Kernel-based sensitivity indices for any model behavior and screening

Matieyendou Lamboni[1,2]

[1] University of Guyane, Department DFR-ST, 97346 Cayenne, French Guiana, France
[2] 2228-UMR Espace-Dev, Univ. of Guyane, Univ. of Réunion, IRD, Univ. of Montpellier, France

Abstract

Complex models are often used to understand interactions and drivers of human-induced and/or natural phenomena. It is worth identifying the input variables that drive the model output(s) in a given domain and/or govern specific model behaviors such as contextual indicators based on socioenvironmental models. Using the theory of multivariate weighted distributions to characterize specific model behaviors, we propose new measures of association between inputs and such behaviors. Our measures rely on sensitivity functionals (SFs) and kernel methods, including variance-based sensitivity analysis. The proposed ℓ1-based kernel indices account for interactions among inputs, higher-order moments of SFs, and their upper bounds are somehow equivalent to the Morris-type screening measures, including dependent elementary effects. Empirical kernel-based indices are derived, including their statistical properties for the computational issues, and numerical results are provided.

1. Introduction

Complex models or computer codes such as socio-environmental models (SEMs) are increasingly being developed and used to understand interactions and drivers that affect human well-being and the sustainability of the environment. Such information are used to support decision making for a wide range of problems and issues where the context may depend on the scales of the problem and the issues at stake. SEMs can be very complex and non-linear, and uncertainty arises in many guises, such as in the representation of the system of interest and in the data available to study them. Sensitivity analysis (SA) is a valuable tool to quantify the strength of the various un-controllable (like climate and other exogenous factors) and controllable (like policy and management levers) drivers on contextual indicators of well-being and sustainability (Elsawah et al., 2020).

Contextual indicators of well-being and sustainability often represent specific system behaviors. Conceptually, specific model behaviors are transformations of complex models or systems. Specific model behaviors encountered in the literature are mainly twofold regarding their impacts on the initial distribution of the model inputs. Firstly, the class of objective functions (e.g., Saltelli & Tarantola, 2002; Saltelli et al., 2004) consists of all transformations of the model outputs that do not modify the initial distribution of the model inputs. It includes transformations done by i) projecting the model outputs onto a given basis (Campbell et al., 2006; Lamboni et al., 2011; Xiao et al., 2018), using the kernel-based principal components, using feature maps of the outputs (Aronszajn, 1950; Schölkopf & Smola, 2002; Berlinet et al., 2004); ii) considering the probabilities of the

stochastic and dynamic outputs to exceed a given threshold (Lamboni et al., 2014), iii) using a regression-based classifier (Lamboni et al., 2016), and iv) considering the membership functions from either a crisp (a.k.a binary) or fuzzy clustering (Roux et al., 2021).

Secondly, we distinguish transformations that alter the initial distribution of the inputs such as considering the outputs values within a given cluster; constrained outputs or inputs; factor mapping (Spear & Hornberger, 1980; Rose et al., 1991; Saltelli et al., 2008; Lamboni et al., 2014) or excursion sets (Chevalier et al., 2013; French & Sain, 2013; Fossum et al., 2020), that is, the input values leading to a response to exceed a given threshold; considering the model output values with different chances of being included in the analysis of interest. For instance, membership functions (MFs) from a crisp clustering is equal to one for the output values that comply with the model behavior of interest and zero otherwise, while those from a fuzzy clustering give mainly non-negative values to the output values according to the way that such values meet a criterion of interest. We can see that the crisp clustering MFs will lead to consider only a subset of the output values such as the sustainability, safety or criticality.

Performing global SA (e.g., Sobol, 1993; Borgonovo et al., 2014; Owen, 2014; Gamboa et al., 2014; Mara et al., 2015; Lamboni & Kucherenko, 2021; Fort et al., 2021) on MFs allows for identifying the input variables that drive the model output(s) into a given domain of interest or a specific model behavior (Lamboni et al., 2014; Roux et al., 2021). However, such analysis does not provide the drivers of specific model behaviors. Moreover, considering contextual indicators of well-being and sustainability for decision supporting may implicitly introduce new dependency structures among the factors and levers, and new interactions among such inputs. It may lead to rare events, where outputs do not follow the Gaussian distribution. The recent dependent multivariate SA (dMSA, Lamboni & Kucherenko, 2021; Lamboni, 2021) requires the distribution of the model inputs that complies with the model behavior of interest (i.e., the target inputs distribution) for performing SA. Since dMSA is well-suited for Gaussian distributed sensitivity functionals (SFs), it may be insufficient for exploring contextual indicators in presence of higher-order moments of SFs.

Our paper presents an approach that helps addressing such modeling issues in that it develops new and generic measures of association between specific model behaviors and the drivers or inputs of such models. Our approach accounts for i) dependencies among input factors and levers, ii) interactions among the input factors and levers even for non-independent factors, and iii) higher-order moments of SFs. The proposed measures rely on kernel methods (Aronszajn, 1950; Berlinet et al., 2004) and the theory of multivariate weighted distributions (Fisher, 1934; Rao, 1965; Patil & Rao, 1978; Mahfoud & Patil, 1982; Jain & Nanda, 1995; Navarro et al., 2006; Arnold et al., 2017). The well-known variance-based SA (Sobol, 1993; Saltelli et al., 2008; Borgonovo et al., 2014) and multivariate SA (Lamboni, 2019; Lamboni & Kucherenko, 2021), which account for only the second-order moments of SFs, are particular cases of our new measures. We also extend the Morris screening method (Morris, 1991).

The paper is organized as follows: after characterizing the model behavior of interest by i) identifying the appropriate weight function or the multivariate weighted distribution, and ii) deriving the cumulative distribution function (CDF) of the target inputs and its dependency model in Section 2, we introduce the new measures of association between the target inputs and the model outputs in Section 3. We then provide the first-order and total kernel-based indices with the former index less than the latter. We also introduce a new upper bound of the total index. For the ℓ1-based kernel, the associated upper bound of the total index is equivalent to the Morris-type screening measures, including the dependent elementary effects (Morris, 1991; Lamboni & Kucherenko, 2021). Empirical kernel-based indices and their properties are provided in Section 4. Section 5 provides illustrations for some model behaviors of interest such as the critical domain and cluster of the output values, and Section 6 concludes this work.

General notation

For $d \in \mathbb{N} \setminus \{0\}$, let $\mathbf{X} := (X_1, \dots, X_d)$ be a random vector having F as the joint CDF and $F_j, j = 1, \dots, d$ as marginals. We use F_j^{-1} for the inverse of F_j, and $Y \overset{d}{=} Z$ when Y, Z have the same CDF. For a non-empty subset $u \subseteq \{1, \dots, d\}$, we use $|u|$ for its cardinality (i.e., the number of elements in u) and $(\sim u) := \{1, \dots, d\} \setminus u$. For a given u, we use $\mathbf{X}_u := (X_j, \forall j \in u)$ for a subset of inputs; $\mathbf{X}_{\sim u} := (X_j, \forall j \in (\sim u))$ and we have the partition $\mathbf{X} = (\mathbf{X}_u, \mathbf{X}_{\sim u})$.

We use $\|\cdot\|_p$ for the p-norm on $\mathbb{R}^n$, $\mathbb{E}(\cdot)$ for the expectation, $\mathbb{V}(\cdot)$ for the variance, $\overset{D}{\to}$ and $\overset{P}{\to}$ for the convergence in distribution and in probability, respectively.

2.   Target inputs and outputs distribution

This section aims at identifying the weight functions for some specific model behaviors and determining the distribution of the target inputs, that is the distribution of the inputs that complies with the model behaviors of interest.

2.1   Weight functions for some target outputs

2.1.1   Case of multivariate response models

For a vector-valued function $f : \mathbb{R}^d \to \mathbb{R}^n$, the inputs $\mathbf{X}$ have F as the initial CDF (i.e., $\mathbf{X} \sim F$), ρ as the density, and $f(\mathbf{X})$ is the initial model outputs. It is clear that the outputs distribution is determined by the inputs distribution.

Namely, for a given domain of interest D, let us consider the outputs of interest (called the target outputs) given by $\begin{cases} f(\mathbf{X}), \mathbf{X} \sim F \\ \text{s.t. } f(\mathbf{X}) \in D \end{cases}$. When the set $\{x \in \mathbb{R}^d : f(x) \in D\}$ is not empty, there exists a CDF (i.e., F^w) of the target inputs such that (see Lamboni, 2022a)

$$\begin{cases} f(\mathbf{X}) \text{ with } \mathbf{X} \sim F \\ \text{s.t. } f(\mathbf{X}) \in D \end{cases} \overset{d}{=} f(\mathbf{X}^w), \text{ with } \mathbf{X}^w \sim F^w . \tag{1}$$

By considering the weight function $w(x) = \mathbb{1}_D(f(x))$ with $\mathbb{1}_D(x)$ the indicator function, we can see that the density function (i.e., ρ^w) of $\mathbf{X}^w$ given by

$$\rho^w(\mathbf{x}) \propto w(\mathbf{x})\rho(\mathbf{x}) , \tag{2}$$

allows for characterizing the associated target outputs or the behavior of interest. Such weight function is identical to the crisp or binary clustering of the initial model outputs. For any classifier $c : \mathbb{R}^n \to \mathbb{R}^p$ with $p \in \mathbb{N} \setminus \{0\}$, we can extend the above weight function as follows:

$$w_1(\mathbf{x}) := \mathbb{1}_D(c \circ f(\mathbf{x})) ,$$

where $c \circ f$ stands for the composition of c by f. Thus, we can cover many clustering approaches (such as PCA, the kernel methods, random forest and logistic regression) by choosing the classifier c. To work with continuous weight functions, $\mathbb{1}_D$ can be replaced with smooth functions such as the logistic function, and this leads to the second kind of weight functions, that is,

$$w_2(\mathbf{x}) := m(c \circ f(\mathbf{x})), \text{ with } m : \mathbb{R}^p \to \mathbb{R}_+ .$$

The membership functions from a fuzzy clustering (Bezdek et al., 1984; Höppner et al., 1999) are particular cases of w_2. The general expression of a weight function is given as follows:

$$w_3(\mathbf{x}) := m(c \circ f(\mathbf{x}))\mathbb{1}_D\left(m_2(c_2 \circ f(\mathbf{x}))\right) ,$$

where $c_2 : \mathbb{R}^n \to \mathbb{R}^p$ and $m_2 : \mathbb{R}^p \to \mathbb{R}_+$ are given functions. The weight function w_3 can still be used for constraining, restricting and modifying directly the distribution of the initial inputs by taking $c \circ f$ and $c_2 \circ f$ as identity functions, that is, $c \circ f(\mathbf{x}) = \mathbf{x}$. These types of weight functions directly or indirectly affect the initial distribution of the model inputs as well as the initial outputs distribution because the target inputs distribution F^w induced the target outputs distribution and vice versa for deterministic functions or models. We then include them in a wide class of the outputs distributions of interest called the target outputs distributions.

Since the weight functions guide the initial model outputs toward the behavior of interest, performing SA using the weight functions helps identifying the input variables that drive the initial outputs toward the target outputs defined by such weight functions (see e.g. Lamboni et al. (2014); Roux et al. (2021) for independent initial inputs). Such works can be extended to cope with weight functions including dependent and/or correlated input variables by using the dependent SA introduced in Lamboni & Kucherenko (2021) and Lamboni (2021).

2.1.2   Extension to multivariate and functional outputs

Consider $\Theta \subseteq \mathbb{R}$ and a function $f : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^n$ given by $f(\mathbf{X}, \theta) \in \mathbb{R}^n$ with $\theta \in \Theta$. The function f represents the multivariate and functional response model, including dynamic models. When $\Theta = \{\theta_0\}$, it becomes a vector-valued function. For such functions, one may consider the following weight functions

$$w_5(\mathrm{x}) := L_\Theta(c \circ f(\mathrm{x}, \theta), \beta) \mathbb{1}_D(m_2(c_2 \circ f(\mathrm{x}), \theta), \forall \theta \in \Theta),$$

with L_Θ a desirability measure over all $\theta \in \Theta$ such as a loss function and β some thresholds (see Lamboni et al. (2014) for more details).

2.2   Distribution of the target inputs

This section deals with the distribution and conditional distributions of continuous variables following multivariate weighted distributions such as the target inputs.

In the theory of multivariate weighted distributions (Fisher, 1934; Rao, 1965; Patil & Rao, 1978; Mahfoud & Patil, 1982; Jain & Nanda, 1995; Navarro et al., 2006; Arnold et al., 2017), the density function of X^w (i.e., ρ^w) given by Equation (2) is known as the weighted density associated with the initial density ρ and the weight function w. The weight function aims at altering the density of the initial inputs by assigning unequal chances to the initial observations of being included in the posterior analysis. The formal definition of weight functions and the weighted probability density function (PDF) are given below. We use $\mathbb{E}_F$ for the expectation taking w.r.t. F and $\mathrm{x} \in \mathbb{R}^d$.

Definition 1   *(Fisher, 1934; Rao, 1965; Patil & Rao, 1978)*
Let $\mathrm{w} : \mathbb{R}^d \rightarrow \mathbb{R}_+$ *be a non-negative function.*

   (i)   When $\mathbb{E}_F[w(\mathbf{X})] < \infty$, then w is a weight function.

   (ii)   The weighted density (i.e., ρ^w) associated with w and ρ is given by

$$\rho^{\mathrm{w}}(\mathbf{x}) := \frac{\mathrm{w}(\mathbf{x})}{\mathbb{E}_F[\mathrm{w}(\mathbf{X})]} \rho(\mathbf{x}) . \tag{3}$$

   (iii)   The random vector $\mathbf{X}^{\mathrm{w}}$ having ρ^{w} as PDF and F^{w} as CDF (i.e., $\mathbf{X}^{\mathrm{w}} \sim \mathrm{F}^{\mathrm{w}}$) is called the weighted random vector.

The marginal PDF of $\mathbf{X}_u^w$ and the PDF of $\mathbf{X}_{\sim u}^w$ conditional on $\mathbf{X}_u^W$ (i.e., $\mathbf{X}_{\sim u}^w \mid \mathbf{X}_u^W$) are given by
(Mahfoud & Patil, 1982; Kocherlakota, 1995; Jain & Nanda, 1995; Navarro et al., 2006)

$$\rho_u^w(\mathbf{x}_u) := \frac{\mathbb{E}_{F_{\sim u|u}}[w(\mathbf{x}_u, \mathbf{X}_{\sim u})]}{\mathbb{E}_F[w(\mathbf{X})]} \rho_u(\mathbf{x}_u) ; \tag{4}$$

$$\rho_{\sim u|u}^w(\mathbf{x}_{\sim u} \mid \mathbf{x}_u) := \frac{w(\mathbf{x})}{\mathbb{E}_{F_{\sim u|u}}[w(\mathbf{x}_u, \mathbf{X}_{\sim u})]} \rho_{\sim u|u}(\mathbf{x}_{\sim u} \mid \mathbf{x}_u) , \tag{5}$$

where F_u (resp. $F_{\sim u|u}$) denotes the CDF of $\mathbf{X}_u$ (resp. $\mathbf{X}_{\sim u} \mid \mathbf{X}_u$), and ρ_u (resp. $\rho_{\sim u|u}$) denotes the PDF of $\mathbf{X}_u$ (resp. $\mathbf{X}_{\sim u} \mid \mathbf{X}_u$). Likewise, we are going to use F_u^W (resp. $F_{\sim u|u}^W$) for the distribution of $\mathbf{X}_u^W$ (resp. $\mathbf{X}_{\sim u}^W \mid \mathbf{X}_u^W$).

In what follows, we use $F_{\mathrm{ind}} := \prod_{j=1}^d F_j$ and $F_{\mathrm{ind}}(\mathrm{x}) := \prod_{j=1}^d F_j(x_j)$. For independent initial variables, it is clear that $F = F_{\mathrm{ind}}$ or $\rho = \prod_{j=1}^d \rho_j$ with ρ_j the PDF of X_j. Thus, F_{ind}^w denotes the weighted distribution associated with F_{ind} and w.

Remark 1   *When $F = F_{ind}$, we have*

$$\rho^w(\mathbf{x}) := \frac{w(\mathbf{x})}{\mathbb{E}_F[w(\mathbf{X})]} \prod_{j=1}^d \rho_j(x_j); \quad \rho_u^w(\mathbf{x}_u) = \frac{\mathbb{E}_{F_{\sim u}}[w(\mathbf{x}_u, \mathbf{X}_{\sim u})]}{\mathbb{E}_F[w(\mathbf{X})]} \prod_{j \in u} \rho_j(x_j);$$

$$\rho_{\sim u|u}^w(\mathbf{x}_{\sim u} \mid \mathbf{x}_u) = \frac{w(\mathbf{x})}{\mathbb{E}_{F_{\sim u}}[w(\mathbf{x}_u, \mathbf{X}_{\sim u})]} \prod_{j \in \{1, \dots, d\} \setminus u} \rho_j(x_j).$$

Since the expressions of the marginal and conditional PDFs in Remark 1 are much easier to work with, it is worth expressing any weighted random vector $\mathbf{X}^w$ as the random vector associated with $\mathbf{Y} \sim F_{ind}$ and a new weight function (i.e., w_e) (see Proposition 1). To that end, we use $C: [0,1]^d \to [0,1]$ for the copula of the initial inputs $\mathbf{X}$, that is, $F(\mathrm{x}) = C\big(F_1(x_1), \dots, F_d(x_d)\big)$ (Sklar, 1959; Nelsen, 2006; Joe, 2014).

Proposition 1 *Let $\mathbf{Y} \sim F_{ind}$ and assume that the copula C has $c(\cdot)$ as the joint PDF. Then, the PDF of $\mathbf{X}^w$ is given by*

$$\rho^w(\mathbf{x}) = \frac{w_e(\mathbf{x})}{\mathbb{E}_{F_{\mathrm{ind}}}[w_e(\mathbf{Y})]} \prod_{j=1}^{d} \rho_j(x_j); \quad w_e(\mathbf{x}) := w(\mathbf{x}) c\big(F_1(x_1), \dots, F_d(x_d)\big).$$

Proof *See Supplementary Material A.*

It comes out from Proposition 1 that the copula-based expressions of the PDF of **X** allow for writing

- $F = F_{\mathrm{ind}}^c$ to say that the initial distribution F is an adjustment of the distribution F_{ind} using the weight function c;
- $F^w = F_{ind}^{w_e}$ to say F^w is the weighted distribution associated with F_{ind} and w_e.

To deduce the CDF of the random vector $\mathbf{X}^w \sim F^w = F_{ind}^{w_e}$, including its copula (see Theorem 1), we use $\boldsymbol{\pi} := \big(\pi_1, \dots, \pi_{|\pi|}\big)$ for an arbitrary permutation of $(\sim u) := \{1, \dots, d\} \setminus u$ with the cardinality $|\boldsymbol{\pi}| = d - |u|$. We also use $\mathbf{u}_\pi := \big(u_{\pi_1}, \dots, u_{\pi_{|\pi|}}\big) \in [0,1]^{d-|u|}$; $\mathbf{V}_\pi := \big(V_{\pi_k} \sim \mathcal{U}(0, u_{\pi_k}), k = 1, \dots, |\boldsymbol{\pi}|\big)$ for a random vector of independent variables.

Theorem 1 *Consider the random vector $\mathbf{X}^w \sim F^w = F_{\mathrm{ind}}^{w_e}$ and $\mathbf{Y} \sim F_{\mathrm{ind}}$. Assume that $\mathbf{X}^w$ and $\mathbf{Y}$ are continuous random vectors. Then, there exist a CDF $W(\cdot; \mathbf{x}_u^w): [0,1]^{d-|u|} \to [0,1]$ given by*

$$W\big(\mathbf{u}_{|\pi|}; \mathbf{x}_u^w\big) = \frac{\mathbb{E}_{\mathbf{V}_\pi}\left[w_e\left(\mathbf{x}_u^w, F_{\pi_1}^{-1}(V_{\pi_1}), \dots, F_{\pi_{|\pi|}}^{-1}\big(V_{\pi_{|\pi|}}\big) \right) \right]}{\mathbb{E}_{F_{ind}}[w_e(\mathbf{x}_u^w, \mathbf{Y}_{\sim u})]} \prod_{j=1}^{d-|u|} u_{\pi_j}, \tag{6}$$

such that

$$F_{\sim u|u}^w(\mathbf{x}_{\sim u} \mid \mathbf{x}_u^w) = W\left(F_{\pi_1}\big(x_{\pi_1}\big), \dots, F_{\pi_{|\pi|}}\big(x_{\pi_{|\pi|}}\big); \mathbf{x}_u^w \right). \tag{7}$$

Proof *See Supplementary Material B.*

It is worth noting that $W(\cdot; \mathrm{x}_u^w)$ is a CDF of a random vector having $(0,1)^{d-|u|}$ as the joint support, and it involves only the weight function and the marginal CDFs of **X**. When $u = \emptyset$, we obtain the following expressions of W:

$$W(u_1, \dots, u_d) = \frac{\mathbb{E}_{\mathbf{V}}\big[w_e\big(F_1^{-1}(V_1), \dots, F_d^{-1}(V_d)\big) \big]}{\mathbb{E}_{F_{\mathrm{ind}}}[w_e(\mathbf{Y})]} \prod_{k=1}^{d} u_k, \tag{8}$$

and

$$F^w(\mathrm{x}) = W\big(F_1(x_1), \dots, F_d(x_d)\big) \tag{9}$$

Equations (6)-(9) are useful for sampling random values of $\mathbf{X}_{\sim u}^w \mid \mathbf{X}_u^w$ and $\mathbf{X}^w$, and for deriving the dependency models of $\mathbf{X}^w$.

3. Effects of the target inputs on the target outputs

Since introducing weight functions in our analysis often leads to dependent target inputs, and the target inputs can lead to rare events, this section aims at proposing and studying kernel-based sensitivity indices for identifying the input variables that govern the target outputs.

3.1   Sensitivity functionals based on weight functions

Sensitivity functionals (SFs) contain the primary information about the contribution of the model inputs (see Lamboni (2016, 2020b, 2020a, 2019, 2022b) for independent input variables, and Lamboni & Kucherenko (2021) and Lamboni (2021) for dependent and/or correlated input variables). The first-order SF allows for quantifying the single contribution of the inputs, while the total SF is used to measure the overall contribution of the inputs, including interactions.

For dependent weighted random vector (i.e., $\mathbf{X}^w \sim F^w$) having the CDFs given by Equations (6), (7) and (8), the dependency models of $\mathbf{X}^w$ (Skorohod, 1976; Lamboni & Kucherenko, 2021; Lamboni, 2021, 2022a, 2023) are given by

$$\mathbf{X}^w_{\sim u} = r(\mathbf{X}^w_u, \mathbf{U}_\pi) \stackrel{d}{=} \left(F^{-1}_{\pi_1}(Z_{\pi_1}), \dots, F^{-1}_{\pi_{|\pi|}}\left(Z_{\pi_{|\pi|}}\right) \right), \tag{10}$$

where $\mathbf{U}_\pi := \left(U_{\pi_1}, \dots, U_{\pi_{|\pi|}} \right) \sim \mathcal{U}(0,1)^{d-|u|}$ and

$$\begin{bmatrix} Z_{\pi_1} := W^{-1}_{\pi_1}\left(U_{\pi_1} \mid \mathbf{X}^w_u\right) \\ Z_{\pi_2} := W^{-1}_{\pi_2|\pi_1}\left(U_{\pi_2} \mid \mathbf{X}^w_u, Z_{\pi_1}\right) \\ \vdots \\ Z_{\pi_{|\pi|}} := W^{-1}_{\pi_{|\pi|}|\sim\pi_{|\pi|}}\left(U_{\pi_{|\pi|}} \mid \mathbf{X}^w_u, Z_{\pi_1}, \dots, Z_{\pi_{|\pi|-1}}\right) \end{bmatrix} \sim W(\cdot \mid \mathbf{X}^w_u),$$

with $W^{-1}_{\pi_2|\pi_1}$ the inverse of the CDF of Z_{π_2} conditional on Z_{π_1}. Composing the target outputs $f(\mathbf{X}^w, \theta)$ by the dependency model given by (10) yields

$$f(\mathbf{X}^w, \theta) \stackrel{d}{=} f(\mathbf{X}^w_u, r(\mathbf{X}^w_u, \mathbf{U}_\pi), \theta).$$

The first-order and total SFs associated with $\mathbf{X}^w_u$ are respectively defined by

$$f^{fo}_u(\mathbf{X}^w_u, \theta) := \mathbb{E}_U[f(\mathbf{X}^w_u, r(\mathbf{X}^w_u, \mathbf{U}_\pi), \theta)] - \mathbb{E}[f(\mathbf{X}^w_u, r(\mathbf{X}^w_u, \mathbf{U}_\pi), \theta)],$$

$$f^{tot}_u(\mathbf{X}^w, \mathbf{U}_\pi, \theta) := f(\mathbf{X}^w_u, r(\mathbf{X}^w_u, \mathbf{U}_\pi), \theta) - \mathbb{E}_{\mathbf{X}^w_u}[f(\mathbf{X}^w_u, r(\mathbf{X}^w_u, \mathbf{U}_\pi), \theta)].$$

Using the fact that $F^w = F^{w_e}_{ind}$, new expressions of SFs based on w_e are derived in Proposition 2. We use F_U for the CDF of $\mathcal{U}(0,1)^{d-|u|}$ in what follows.

Proposition 2 *Let $\mathbf{X}^w \sim F^w = F^{w_e}_{ind}$, $\mathbf{Y} \sim F_{ind}$ and $\mathbf{U} \sim F_U$ be independent random vectors. Then, the first-order and total SFs related to $\mathbf{X}^w_u$ are given by*

$$f^{fo}_u(\mathbf{X}^w_u, \theta) = \mathbb{E}_U[f(\mathbf{X}^w_u, r(\mathbf{X}^w_u, \mathbf{U}), \theta)] - \frac{\mathbb{E}[f(\mathbf{Y}_u, r(\mathbf{Y}_u, \mathbf{U}), \theta)w_e(\mathbf{Y})]}{\mathbb{E}[w_e(\mathbf{Y})]}; \tag{11}$$

$$f^{tot}_u(\mathbf{X}^w_u, \mathbf{U}, \theta) = f(\mathbf{X}^w_u, r(\mathbf{X}^w_u, \mathbf{U}), \theta) - \frac{\mathbb{E}_{\mathbf{Y}}[f(\mathbf{Y}_u, r_u(\mathbf{Y}_u, \mathbf{U}), \theta)w_e(\mathbf{Y})]}{\mathbb{E}[w_e(\mathbf{Y})]}. \tag{12}$$

Proof *The results are straightforward bearing in mind Remark 1 (see also Lamboni, 2023).*

It is worth noting that the above SFs can be adapted for structured model outputs such as categorical variables, sequences, graphs and labels.

3.2   Definition and properties of the kernel-based sensitivity indices

For an output set $\mathcal{Y}$ and $\mathbf{y}, \mathbf{y}' \in \mathcal{Y}$, it is realistic i) to map a class of labels or contentious variables $\mathcal{Y}$ into a feature space $\mathcal{F}_k \subseteq \mathbb{R}^L$ using the map $\mathbf{y} \mapsto \phi(\mathbf{y}) := \{\psi_\ell(\mathbf{y})\}^L_{\ell=1}$ with the possibility $L = \infty$; and ii) to describe the desirable structure by a similarity measure between $\mathbf{y}, \mathbf{y}'$ a.k.a. kernel, that is,

$$k(\mathbf{y}, \mathbf{y}') := \langle \phi(\mathbf{y}), \phi(\mathbf{y}') \rangle_{\mathcal{F}_k},$$

where $\langle \cdot, \cdot \rangle_{\mathcal{F}_k}$ stands for the inner product. Such kernel is symmetric and positive definite (SPD), and the Moore-Aronszajn theorem (Aronszajn, 1950; Berlinet et al., 2004) ensures that the feature space $\mathcal{F}_k$ is a RKHS endowed

with the inner product $k(\cdot,\cdot)$. The feature map allows not only for dealing with the structured outputs but also for considering derivable moments. In what follows, assume that

(A1): $k(\mathbf{y},\mathbf{y}')$ is SPD; $\phi(\mathbf{y}) = k(\mathbf{y},\cdot)$ is convex and $k(\mathbf{y},\mathbf{0}) = 0$.

Example: the ℓ_p-based kernel and the quadratic kernel given by

$$k_p(\mathbf{y},\mathbf{y}') := \left\langle \parallel \mathbf{y} \parallel_p^p, \parallel\mathbf{y}'\parallel_p^p \right\rangle; \; k_q(\mathbf{Y},\mathbf{Y}') := \langle \mathbf{Y},\mathbf{Y}'\rangle^2,$$

respectively, satisfy (A1).

Namely, we use $\mathbf{X}^{w'}$ (resp. $\mathbf{U}'$) for an i.i.d. copy of $\mathbf{X}^w \sim F_{ind}^{We}$ (resp. $\mathbf{U} \sim F_U$), and we consider the following functionals

$$f^c(\mathbf{X}_u^w,\mathbf{U},\theta) := f(\mathbf{X}_u^w,r(\mathbf{X}_u^w,\mathbf{U}),\theta) - \mathbb{E}[f(\mathbf{X}_u^w,r(\mathbf{X}_u^w,\mathbf{U}),\theta)]\,;$$

$$f_u^*\big(\mathbf{X}_u^w,\mathbf{X}_u^{w'},\mathbf{U},\theta\big) := f(\mathbf{X}_u^w,r(\mathbf{X}_u^w,\mathbf{U}),\theta) - f\big(\mathbf{X}_u^{w'},r(\mathbf{X}_u^{w'},\mathbf{U}),\theta\big)\,.$$

For a mensurable and SPD kernel k, we also assume that

(A2): $0 < \int_\Theta \mathbb{E}\left[k\left(f^c(\mathbf{X}_u^w,\mathbf{U},\theta),f^c(\mathbf{X}_u^{w'},\mathbf{U}',\theta)\right)\right] d\theta < +\infty.$

3.2.1 Kernel-based sensitivity indices for vector-valued functions

This section deals with a vector-valued function $f: \mathbb{R}^d \to \mathbb{R}^n$, which is a particular case of $f(\mathbf{X},\theta)$ with $\theta \in \Theta = \{\theta_0\}$. Thus, the above first-order and total SFs, the centered outputs f_u^c and f_u^* become respectively

$$f_u^{fo}(\mathbf{X}_u^w); \quad f_u^{tot}(\mathbf{X}_u^w,\mathbf{U}); \quad f^c(\mathbf{X}_u^w,\mathbf{U}); \quad f_u^*\big(\mathbf{X}_u^w,\mathbf{X}_u^{w'},\mathbf{U}\big)\,.$$

We also use

$$\mu_u^{*p} := \mathbb{E}\left[\left\|f_u^*\big(\mathbf{X}_u^w,\mathbf{X}_u^{w'},\mathbf{U}\big)\right\|_p^p\right]\,, \tag{13}$$

and we are going to see latter that the functional $f_u^*\big(\mathbf{X}_u^w,\mathbf{X}_u^{w'},\mathbf{U}\big)$ will lead to the upper bounds of the kernel-based SIs (Kb-SIs). Definition 2 formally introduces such indices.

Definition 2 *Let $X^{w'}, X^{w''}$, and $X^{w'''}$ be i.i.d. copies of X^w, and assume that (A1)-(A2) hold. Then, the first-order and total Kb-SIs of X_u^w are defined as follows*:

$$S_u^k := \frac{\mathbb{E}\left[k\left(f_u^{fo}(\mathbf{X}_u^w),f_u^{fo}(\mathbf{X}_u^{w'})\right)\right]}{\mathbb{E}\left[k\left(f^c(\mathbf{X}_u^w,\mathbf{U}),f^c(\mathbf{X}_u^{w'},\mathbf{U}')\right)\right]}\,; \tag{14}$$

$$S_{T_u}^k := \frac{\mathbb{E}\left[k\left(f_u^{tot}(\mathbf{X}_u^w,\mathbf{U}),f_u^{tot}(\mathbf{X}_u^{w'},\mathbf{U}')\right)\right]}{\mathbb{E}\left[k\left(f^c(\mathbf{X}_u^w,\mathbf{U}),f^c(\mathbf{X}_u^{w'},\mathbf{U}')\right)\right]}\,, \tag{15}$$

respectively. We also define

$$\Upsilon_u^k := \frac{\mathbb{E}\left[k\left(f_u^*(\mathbf{X}_u^w,\mathbf{X}_u^{w'},\mathbf{U}),f_u^*(\mathbf{X}_u^{w''},\mathbf{X}_u^{w'''},\mathbf{U}')\right)\right]}{\mathbb{E}\left[k\left(f^c(\mathbf{X}_u^w,\mathbf{U}),f^c(\mathbf{X}_u^{w'},\mathbf{U}')\right)\right]}\,. \tag{16}$$

For the ℓ_p-based kernel, we can see that $\Upsilon_u^{k_p} = \frac{\left(\mu_u^{*p}\right)^2}{\mathbb{E}\left[k_p\left(f^c(\mathbf{X}_u^w,\mathbf{U}),f^c(\mathbf{X}_u^{w'},\mathbf{U}')\right)\right]}$.

Lemma 1 guarantees the usual and interesting properties of sensitivity indices one should expect such as taking into account the interactions among the target inputs does not reduce the effects of the associated inputs. It also provides the upper bound of the total index of $\mathbf{X}_u^w$.

Lemma 1 *Consider a mensurable and SPD kernel k and assume that (A1)-(A2) hold. Then, we have*

$$0 \leq S_u^k \leq S_{T_u}^k \leq 1 \,, \tag{17}$$

$$S_{T_u}^k \leq \Upsilon_u^k \,. \tag{18}$$

Proof *See Supplementary Material C.*

It comes out from Lemma 1 that the upper bound of the total index given by (16) does not use the model derivatives. It is also worth noting that the numerator of the upper bound Υ_u^k relies on the difference of two model evaluations (i.e., f_u^*), while the numerator of the total index requires the evaluation of the conditional expectation of the target outputs. As a matter of fact, the computation of Υ_u^k should require less model evaluations to converge compared to the total index in general. Thus, Υ_u^k can be used for screening the target inputs. Furthermore, we can significantly reduce the number of model evaluations for computing Υ_u^k by making use of the Morris sequence (Morris, 1991) and the transformation $X_j = F_j^{-1}(U_j)$. Since the kernel-based SIs include Sobol' indices, generalized sensitivity indices (GSIs) and dependent GSIs (see Section 3.3), Υ_u^k is also the upper bounds of such indices.

3.2.2 Extension to multivariate and functional outputs

To deal with spatio-temporal model and dynamic models, we extend the kernel-based SIs from Definition 2 as follows:

$$S_u^{k,f} := \frac{\int_\Theta \mathbb{E}\left[k\left(f_u^{fo}(\mathbf{X}_u^w, \theta), f_u^{fo}(\mathbf{X}_u^{w'}, \theta)\right)\right] d\theta}{\int_\Theta \mathbb{E}\left[k\left(f^c(\mathbf{X}_u^w, \mathbf{U}, \theta), f^c(\mathbf{X}_u^{w'}, \mathbf{U}', \theta)\right)\right] d\theta} \,; \tag{19}$$

$$S_{T_u}^{k,f} := \frac{\int_\Theta \mathbb{E}\left[k\left(f_u^{to}(\mathbf{X}^w, \mathbf{U}, \theta), f_u^{tot}(\mathbf{X}_u^{w'}, \mathbf{U}', \theta)\right)\right] d\theta}{\int_\Theta \mathbb{E}\left[k\left(f^c(\mathbf{X}_u^w, \mathbf{U}, \theta), f^c(\mathbf{X}_u^{w'}, \mathbf{U}', \theta)\right)\right] d\theta} \,; \tag{20}$$

$$\Upsilon_u^{k,f} := \frac{\int_\Theta \mathbb{E}\left[k\left(f_u^*(\mathbf{X}_u^w, \mathbf{X}_u^{w'}, \mathbf{U}, \theta), f_u^*(\mathbf{X}_u^{w''}, \mathbf{X}_u^{w'''}, \mathbf{U}', \theta)\right)\right] d\theta}{\int_\Theta \mathbb{E}\left[k\left(f^c(\mathbf{X}_u^w, \mathbf{U}, \theta), f^c(\mathbf{X}_u^{w'}, \mathbf{U}', \theta)\right)\right] d\theta} \,. \tag{21}$$

Such indices also satisfy the usual and interesting properties provided in Lemma 1. Moreover, when $\Theta = \{\theta_0\}$, we can remark that $S_u^{k,f}, S_{T_u}^{k,f}$, and $\Upsilon_u^{k,f}$ come down to $S_u^k, S_{T_u}^k, \Upsilon_u^k$, respectively.

3.3 Link with existing global sensitivity indices

3.3.1 On the ℓ_1-based kernel indices and Morris-type screening methods

When $n = 1, k_1(y, y') := |y||y'|$, the numerator of $\Upsilon_j^{k_1}$ for any $j \in \{1, \dots, d\}$ is given by

$$\left(\mu_j^{*1}\right)^2 := \mathbb{E}^2\left[\left|f_j(\mathbf{X}_j^w, \mathbf{U}) - f_j(\mathbf{X}_j^{w'}, \mathbf{U})\right|\right].$$

Using the Morris sequence, we can see that the empirical expression of $\mu_j^{*1} = \mathbb{E}\left[\left|f_j(\mathbf{X}_j^w, \mathbf{U}) - f_j(\mathbf{X}_j^{w'}, \mathbf{U})\right|\right]$ is proportional to the Morris screening measure (noted μ_j^*) and its extension $d\mu_j^*$ provided in Lamboni & Kucherenko (2021) for dependent inputs. Thus, the Morris screening measure μ_j^* and its extension $d\mu_j^*$ are the upper bounds of $\sqrt{S_{T_j}^{k_1}}$ up to a constant of normalization. The normalized upper bound $\Upsilon_j^{k_1}$ enhances such screening measures. As a consequence, the Morris-type screening measures implicitly require working with the ℓ_1-norm or applying the ℓ_1-norm on SFs. Moreover, as small norms allow for measuring the small variations, the new ℓ_1-based kernel SIs and the Morris-type screening measures are well-suited for quantifying the small variations of the inputs effects.

3.3.2 Owen, Dick and Chen's L^p measure of dependence

For independent random variables and the kernel $k_{l_p}(Y, Y') := |Y|^{2p}|Y'|^{2p}$, we obtain the L^p-measure of dependence for the first-order only, which has been proposed in Owen et al. (2014), that is, $\sqrt{S_u^{k_{l_p}}} = \mathbb{E}\left[\left|f_u^{fo}(\mathbf{X}_u)\right|^{2p}\right]$. When $p = 1$ and inputs are independent, we obtain Sobol' indices.

3.3.3 The ℓ_2-based kernel indices: Sobol' indices and generalized sensitivity indices

When $n = 1$, $k_2(y, y') := |y|^2|y'|^2$ and the inputs are independent, we can check that the first-order and total Sobol indices are the square root of the ℓ_2-based kernel SIs, that is,

$$S_u = \sqrt{S_u^{k_2}}, \qquad S_{T_u} := \sqrt{S_{T_u}^{k_2}}.$$

Moreover, we have $\sqrt{S_{T_u}^{k_2}} = \frac{1}{2}\sqrt{\Upsilon_u^{k_2}}$.

For $n > 1$, $k_2(\mathbf{y}, \mathbf{y}') := \|\mathbf{y}\|_{L^2}^2 \|\mathbf{y}'\|_{L^2}^2$, we obtain the square of the generalized sensitivity indices (GSIs) of the first-type (see Lamboni et al., 2011; Lamboni, 2019, 2022b). Furthermore, when the target inputs are dependent, the ℓ_2-based kernel SIs are the square of the dependent GSIs (including GSIs) of the first-type introduced in Lamboni & Kucherenko (2021), that is,

$$dGSI_u^{1,M} = \sqrt{S_u^{k_2}} \qquad dGSI_{T_u}^{1,M} = \sqrt{S_{T_u}^{k_2}}.$$

3.3.4 Quadratic kernel indices and dependent generalized sensitivity indices

For $n \geq 1$, the kernel $k_q(\mathbf{y}, \mathbf{y}') := \langle \mathbf{y}, \mathbf{y}' \rangle^2$ leads to the square of i) GSIs of the second-type for independent variables (see Lamboni et al., 2011; Lamboni, 2019, 2022b), and ii) the dependent GSIs of the second-type (see Lamboni & Kucherenko, 2021; Lamboni, 2021), that is,

$$dGSI_u^{2,M} = \sqrt{S_u^{k_q}} \qquad dGSI_{T_u}^{2,M} = \sqrt{S_{T_u}^{k_q}} \leq \sqrt{\Upsilon_u^{k_p}}.$$

Thus, $\sqrt{\Upsilon_u^{k_p}}$ becomes the upper bound of the total GSIs of the second-type.

4. Computing the kernel-based sensitivity indices

For computing the first-order and total Kb-SIs in one hand, and the screening measure Υ_u^k in the other hand, we are going to express such indices using the weight function w_e and the CDF F_{ind} (see Proposition 3) because the dependency models of $\mathbf{X}^w \sim F_{ind}^{w_e}$ have been derived using w_e. To that end, we use $\mathbf{Y}'$, $\mathbf{Y}''$ and $\mathbf{Y}'''$ for i.i.d. copies of $\mathbf{Y} \sim F_{ind}$.

Proposition 3 *Assume that (A1)-(A2) hold. Then, we have*

$$S_u^k := \frac{\mathbb{E}\left[k\left(f_u^{fo}(\mathbf{Y}_u), f_u^{fo}(\mathbf{Y}_u')\right)w_e(\mathbf{Y})w_e(\mathbf{Y}')\right]}{\mathbb{E}\left[k\left(f^c(\mathbf{Y}_u, \mathbf{U}), f^c(\mathbf{Y}_u', \mathbf{U}')w_e(\mathbf{Y})w_e(\mathbf{Y}')\right)\right]}; \tag{22}$$

$$S_{T_u}^k := \frac{\mathbb{E}\left[k\left(f_u^{tot}(\mathbf{Y}_u, \mathbf{U}), f_u^{tot}(\mathbf{Y}_u', \mathbf{U}')\right)w_e(\mathbf{Y})w_e(\mathbf{Y}')\right]}{\mathbb{E}\left[k\left(f^c(\mathbf{Y}_u, \mathbf{U}), f^c(\mathbf{Y}_u', \mathbf{U}')\right)w_e(\mathbf{Y})w_e(\mathbf{Y}')\right]}; \tag{23}$$

$$\Upsilon_u^k := \frac{\mathbb{E}\left[k\left(f_u^*(\mathbf{Y}_u, \mathbf{Y}_u', \mathbf{U}), f_u^*(\mathbf{Y}_u'', \mathbf{Y}_u''', \mathbf{U}')\right)w_e(\mathbf{Y})w_e(\mathbf{Y}')w_e(\mathbf{Y}'')w_e(\mathbf{Y}''')\right]}{\mathbb{E}\left[k\left(f^c(\mathbf{Y}_u, \mathbf{U}), f^c(\mathbf{Y}_u', \mathbf{U}')\right)w_e(\mathbf{Y})w_e(\mathbf{Y}')\right]\mathbb{E}^2[w_e(\mathbf{Y})]}. \tag{24}$$

Proof *Using Definition 2, the results are straightforward bearing in mind Remark 1.*

To construct the estimators of the Kb-SIs, we are given two independent samples, that is, $\{\mathbf{Y}_i, \mathbf{Y}'_i, \mathbf{U}_i, \mathbf{U}'_i\}_{i=1}^{m_1}$ and $\{\mathbf{Y}_i, \mathbf{Y}'_i, \mathbf{U}_i, \mathbf{U}'_i\}_{i=1}^{m}$ from $(\mathbf{Y}, \mathbf{Y}', \mathbf{U}'\mathbf{U}')$, and we consider the following consistent estimators:

$$\hat{\mu}(\mathbf{Y}_u) := \frac{1}{m_1}\sum_{i=1}^{m_1} f(\mathbf{Y}_u, r(\mathbf{Y}_u, \mathbf{U}_i)) \xrightarrow{P} \mathbb{E}_{\mathbf{U}}[f(\mathbf{Y}_u, r(\mathbf{Y}_u, \mathbf{U}))] \; ;$$

$$\hat{\mu}(\mathbf{U}) := \frac{\sum_{i=1}^{m_1} f\left(\mathbf{Y}_{i,u}, r(\mathbf{Y}_{i,u}, \mathbf{U})\right) w_e(\mathbf{Y}_i)}{\sum_{i=1}^{m_1} w_e(\mathbf{Y}_i)} \xrightarrow{P} \frac{\mathbb{E}_{\mathbf{Y}}[f(\mathbf{Y}_u, r(\mathbf{Y}_u, \mathbf{U})) w_e(\mathbf{Y})]}{\mathbb{E}[w_e(\mathbf{Y})]} \; ,$$

$$\hat{\mu} := \frac{\sum_{i=1}^{m_1} f\left(\mathbf{Y}_{i,u}, r(\mathbf{Y}_{i,u}, \mathbf{U}_i)\right) w_e(\mathbf{Y}_i)}{\sum_{i=1}^{m_1} w_e(\mathbf{Y}_i)} \xrightarrow{P} \frac{\mathbb{E}[f(\mathbf{Y}_u, r(\mathbf{Y}_u, \mathbf{U})) w_e(\mathbf{Y})]}{\mathbb{E}[w_e(\mathbf{Y})]} \; .$$

Using the method of moments, we derive the estimators of the first-order and total SFs, the centered model outputs as follows:

$$\widehat{f_u^{fo}}(\mathbf{Y}_u) := \hat{\mu}(\mathbf{Y}_u) - \hat{\mu} \xrightarrow{P} f_u^{fo}(\mathbf{Y}_u) \; ;$$

$$\widehat{f_u^{\text{tot}}}(\mathbf{Y}_u, \mathbf{U}) := f(\mathbf{Y}_u, r(\mathbf{Y}_u, \mathbf{U})) - \hat{\mu}(\mathbf{U}) \xrightarrow{P} f_u^{\text{tot}}(\mathbf{Y}_u, \mathbf{U}) \; ;$$

$$\widehat{f^c}(\mathbf{Y}_u, \mathbf{U}) := f(\mathbf{Y}_u, r(\mathbf{Y}_u, \mathbf{U})) - \hat{\mu} \xrightarrow{P} f^c(\mathbf{Y}_u, \mathbf{U}) \; .$$

Again the method of moments and the plug-in approach allow for deriving the consistent estimators of the Kb-SIs in Theorem 2. To that end, we use

$$\sigma_u^{fo} := \mathbb{V}\left[k\left(f_u^{fo}(\mathbf{Y}_u), f_u^{fo}(\mathbf{Y}'_u)\right) w_e(\mathbf{Y}) w_e(\mathbf{Y}')\right] ,$$

$$\sigma_u^{tot} := \mathbb{V}\left[k\left(f_u^{tot}(\mathbf{Y}_u, \mathbf{U}), f_u^{tot}(\mathbf{Y}'_u, \mathbf{U}')\right) w_e(\mathbf{Y}) w_e(\mathbf{Y}')\right] ,$$

$$\mu_c^k := \mathbb{E}\left[k\left(f^c(\mathbf{Y}_u, \mathbf{U}), f^c(\mathbf{Y}'_u, \mathbf{U}')\right) w_e(\mathbf{Y}) w_e(\mathbf{Y}')\right] .$$

To derive the asymptotic distributions of the estimators of Kb-SIs, we suppose that we use a sample of size $M \gg m$ to estimate μ_c^k. Indeed, for estimating the non-normalized Kb-SIs of $\mathbf{X}_u$ for all $u \subseteq \{1, \dots, d\}$, different samples are going to be used, and some of such samples can be combined for estimating μ_c^k.

Theorem 2 *Assume that (A1)-(A2) hold and k is differentiable almost everywhere.*

(i) *The consistent estimators of S_u^k and $S_{T_u}^k$ are respectively given by*

$$\widehat{S_u^k} := \frac{\frac{1}{m}\sum_{i=1}^{m} k\left(\widehat{f_u^{fo}}(\mathbf{Y}_{i,u}), \widehat{f_u^{fo}}(\mathbf{Y}'_{i,u})\right) w_e(\mathbf{Y}_i) w_e(\mathbf{Y}'_i)}{\frac{1}{M}\sum_{i=1}^{M} k\left(\widehat{f^c}(\mathbf{Y}_{i,u}, \mathbf{U}_i), \widehat{f^c}(\mathbf{Y}'_{i,u}, \mathbf{U}'_i)\right) w_e(\mathbf{Y}_i) w_e(\mathbf{Y}'_i)} \; ; \tag{25}$$

$$\widehat{S_{T_u}^k} := \frac{\frac{1}{m}\sum_{i=1}^{m} k\left(\widehat{f_u^{\text{tot}}}(\mathbf{Y}_{i,u}, \mathbf{U}_i), \widehat{f_u^{\text{tot}}}(\mathbf{Y}'_{i,u}, \mathbf{U}'_i)\right) w_e(\mathbf{Y}_i) w_e(\mathbf{Y}'_i)}{\frac{1}{M}\sum_{i=1}^{M} k\left(\widehat{f^c}(\mathbf{Y}_{i,u}, \mathbf{U}_i), \widehat{f^c}(\mathbf{Y}'_{i,u}, \mathbf{U}'_i)\right) w_e(\mathbf{Y}_i) w_e(\mathbf{Y}'_i)} \; . \tag{26}$$

(ii) *If $m_1, m, M \to \infty$ with $\frac{m}{M} \to 0; \frac{m_1}{M} \to 0$, then*

$$\sqrt{m}(\widehat{S_u^k} - S_u^k) \xrightarrow{D} \mathcal{N}\left(0, \frac{\sigma_u^{fo}}{(\mu_c^k)^2}\right); \qquad \sqrt{m}\left(\widehat{S_{T_u}^k} - S_{T_u}^k\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{\sigma_u^{\text{tot}}}{(\mu_c^k)^2}\right).$$

Proof *See Supplementary Material D.*

5. Test cases

We consider two functions to illustrate our approach. The first test case treats a cluster of the output values by making use of the CDF W (see Theorem 1) to derive analytical target input distributions and SFs. The second test case deals with a vector-valued function and a polynomial weight function. To compute the kernel-based SIs, we consider the ℓ_1-based kernel and the quadratic kernel.

5.1 Quadratic function ($d = 3, \ n = 1$)

In this section, we consider the following single response model:

$$f(\mathbf{X}) = X_1^2 + X_2^2 + X_3^2,$$

which includes three independent input factors $X_j \sim \mathcal{N}(0,1)$ with $j = 1,2,3$. For a given threshold $c \in \mathbb{R}_+$, we are interested in the target input distribution given by

$$dF^w(\mathrm{x}) = \frac{\mathbb{1}_{]-\infty,c]}(f(\mathrm{x}))}{\mathbb{E}_{F_{ind}}\left[\mathbb{1}_{]-\infty,c]}(f(\mathbf{X}))\right]} dF_{ind}(\mathrm{x}); \ \forall \mathrm{x} \in \mathbb{R}^d.$$

Such model behavior represents a safe domain such as the concentration of a biological hazard below a given threshold c, that is, $f(\mathbf{X}) \leq c$. To derive the analytical SFs, it is known that a dependency model of

$$\mathbf{X}^w := (X_1^w, X_2^w, X_3^w) := \left\{ X_j \sim \mathcal{N}(0,1), j = 1,2,3 : X_1^2 + X_2^2 + X_3^2 \leq c \right\}$$

is given by $(X_2^w)^2 = Z_2(c - (X_1^w)^2)$, $(X_3^w)^2 = Z_3(c - (X_1^w)^2)(1 - Z_2)$, where $(X_1^w)^2 \sim B1(c, 1/2, 2)$, $Z_2 \sim Beta(1/2, 3/2)$ and $Z_3 \sim Beta(1/2, 1)$ are independent with $B1$ the beta distribution of first-kind (see Lamboni, 2022a), Corollaries 2, 4 for more details). Thus, we have (see Supplementary Material E)

$$f(\mathbf{X}^w) := (X_1^w)^2\left(1 - Z_2 - Z_3(1 - Z_2)\right) + cZ_2 + cZ_3(1 - Z_2);$$

$$f^c(X_1^w, Z_2, Z_3) = (X_1^w)^2\left(1 - Z_2 - Z_3(1 - Z_2)\right) + cZ_2 + cZ_3(1 - Z_2) - \frac{3}{5}c.$$

Moreover, the first-order and total SFs of X_1^w and f^* are given by

$$f_1^{fo}(X_1^w) = \frac{1}{2}[(X_1^w)^2 - c/5]; \qquad f_1^{tot}(X_1^w, Z_2, Z_3) = [(X_1^w)^2 - c/5](1 - Z_2)(1 - Z_3);$$

$$f^*\left(X_1^w, X_1^{w\prime}, Z_2, Z_3\right) = \left[(X_1^w)^2 - \left(X_1^{w\prime}\right)^2\right](1 - Z_2)(1 - Z_3).$$

The computations of Kb-SIs using the above functionals give the following results. For the ℓ_1-based kernel, the indices of X_j^w are $\sqrt{S_j^{k_1}} = \sqrt{S_{T_j}^{k_1}} = 0.385$ and $\sqrt{\Upsilon_j^{k_1}} = 0.505$ with $j = 1, 2, 3$. For the quadratic kernel, we have obtained $\sqrt{S_j^{k_q}} = 0.167 \sqrt{S_{T_j}^{k_q}} = 0.223$ and $\sqrt{\Upsilon_j^{k_q}} = 0.445$ with $j = 1, 2, 3$. These indices are invariant by changing the threshold c.

5.2 Multivariate g-Sobol function ($d = 10, \ n = 4$)

We consider the multivariate g-Sobol function, which includes ten independent input variables following the standard uniform distribution (i.e., $X_j \sim \mathcal{U}(0,1), j = 1, \dots, 10$) and provides four outputs. It is given by (Lamboni, 2019, 2022b)

$$f(\mathbf{x}) := \begin{bmatrix} \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[1,j]}{1 + \mathcal{A}[1,j]} \\ \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[2,j]}{1 + \mathcal{A}[2,j]} \\ \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[3,j]}{1 + \mathcal{A}[3,j]} \\ \prod_{j=1}^{d=10} \frac{|4x_j - 2| + \mathcal{A}[4,j]}{1 + \mathcal{A}[4,j]} \end{bmatrix}, \mathcal{A} = \begin{bmatrix} 0 & 0 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 \\ 0 & 1 & 4.5 & 9 & 99 & 99 & 99 & 99 & 99 & 99 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 \end{bmatrix}.$$

We are interested in the model behavior defined by $w(\mathbf{x}) := \prod_{j=1}^{10} x_j^{\alpha_j}$ for every $\alpha_j \in \mathbb{R}_+$. We can see that the target inputs distribution is given by $\rho^w(\mathbf{x}) = \dfrac{\prod_{j=1}^{10} x_j^{\alpha_j} \mathbb{1}_{[0,1]}(x_j)}{\mathbb{E}_{F_{ind}}\left[\prod_{j=1}^{10} X_j^{\alpha_j}\right]}$, which implies that $X_j^w \sim Beta(\alpha_j + 1, 1)$ for any $j \in \{1, \ldots, d\}$.

Tables 1-2 provide the estimates of sensitivity indices using the ℓ_1-based kernel and the quadratic kernel. Obviously, when $q = 1/2$ and $\alpha = (0, \ldots, 0)$ (see Table 1), we obtain the estimations of GSIs provided in Lamboni (2022b) for the quadratic kernel. Only X_1 and X_2 are important according to the total GSIs and their upper bounds. Similar results were obtained using the ℓ_1-based kernel SIs if we fix the threshold at $T = 0.2$. Decreasing the threshold to 0.1 shows that two more inputs, that is, X_3, X_4 are important. Such difference is probably due to the fact that the ℓ_1-based kernel measures the small variations of SFs compared to the quadratic kernel, which is associated with the L_2-norm for single valued functions.

Table 1: Square root of kernel-based SIs for the multivariate Sobol function associated with $\alpha = (0, \ldots, 0)$.

Kernels	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
					First-order Kb-SIs					
ℓ_1-based	0.756	0.556	0.176	0.136	0.099	0.092	0.086	0.082	0.079	0.076
Quadratic	0.536	0.328	0.027	0.016	0.011	0.009	0.008	0.007	0.007	0.006
					Total Kb-SIs					
ℓ_1-based	0.756	0.555	0.178	0.148	0.099	0.091	0.085	0.082	0.078	0.076
Quadratic	0.637	0.436	0.038	0.026	0.015	0.012	0.011	0.011	0.009	0.010
					Upper bounds of Kb-SIs					
ℓ_1-based	1.002	0.747	0.235	0.188	0.133	0.124	0.114	0.114	0.105	0.100
Quadratic	1.230	0.844	0.074	0.048	0.031	0.027	0.021	0.024	0.020	0.018

By taking $\boldsymbol{\alpha} = (20, 20, 10, 10, 10, 10, 10, 1, 1, 1)$, the ℓ_1-based kernel SIs show that all the inputs are important ($T = 0.1$) while the quadratic kernel SIs identify $X_j^w, j = 3, \ldots, 7$ as non important variables (see Table 2). As expected, the results provided in Tables 1-2 prove that changing the model behavior of interest (thanks to the weight functions) can significantly modify the impacts of the input variables. The polynomial weight function $x_j^{\alpha_j}$ aims at emphasizing the selection of the initial input values of X_j that are close to 1 when α_j is higher. Thus, it restricts the support of X_j, which results in decreasing the effect of X_j^w on the model outputs.

Table 2: Square root of kernel-based SIs for the multivariate Sobol function associated with $\alpha = (20, 20, 10, 10, 10, 10, 10, 1, 1, 1)$.

Kernels	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
					First-order Kb-SIs					
ℓ_1-based	0.465	0.388	0.277	0.229	0.173	0.165	0.157	0.319	0.309	0.301
Quadratic	0.261	0.190	0.075	0.058	0.041	0.040	0.035	0.121	0.116	0.113
					Total Kb-SIs					
ℓ_1-based	0.465	0.388	0.276	0.229	0.173	0.164	0.157	0.320	0.309	0.300
Quadratic	0.267	0.197	0.074	0.059	0.043	0.041	0.037	0.123	0.120	0.116
					Upper bounds of Kb-SIs					
ℓ_1-based	0.615	0.530	0.374	0.303	0.238	0.224	0.211	0.428	0.417	0.402
Quadratic	0.477	0.441	0.148	0.104	0.088	0.081	0.069	0.252	0.252	0.233

6. Conclusion

We have proposed a methodology for exploring specific model behaviors such as identifying the input variables that drive the initial model output(s) toward a domain of interest (e.g., failure, safe and sustainable domain) and/or govern the target output(s) defined via weight functions, including the output values within a given cluster. Weight functions include rule-based ensembles, non-negative desirability measures of the model output(s) to meet a given criterion, membership functions from crisp or fuzzy clustering, and any non-negative

function based on classifiers such as PCA, kernel PCA, logistic regression, random forest. The proposed approach is well-suited for performing SA not only for the target output(s) but also for models with inputs following the multivariate weighted distribution, including the copula-based distributions.

We have introduced the kernel-based SIs, including the ℓ_1-based SIs that aim at addressing the issues of high-order moments of SFs and interactions among the model drivers even non-independent. The well-known variance-based SA and dependent multivariate SA are particular cases of our new measures of association between model drivers and outputs. We have provided consistent estimators of kernel-based SIs for computing such indices by distinguishing the case of multivariate and functional outputs (including spatio-temporal and dynamic models) and the case of multivariate response models, including single response models.

Analytical and numerical results show the relevance of our approach for analyzing a given model behavior. For the second test case, our results are equal to those provided in Lamboni (2022b); Lamboni & Kucherenko (2021); Lamboni (2021), emphasizing the extension of generalized sensitivity indices and dependent generalized sensitivity indices. In next future, it is worth investigating the extension of this work to cope with most kernels and the original space of labels used as outputs.

Acknowledgements

Supplementary Material

The Supplementary Material for this article can be found online at https://sesmo.org/article/view/18566/18117.

References

Arnold, B. C., Ghosh, I., & Alzaatreh, A. (2017). Construction of bivariate and multivariate weighted distributions via conditioning. Communications in Statistics - Theory and Methods, 46(18): 8897–8912.

Aronszajn, N. (1950). Theory of reproducing kernels. Transactions of the American Mathematical Society, 68: 337–404.

Berlinet, A., Thomas, C., & Gnan, T. A. (2004). Reproducing Kernel Hilbert Space in probability and statistics. Kluwer Academic.

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). Fcm: The fuzzy c-means clustering algorithm. Computers & Geosciences, 10(2-3):191–203.

Borgonovo, E., Tarantola, S., Plischke, E., & Morris, M. D. (2014). Transformations and invariance in the sensitivity analysis of computer experiments. Journal of the Royal Statistical Society Series B, 76(5): 925–947.

Campbell, K., McKay, M. D., & Williams, B. J. (2006). Sensitivity analysis when model outputs are functions. Reliability Engineering & System Safety, 91(10-11): 1468 – 1472.

Chevalier, C., Ginsbourger, D., Bect, J., & Molchanov, I. (2013). Estimating and quantifying uncertainties on level sets using the vorob'ev expectation and deviation with gaussian process models. In: Ucinski D., Atkinson A., Patan M. (eds) mODa 10 - Advances in Model-Oriented Design and Analysis. Contributions to Statistics, pages 35–43.

Elsawah, S., Filatova, T., Jakeman, A., Kettner, A., Zellner, M., Athanasiadis, I., Hamilton, S., Axtell, R., Brown, D., Gilligan, J., Janssen, M., Robinson, D., Rozenberg, J., Ullah, I., & Lade, S. (2020). Eight grand challenges in socio-environmental systems modeling. Socio-Environmental Systems Modelling, 2, 16226.

Fisher, R. A. (1934). The effect of methods of ascertainment upon the frequencies. Ann. Eugen., 6: 13–25.

Fort, J.-C., Klein, T., & Lagnoux, A. (2021). Global sensitivity analysis and Wasserstein spaces. SIAM/ASA Journal on Uncertainty Quantification, 9(2): 880–921.

Fossum, T. O., Travelletti, C., Eidsvik, J., Ginsbourger, D., & Rajan, K. (2020). Learning excursion sets of vector-valued gaussian random fields for autonomous ocean sampling. ArXiv, abs/2007.03722.

French, J. P. & Sain, S. R. (2013). Spatio-temporal exceedance locations and confidence regions. The Annals of Applied Statistics, 7(3): 1421–1449.

Gamboa, F., Janon, A., Klein, T., & Lagnoux, A. (2014). Sensitivity analysis for multidimensional and functional outputs. Electron. J. Statist., 8(1): 575–603.

Höppner, F., Klawonn, F., Kruse, R., & Runkler, T. (1999). Fuzzy cluster analysis: methods for classification, data analysis and image recognition. John Wiley & Sons.

Jain, K. & Nanda, A. K. (1995). On multivariate weighted distributions. Communications in Statistics Theory and Methods, 24(10): 2517–2539.

Joe, H. (2014). Dependence Modeling with Copulas. Chapman & Hall/CRC, London.

Kocherlakota, S. (1995). Discrete bivariate weighted distributions under multiplicative weight function. Communications in Statistics - Theory and Methods, 24(2):533–551.

Lamboni, M. (2016). Global sensitivity analysis: a generalized, unbiased and optimal estimator of total-effect variance. Statistical Papers, pages 1–26.

Lamboni, M. (2019). Multivariate sensitivity analysis: Minimum variance unbiased estimators of the first-order and total-effect covariance matrices. Reliability Engineering & System Safety, 187: 67 – 92.

Lamboni, M. (2020a). Derivative-based generalized sensitivity indices and Sobol' indices. Mathematics and Computers in Simulation, 170: 236 – 256.

Lamboni, M. (2020b). Uncertainty quantification: a minimum variance unbiased (joint) estimator of the non-normalized Sobol' indices. Statistical Papers, 61: 1939–1970.

Lamboni, M. (2021). On dependency models and dependent generalized sensitivity indices. arXiv preprint arXiv2104.12938.

Lamboni, M. (2022a). Efficient dependency models: Simulating dependent random variables. Mathematics and Computers in Simulation, 200: 199–217.

Lamboni, M. (2022b). Weak derivative-based expansion of functions: ANOVA and some inequalities. Mathematics and Computers in Simulation, 194: 691–718.

Lamboni, M. (2023). On exact distribution for multivariate weighted distributions and classification. Methodology and Computing Applied Probability, 25: 1–41.

Lamboni, M., Koeble, R., & Leip, A. (2016). Multi-scale land-use disaggregation modelling: Concept and application to EU countries. Environmental Modelling & Software, 82: 183–217.

Lamboni, M. & Kucherenko, S. (2021). Multivariate sensitivity analysis and derivative-based global sensitivity measures with dependent variables. Reliability Engineering & System Safety, 212: 107519.

Lamboni, M., Monod, H., & Makowski, D. (2011). Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. Reliability Engineering and System Safety, 96: 450–459.

Lamboni, M., Sanaa, M., & Tenenhaus-Aziza, F. (2014). Sensitivity analysis for critical control points determination and uncertainty analysis to link FSO and process criteria: Application to listeria monocytogenes in soft cheese made from pasteurized milk. Risk Analysis, 34(4): 751–764.

Mahfoud, M. & Patil, G. (1982). On weighted distributions, pages 479–492. Institute of Mathematical Statistics, Amsterdam: North Holland.

Mara, T. A., Tarantola, S., & Annoni, P. (2015). Non-parametric methods for global sensitivity analysis of model output with dependent inputs. Environmental Modelling & Software, 72: 173 – 183.

Morris, M. (1991). Factorial sampling plans for preliminary computational experiments. Technometrics, 33: 161–174.

Navarro, J., Ruiz, J. M., & Aguila, Y. D. (2006). Multivariate weighted distributions: a review and some extensions. Statistics, 40(1): 51–64.

Nelsen, R. (2006). An Introduction to Copulas. Springer-Verlag, New York.

Owen, A. (2014). Sobol' indices and Shapley value. Journal on Uncertainty Quantification, 2: 245–251.

Owen, A. B., Dick, J., & Chen, S. (2014). Higher order Sobol' indices. Information and Inference: A Journal of the IMA, 3(1): 59–81.

Patil, G. P. & Rao, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. Biometrics, 34(2): 179–189.

Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment. Sankhya: The Indian Journal of Statistics, Series A (1961-2002), 27(2/4):311–324.

Rose, K. A., Smith, E. P., Gardner, R. H., Brenkert, A. L., & Bartell, S. M. (1991). Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty. Journal of Forecasting, 10: 117–133.

Roux, S., Buis, S., Lafolie, F., & Lamboni, M. (2021). Cluster-based gsa: Global sensitivity analysis of models with temporal or spatial outputs using clustering. Environmental Modelling & Software, 140: 105046.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). Global Sensitivity Analysis: the Primer. Wiley, New York.

Saltelli, A. & Tarantola, S. (2002). On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal. Journal of the American Statistical Association, 97: 702–709.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). Sensitivity analysis in practice. Probability and Statistics. John Wiley and Sons, Chichester.

Schölkopf, B. & Smola, A. J. (2002). Learning with Kernels. MIT Press, Cambridge, MA.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut Statistique de l'Université de Paris, 8: 229–231.

Skorohod, A. V. (1976). On a representation of random variables. Theory Probability and its Applications, 21(3): 645–648.

Sobol, I. M. (1993). Sensitivity analysis for non-linear mathematical models. Mathematical Modelling and Computational Experiments, 1: 407–414.

Spear, R. & Hornberger, G. (1980). Eutrophication in peel inlet-II. identification of critical uncertainties via generalized sensitivity analysis. Water Research, 14(1): 43–49.

Xiao, S., Lu, Z., & Wang, P. (2018). Multivariate global sensitivity analysis for dynamic models based on wavelet analysis. Reliability Engineering & System Safety, 170: 20–30.