exponentiate a square root diagonal of $V_{\mathbf{M}}$, and

$$\begin{aligned}[Q_{\mathbf{m}}]_{(\mathbf{L}\times\mathbf{L}')^2} := {} & [W_{\mathbf{mm}}]_{(\mathbf{L}\times\mathbf{L}')^2} - [V_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} \circ \sum_{\mathbf{L}^\circ\in\{\mathbf{L},\mathbf{L}'\}} \frac{[W_{\mathbf{Mm}}]_{\mathbf{L}^{\circ 2}\times\mathbf{L}\times\mathbf{L}'}}{[V_{\mathbf{M}}]^{2/2}_{\mathbf{L}^{\circ 2}}} \\ & + \frac{[V_{\mathbf{m}}]^2_{\mathbf{L}\times\mathbf{L}'}}{4} \circ \sum_{\mathbf{L}^\circ\in\{\mathbf{L},\mathbf{L}'\}} \frac{[W_{\mathbf{MM}}]_{\mathbf{L}^{\circ 2}\times\mathbf{L}^2}}{[V_{\mathbf{M}}]^{2/2}_{\mathbf{L}^{\circ 2}} \otimes [V_{\mathbf{M}}]^{2/2}_{\mathbf{L}^2}} + \frac{[W_{\mathbf{MM}}]_{\mathbf{L}^{\circ 2}\times\mathbf{L}'^2}}{[V_{\mathbf{M}}]^{2/2}_{\mathbf{L}^{\circ 2}} \otimes [V_{\mathbf{M}}]^{2/2}_{\mathbf{L}'^2}}\end{aligned} \tag{2.14}$$

$$[W_{\mathbf{mm}'}]_{\mathbf{L}^4} := \mathbb{V}_M[\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}], \mathbb{V}_{\mathbf{m}'}[\mathsf{y}_{\mathbf{m}'}]]_{\mathbf{L}^4} \tag{2.15}$$

It is satisfying to note that these Equations enforce

$$[T_{\mathbf{M}}]_{l^2} = 0 \quad \text{on the diagonal} \quad [\mathsf{S}_{\mathbf{M}}]_{l^2} = 1 \tag{2.16}$$

In practice it may be best to retain only the term in $W_{\mathbf{mm}}$, ignoring the uncertainty in $V_{\mathbf{M}}$ conveyed by $W_{\mathbf{Mm}}, W_{\mathbf{MM}}$, because these may drastically reduce uncertainty whenever $V_{\mathbf{m}} \approx V_{\mathbf{M}}$, which is the circumstance of greatest interest. In a similar vein, the total index is defined as the difference between the closed index and $[\mathsf{S}_{\mathbf{M}}]_{\mathbf{L}\times\mathbf{L}'}$, which is exactly 1 on the diagonal and usually highly correlated with $[\mathsf{S}_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'}$ off the diagonal. Tiny differences between terms in the correlated case are swamped by numerical errors in practice. Extending square root diagonal notation naturally to 2 dimensions, the conservative standard deviation

$$\left[T^T_{\mathbf{M}-\mathbf{m}}\right]_{\mathbf{L}\times\mathbf{L}'} := [T_{\mathbf{M}}]_{\mathbf{L}\times\mathbf{L}'} + [T_{\mathbf{m}}]_{\mathbf{L}\times\mathbf{L}'} \geq \sqrt{\mathbb{V}_M\left[\mathsf{S}^T_{\mathbf{M}-\mathbf{m}}\right]_{(\mathbf{L}\times\mathbf{L}')^2}} \tag{2.17}$$

achieves equality on the diagonal, and is robust and sufficiently precise for most practical purposes.

The remainder of this paper is devoted to calculating these two quantities – the generalized closed Sobol' Index $S_{\mathbf{m}}$ and its standard deviation due to ungoverned noise $T_{\mathbf{m}}$.

3. Stochastic Process estimates. The central problem in calculating errors on Sobol' indices is that they involve ineluctable covariances between differently marginalized SPs, via their moments over ungoverned noise. But marginalization and moment determination are both a matter of taking expectations. So the ineluctable can be avoided by reversing the order of expectations – taking moments over ungoverned noise, then marginalizing. To this end, adopt as design matrix a triad of inputs to condition $[\mathsf{u}]_{(\mathbf{M}+\mathbf{1})\times\mathbf{3}}$, eliciting the response

$$[\mathsf{y}]_{\mathbf{L}\times\mathbf{3}} := \mathbb{E}_{\mathbf{M}}\,\mathbb{E}_{\mathbf{M}'-\mathbf{m}'}\,\mathbb{E}_{\mathbf{0}''}\left[y([\mathsf{u}]_{(\mathbf{M}+\mathbf{1})\times\mathbf{3}}) \,\middle|\, [[u]_{\mathbf{0}}\,, [u]_{\mathbf{m}'}\,, [u]_{\mathbf{M}''}]\right] \tag{3.1}$$

Primes mark independent inputs, otherwise expectations are shared by all three members of the triad. It is not always obvious whether inputs are independent or shared by the triad, but this can be mechanically checked against the measure of integration behind an expectation. Repeated expectations over the same axis are rare here, usually indicating that apparent repetitions must be "primed". The purpose of the triad is to interrogate its response for moments in respect of ungoverned noise (which is shared by the triad members)

$$[\mu_n]_{(\mathbf{L}\times\mathbf{3})^n} := \mathbb{E}_M\left[[\mathsf{y}]^n_{\mathbf{L}\times\mathbf{3}}\right] \quad \forall n \in \mathbb{Z}^+ \tag{3.2}$$

for these embody

$$[\mu_{\mathbf{m}'\ldots\mathbf{m}^{n\prime}}]_{\mathbf{L}^n} := [\mu_n]_{\prod_{j=1}^n(\mathbf{L}\times i_j)} = \mathbb{E}_M[[\mathsf{y}_{\mathbf{m}'}]_{\mathbf{L}} \otimes \cdots \otimes [\mathsf{y}_{\mathbf{m}^{n\prime}}]_{\mathbf{L}}]$$

where $i_j \in \mathbf{3}$ corresponds to $\mathbf{m}^{j\prime} \in \{\mathbf{0}, \mathbf{m}, \mathbf{M}\}$. This expression underpins the quantities we seek. The reduction which follows repeatedly realizes the iterated expectation law

$$[\mu_{\mathbf{0}\ldots\mathbf{0}\mathbf{m}^{j\prime}\ldots\mathbf{m}^{n\prime}}]_{\mathbf{L}^n} := \mathbb{E}_{\mathbf{M}}[\mu_{\mathbf{M}\ldots\mathbf{M}\mathbf{m}^{j\prime}\ldots\mathbf{m}^{n\prime}}]_{\mathbf{L}^n} = \mathbb{E}_{\mathbf{m}}[\mu_{\mathbf{m}\ldots\mathbf{m}\mathbf{m}^{j\prime}\ldots\mathbf{m}^{n\prime}}]_{\mathbf{L}^n} \tag{3.3}$$

and that y was offset in (2.1) to obey

$$[\mu_{\mathbf{0}}]_{\mathbf{L}} = [0]_{\mathbf{L}} \tag{3.4}$$

and in any case

$$[V_{\mathbf{0}}]_{\mathbf{L}^2} = [\mu_{\mathbf{00}}]_{\mathbf{L}^2} - [\mu_{\mathbf{0}}]^2_{\mathbf{L}} = [0]_{\mathbf{L}^2} \tag{3.5}$$

Defining

$$[\mathsf{e}]_{\mathbf{L}\times\mathbf{3}} := \mathsf{y} - \mu_1 \tag{3.6}$$

the expected conditional variance in (2.11) amounts to

$$\begin{aligned}[V_{\mathbf{m}}]_{\mathbf{L}^2} &= \mathbb{E}_{\mathbf{m}}\mathbb{E}_M\left[[\mathsf{e}_{\mathbf{m}} + \mu_{\mathbf{m}}]^2_{\mathbf{L}}\right] - \mathbb{E}_M\left[[\mathsf{e}_{\mathbf{0}} + \mu_{\mathbf{0}}]^2_{\mathbf{L}}\right] \\ &= \mathbb{E}_{\mathbf{m}}\left[[\mu_{\mathbf{m}}]^2_{\mathbf{L}}\right] - [\mu_{\mathbf{0}}]^2_{\mathbf{L}} + \mathbb{E}_{\mathbf{m}}[\mu_{\mathbf{mm}}]_{\mathbf{L}^2} - [\mu_{\mathbf{00}}]_{\mathbf{L}^2} \\ &= \mathbb{E}_{\mathbf{m}}\left[[\mu_{\mathbf{m}}]^2_{\mathbf{L}}\right]\end{aligned} \tag{3.7}$$

and the covariance between conditional variances in (2.15) is

$$\begin{aligned}[W_{\mathbf{mm}'}]_{\mathbf{L}^4} &:= \mathbb{V}_M[\mathbb{V}_{\mathbf{m}}[\mathsf{y}_{\mathbf{m}}], \mathbb{V}_{\mathbf{m}'}[\mathsf{y}_{\mathbf{m}'}]] \\ &= \mathbb{V}_M\left[\mathbb{E}_{\mathbf{m}}\left[[\mathsf{y}_{\mathbf{m}}]^2_{\mathbf{L}} - [\mathsf{y}_{\mathbf{0}}]^2_{\mathbf{L}}\right], \mathbb{E}_{\mathbf{m}'}\left[[\mathsf{y}_{\mathbf{m}'}]^2_{\mathbf{L}} - [\mathsf{y}_{\mathbf{0}}]^2_{\mathbf{L}}\right]\right] \\ &= \mathbb{E}_M\left[\mathbb{E}_{\mathbf{m}}\left[[\mathsf{y}_{\mathbf{m}}]^2_{\mathbf{L}} - [\mathsf{y}_{\mathbf{0}}]^2_{\mathbf{L}}\right] \otimes \mathbb{E}_{\mathbf{m}'}\left[[\mathsf{y}_{\mathbf{m}'}]^2_{\mathbf{L}} - [\mathsf{y}_{\mathbf{0}}]^2_{\mathbf{L}}\right]\right] \\ &\quad - [V_{\mathbf{m}}]_{\mathbf{L}^2} \otimes [V_{\mathbf{m}'}]_{\mathbf{L}^2} \\ &= [A_{\mathbf{mm}'} - A_{\mathbf{0m}'} - A_{\mathbf{m0}} + A_{\mathbf{00}}]_{\mathbf{L}^4}\end{aligned}$$

Here, the inputs within any $\mathbf{m}, \mathbf{m}' \subseteq \mathbf{M}$ clearly vary independently, and

$$\begin{aligned}[A_{\mathbf{mm}'}]_{\mathbf{L}^4} &:= \mathbb{E}_M\mathbb{E}_{\mathbf{m}'}\mathbb{E}_{\mathbf{m}}\left[[\mathsf{y}_{\mathbf{m}}]^2_{\mathbf{L}} \otimes [\mathsf{y}_{\mathbf{m}'}]^2_{\mathbf{L}}\right] - [V_{\mathbf{m}}]_{\mathbf{L}^2} \otimes [V_{\mathbf{m}'}]_{\mathbf{L}^2} \\ &= \mathbb{E}_{\mathbf{m}}\mathbb{E}_{\mathbf{m}'}\mathbb{E}_M\left[[\mathsf{e}_{\mathbf{m}} + \mu_{\mathbf{m}}]^2_{\mathbf{L}} \otimes [\mathsf{e}_{\mathbf{m}'} + \mu_{\mathbf{m}'}]^2_{\mathbf{L}} - [\mu_{\mathbf{m}}]^2_{\mathbf{L}} \otimes [\mu_{\mathbf{m}'}]^2_{\mathbf{L}}\right]\end{aligned}$$

exploiting the fact that $V_{\mathbf{0}} = [0]_{\mathbf{L}^2}$. Equation (3.3) cancels all terms beginning with $[\mathsf{e}_{\mathbf{m}}]^2_{\mathbf{L}}$, first across $A_{\mathbf{mm}'} - A_{\mathbf{0m}'}$ then across $A_{\mathbf{m0}} - A_{\mathbf{00}}$. All remaining terms ending in $[\mu_{\mathbf{m}'}]^2_{\mathbf{L}}$ are eliminated by centralization $\mathbb{E}_M[\mathsf{e}_{\mathbf{m}}] = 0$. Similar arguments eliminate $[\mathsf{e}_{\mathbf{m}'}]^2_{\mathbf{L}}$ and $[\mu_{\mathbf{m}}]^2_{\mathbf{L}}$. Effectively then

$$[A_{\mathbf{mm}'}]_{\mathbf{L}^4} = \sum_{\pi(\mathbf{L}^2)}\sum_{\pi(\mathbf{L}'^2)} \mathbb{E}_{\mathbf{m}}\mathbb{E}_{\mathbf{m}'}[\mu_{\mathbf{m}} \otimes \mu_{\mathbf{mm}'} \otimes \mu_{\mathbf{m}'}]_{\mathbf{L}^2\times\mathbf{L}'^2}$$

so (3.4) entails

$$[W_{\mathbf{mm}'}]_{\mathbf{L}^4} = \sum_{\pi(\mathbf{L}^2)}\sum_{\pi(\mathbf{L}'^2)} \mathbb{E}_{\mathbf{m}}\mathbb{E}_{\mathbf{m}'}[\mu_{\mathbf{m}} \otimes \mu_{\mathbf{mm}'} \otimes \mu_{\mathbf{m}'}]_{\mathbf{L}^2\times\mathbf{L}'^2} \tag{3.8}$$

where each summation is over permutations of tensor axes

$$\pi(\mathbf{L}^2) := \left\{(\mathbf{L} \times \mathbf{L}''), (\mathbf{L}'' \times \mathbf{L})\right\} \quad ; \quad \pi(\mathbf{L}'^2) := \left\{(\mathbf{L}' \times \mathbf{L}'''), (\mathbf{L}''' \times \mathbf{L}')\right\}$$

Primes on constants are for bookeeping purposes only ($\mathbf{L}^{j\prime} = \mathbf{L}$ always), they do not change the value of the constant – unlike primes on variables ($\mathbf{m}^{j\prime}$ need not equal $\mathbf{m}$ in general). One is mainly interested in variances (errors), constituted by the diagonal $\mathbf{L}'^2 = \mathbf{L}^2$, for which the summation in (3.8) is over a pair of transposed pairs.

In order to further elucidate these estimates, we must fill in the details of the underlying SPs, sufficiently identifying the regression y by its first two moments μ_1, μ_2. Then the Sobol' indices are given by (2.11) and (3.7), and their standard deviation by (2.13), (2.14), and (3.8).

REFERENCES

[1] M. A. Alvarez, L. Rosasco, and N. D. Lawrence, *Kernels for vector-valued functions: a review*, 2011, https://arxiv.org/abs/http://arxiv.org/abs/1106.6251v2.

[2] K. Campbell, M. D. McKay, and B. J. Williams, *Sensitivity analysis when model outputs are functions*, Reliability Engineering & System Safety, 91 (2006), pp. 1468–1472, https://doi.org/10.1016/j.ress.2005.11.049.

[3] W. Chen, R. Jin, and A. Sudjianto, *Analytical variance-based global sensitivity analysis in simulation-based design under uncertainty*, Journal of Mechanical Design, 127 (2005), p. 875, https://doi.org/10.1115/1.1904642.

[4] M. Chevreuil, R. Lebrun, A. Nouy, and P. Rai, *A least-squares method for sparse low rank approximation of multivariate functions*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 897–921, https://doi.org/10.1137/13091899X, https://arxiv.org/abs/http://arxiv.org/abs/1305.0030v2.

[5] D. Chicco, M. J. Warrens, and G. Jurman, *The coefficient of determination r-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation*, PeerJ Computer Science, 7 (2021-07), p. e623, https://doi.org/10.7717/peerj-cs.623.

[6] P. G. Constantine, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, Society for Industrial and Applied Mathematics, mar 2015, https://doi.org/10.1137/1.9781611973860.

[7] C. Cortes and V. Vapnik, *Support-vector networks*, Machine Learning, 20 (1995), pp. 273–297, https://doi.org/10.1007/bf00994018.

[8] F. Gamboa, A. Janon, T. Klein, and A. Lagnoux, *Sensitivity indices for multivariate outputs*, Comptes Rendus Mathematique, 351 (2013), pp. 307–310, https://doi.org/10.1016/j.crma.2013.04.016.

[9] O. Garcia-Cabrejo and A. Valocchi, *Global sensitivity analysis for multivariate output using polynomial chaos expansion*, Reliability Engineering & System Safety, 126 (2014), pp. 25–36, https://doi.org/10.1016/j.ress.2014.01.005.

[10] R. Ghanem, D. Higdon, H. Owhadi, et al., *Handbook of uncertainty quantification*, vol. 6, Springer, 2017.

[11] R. G. Ghanem and P. D. Spanos, *Spectral techniques for stochastic finite elements*, Archives of Computational Methods in Engineering, 4 (1997), pp. 63–100, https://doi.org/10.1007/BF02818931.

[12] M. G. Kendall, *Kendall's advanced theory of statistics*, John Wiley & Sons, 1994.

[13] K. Konakli and B. Sudret, *Global sensitivity analysis using low-rank tensor approximations*, Reliability Engineering & System Safety, 156 (2016), pp. 64–83, https://doi.org/10.1016/j.ress.2016.07.012.

[14] M. Lamboni, H. Monod, and D. Makowski, *Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models*, Reliability Engineering & System Safety, 96 (2011), pp. 450–459, https://doi.org/10.1016/j.ress.2010.12.002.

[15] B. Lamoureux, N. Mechbal, and J. R. Massé, *A combined sensitivity analysis and kriging surrogate modeling for early validation of health indicators*, Reliability Engineering and System Safety, 130

(2014), pp. 12–26, https://doi.org/10.1016/j.ress.2014.03.007.

[16] A. Marrel, B. Iooss, B. Laurent, and O. Roustant, *Calculations of sobol indices for the gaussian process metamodel*, Reliability Engineering & System Safety, 94 (2009), pp. 742–751, https://doi.org/10.1016/j.ress.2008.07.008.

[17] J. E. Oakley and A. O'Hagan, *Probabilistic sensitivity analysis of complex models: a bayesian approach*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66 (2004), pp. 751–769, https://doi.org/10.1111/j.1467-9868.2004.05304.x.

[18] C. E. Rasmussen, *Some useful gaussian and matrix equations*, 2016, http://mlg.eng.cam.ac.uk/teaching/4f13/1617/gaussian%20and%20matrix%20equations.pdf.

[19] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning series)*, The MIT Press, 2005.

[20] S. Razavi, A. Jakeman, A. Saltelli, C. Prieur, B. Iooss, E. Borgonovo, E. Plischke, S. L. Piano, T. Iwanaga, W. Becker, S. Tarantola, J. H. Guillaume, J. Jakeman, H. Gupta, N. Melillo, G. Rabitti, V. Chabridon, Q. Duan, X. Sun, S. Smith, R. Sheikholeslami, N. Hosseini, M. Asadzadeh, A. Puy, S. Kucherenko, and H. R. Maier, *The future of sensitivity analysis: An essential discipline for systems modeling and policy support*, Environmental Modelling & Software, 137 (2021-03), p. 104954, https://doi.org/10.1016/j.envsoft.2020.104954.

[21] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales*, Cambridge University Press, 2000.

[22] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, *Design and analysis of computer experiments*, Statistical Science, 4 (1989), pp. 409–423.

[23] A. Saltelli, K. Aleksankina, W. Becker, P. Fennell, F. Ferretti, N. Holst, S. Li, and Q. Wu, *Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices*, Environmental Modelling & Software, 114 (2019-04), pp. 29–39, https://doi.org/10.1016/j.envsoft.2019.01.012.

[24] I. M. Sobol, *Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates*, Mathematics and Computers in Simulation, 55 (2001), pp. 271–280.

[25] A. Srivastava, A. K. Subramaniyan, and L. Wang, *Analytical global sensitivity analysis with gaussian processes*, Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 31 (2017), pp. 235–250, https://doi.org/10.1017/s0890060417000142.

[26] B. Sudret, *Global sensitivity analysis using polynomial chaos expansions*, Reliability Engineering & System Safety, 93 (2008), pp. 964–979, https://doi.org/10.1016/j.ress.2007.04.002.

[27] A. Villani, *Another note on the inclusion $L_p(\mu) \subset L_q(\mu)$*, The American Mathematical Monthly, 92 (1985), p. 485, https://doi.org/10.2307/2322503.

[28] D. Williams, *Probability with Martingales*, Cambridge University Press, 1991.

[29] Z. Wu, D. Wang, P. O. N, F. Hu, and W. Zhang, *Global sensitivity analysis using a gaussian radial basis function metamodel*, Reliability Engineering & System Safety, 154 (2016), pp. 171–179, https://doi.org/10.1016/j.ress.2016.06.006.

[30] S. Xiao, Z. Lu, and F. Qin, *Estimation of the generalized sobol's sensitivity index for multivariate output model using unscented transformation*, Journal of Structural Engineering, 143 (2017), https://doi.org/10.1061/(asce)st.1943-541x.0001721.

[31] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.

[32] D. Xiu and G. E. Karniadakis, *The wiener–askey polynomial chaos for stochastic differential equations*, SIAM Journal on Scientific Computing, 24 (2002), pp. 619–644, https://doi.org/10.1137/s1064827501387826.

[33] K. Zhang, Z. Lu, K. Cheng, L. Wang, and Y. Guo, *Global sensitivity analysis for multivariate output model and dynamic models*, Reliability Engineering & System Safety, 204 (2020), p. 107195, https://doi.org/10.1016/j.ress.2020.107195.