



HAL
open science

Measuring inputs-outputs association for time-dependent hazard models under safety objectives using kernels

Matieyendou Lamboni

► **To cite this version:**

Matieyendou Lamboni. Measuring inputs-outputs association for time-dependent hazard models under safety objectives using kernels. IJUQ, In press. hal-04621193

HAL Id: hal-04621193

<https://hal.science/hal-04621193v1>

Submitted on 23 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Measuring inputs-outputs association for time-dependent hazard models under safety objectives using kernels

Matieyendou Lamboni^{1a,b}

^a*University of Guyane, Department DFRST, 97346 Cayenne, French Guiana, France*

^b*228-UMR Espace-Dev, University of Guyane, University of Réunion, IRD, University of Montpellier, France.*

Abstract

A methodology for assessing the inputs-outputs association for time-dependent predictive models subjected to safety objectives is investigated. Firstly, new dependency models for sampling random values of uncertain inputs that comply with the safety objectives are provided by making use of the desirability measures. Secondly, combining predictive risk-models with such dependency models leads to the development of new kernel-based statistical tests of independence between the (safe) dynamic outputs and inputs. The associated test statistics are then normalized so as to introduce the first-order and total sensitivity indices that account for the desirability measures. Such indices rely on time-dependent sensitivity functionals (SFs) and kernel methods, which allow for treating non-stationary SFs as well as SFs having skewed or heavy-tailed distributions. Our approach is also well-suited for dynamic hazard models with prescribed copulas of inputs.

Keywords: Copulas, Independence tests, Spatio-temporal models, Uncertainty quantification, Weighted distributions.

1. Introduction

Integrated dynamic systems are widely used in engineering for quantitative risk assessment and management. Monitoring and controlling an identified hazard at different time steps are crucial in quantitative risk analysis, including quantitative microbiological risk analysis, so as to bring the systems under control, to prevent natural disasters, and to minimize damages. It is known that quantitative risk assessment (QRA) is one of the most challenging steps in risk analysis when relevant dataset is available. Some QRA approaches aim at evaluating directly the risk output, that is, the probability of occurrence of hazards. Other QRA methods rely on predictive models of the level of hazards at different time steps. Among others,

¹Corresponding author: matieyendou.lamboni[at]gmail.com or [at]univ-guyane.fr; Feb. 27, 2024

site-specific hazard curves for earthquake damage management ([1, 2]); the equity capital of a firm (which should be positive); the Black-Scholes dynamic model for a portfolio value in finance; process-risk models that simulate the concentrations of microbiological hazards ([3, 4, 5]), socio-environmental models ([6]) and flood-risk models ([7, 8, 9]) are examples.

For risk management based on predictive models, making decisions often relies on contextual indicators, involving the hazard levels and thresholds that are used for defining critical and/or failure conditions. Such thresholds (also known as safety objectives) are provided by competent authorities or are available in the literature or can be estimated as the values at risk using previous data. Contextual indicators represent specific system behaviors such as the safety of the products, structural failures, safety systems, and sustainability. Conceptually, only outputs of predictive models that comply with the safety objectives or not are going to be considered, depending on the purpose of the analysis. Mathematically, it consists in clustering the dynamic outputs of hazard models using either the binary classification or the fuzzy classification. Reliable predictive models rely on a suitable choice of parameters and/or input variables (e.g., risk factors) for describing an hazard and its evolution over time, and such inputs are often uncertain. Denote with \mathbf{X} such uncertain inputs.

Performing uncertainty quantification (UQ) and sensitivity analysis (SA) for assessing the strengths of un-controllable and/or controllable inputs (e.g., policy and management levers) on the dynamic outputs subjected to safety objectives (i.e., safe outputs) are worth investigating, but require measuring the ability of each dynamic output or curve to meet the safety objectives ([5]). There are different possibilities to measure such abilities, depending on risk managers. To define one reasonable desirability measure, denote with \mathbf{x} a sample point of the inputs \mathbf{X} of the hazard model given by $Y(t) := f(\mathbf{X}, t)$ with $t \in [0, T]$ and T a time period. Obviously, each sample point will not lead to meet a safety objective (noted S), that is, ($y(t) = f(\mathbf{x}, t) \leq S, \forall t \in [0, T]$). Moreover, two different sample points that meet a safety objective can perform differently. Thus, to bring the system under control, the following desirability measure is considered:

$$D(\mathbf{x}) := \begin{cases} \int_0^T |f(\mathbf{x}, t) - S| dt & \text{if } y(t) \leq S \quad \forall t \in [0, T] \\ 0 & \text{otherwise} \end{cases} . \quad (1)$$

The above desirability measure can be seen as a mixture of the crisp and fuzzy clustering membership functions of the output values according to their abilities to meet the safety objective. Regarding the theory of multivariate weighted distribu-

tions (e.g., [10, 11, 12]), $D(\mathbf{x})$ gives different chances to the sample points of being included in UQ and SA, and it is also known as the weight function. For instance, such a desirability measure includes the payoff of the European call option, that is, $\max\{f(\mathbf{x}, T) - K, 0\}$ with $f(\mathbf{x}, T)$ the price at maturity T and K a strike.

While other desirability measures can be considered, it is worth noting that taking account of such measures is going to modify the initial distribution of the model inputs by introducing or adding new dependence structures (see [5, 13]). Thus, performing UQ and SA require the distributions of the model inputs that account for every desirability measure (i.e., new joint distribution of the inputs). Such joint distribution functions have been derived in [13, 14] as well as the conditional distribution functions and the dependency models, that is, regression-based representations of random vectors following such distributions (see also [15, 16, 17]). Dependency models have been used for developing dependent multivariate sensitivity analysis (DMSA) for models with dependent and/or correlated input variables ([15, 16, 17, 13]), and those derived in [13, 14] rely on the weighted distribution associated with the initial distribution function of inputs rather than the margins of the new joint distribution. Although different statistical moments, including the first-order and total dependent generalized sensitivity indices, can be computed using the weighted distributions ([13, 14]), generating directly random values according to the new joint distribution is much convenient for the sample-based risk assessment.

Moreover, among global SA (GSA) (e.g., [18, 19, 20, 21, 22, 23, 24, 25, 26, 15, 16, 27, 28, 29, 30]), DMSA ([15, 16, 13, 14]), including multivariate sensitivity analysis (MSA, [31, 32, 28, 33, 24, 29]), relies on sensitivity functionals (SFs). Sensitivity functionals for dynamic models are time-dependent random processes that contain different information about the contrition of each input (including interactions) on the whole dynamic outputs. Since MSA and DMSA make use of variance as the importance measure, the resultant sensitivity indices account only for the second-order moments of SFs, and such indices may not exist for some distributions of SFs such as Cauchy's distribution. Therefore, DMSA may be inefficient for capturing all the moments or sufficient moments or desirable moments of SFs associated with the safe outputs because such SFs can follow skewed and heavy-tailed distributions and/or may be non-stationary ([34]).

This paper aims at providing measures of association between inputs and outputs for time-dependent predictive models that incorporate the desirability measures or comply with the safety objectives. Our contributions are twofold: firstly, dependency

models are derived in Section 2 for sampling the random values of inputs that follow the new joint distribution without using the weighted distributions. Note that the new distribution is often that of non-independent variables, keeping in mind Equation (1).

Secondly, by combining predictive models with such dependency models, we propose new dependence measures of inputs-outputs association that rely on statistical tests of independence between the dynamic outputs and inputs. The asymmetric role of the outputs and inputs motivates the construction of our test statistics by making use of time-dependent SFs. Moreover, to account for relevant statistical properties of SFs, SFs are going to be embedded into the appropriate Reproducing Kernel Hilbert Space (RKHS) ([35, 36]) by properly choosing the kernels (see Section 3). The kernel-based test statistics are normalized in Section 4.2 to provide the kernel-based sensitivity indices (Kb-SIs) for time-dependent models of hazards subjected to safety objectives. Such new measures (i.e., Kb-SIs) satisfy the Renyi axioms ([37]) and lead to a statistical foundation of sensitivity indices for time-dependent models or processes. Consistent estimators of the first-order and total Kb-SIs are derived in Section 5, and a case study is provided in Section 6. We conclude this work in Section 7.

General notation

For an integer $d \in \mathbb{N} \setminus \{0\}$, $\mathbf{X} := (X_1, \dots, X_d)$ denotes a random vector of the model inputs. For $u \subseteq \{1, \dots, d\}$, we use $\mathbf{X}_u := (X_j, \forall j \in u)$; $\mathbf{X}_{\sim u} := (X_j, \forall j \in \{1, \dots, d\} \setminus u)$ and we have the following partition $\mathbf{X} = (\mathbf{X}_u, \mathbf{X}_{\sim u})$. We use $\|\cdot\|_p$ for the p -norm, $\mathbb{E}[\cdot]$ for the expectation and $\mathbb{V}[\cdot]$ for the variance-covariance.

2. Distribution of inputs under safety objectives and dependency models

This section deals with the distribution of the inputs that comply with a safety objective and the statistical representation of the inputs, following that distribution. Given $t \in [0, T]$, the time-dependent model $f(\cdot, t) : \mathbb{R}^d \rightarrow \mathbb{R}$ includes \mathbf{X} as inputs and provides $Y(t) = f(\mathbf{X}, t)$ as the output at time t .

Without loss of generality, let us consider a safety condition, which is often described by $(Y(t) \leq S, \forall t \in [0, T])$ with S the safety objective for an identified hazard. Taking account of the desirability measure given by Equation (1) or every non-negative desirability measure to bring the system under control is going to modify the initial distribution of the model inputs. Indeed, the resultant distribution

is going to avoid the failed input values (i.e., \mathbf{x}_f) defined by

$$\{\mathbf{x}_f \in \mathbb{R}^d : \exists t_0 \in [0, T] \text{ such that } f(\mathbf{x}_f, t) > S\},$$

by giving them a zero chance of being included in further analysis under safety objectives. The resultant distribution also gives different or equal weights to other (safe) input values \mathbf{x}_s , satisfying ($f(\mathbf{x}_s, t) \leq S, \forall t \in [0, T]$). Nevertheless, the above failure condition is of great interest since it can be used for identifying the key inputs that drive a system to fail. While in what follows, we are going to treat the safety condition, it is worth noting that the following proposed methodology is well-suited for dealing with failure conditions as well by just replacing Equation (1) with a failure-desirability measure.

Formally, let $\rho : \mathbb{R}^d \rightarrow \mathbb{R}_+$ be the initial probability density function (PDF) or mass function of \mathbf{X} and $D : \mathbb{R}^d \rightarrow \mathbb{R}_+$ a desirability measure. The PDF or mass function of the inputs that comply with the safety objective is given by (see e.g., [10, 11, 12])

$$\rho^s(\mathbf{x}) := \frac{\rho(\mathbf{x}) D(\mathbf{x})}{\mathbb{E}[\rho(\mathbf{X}) D(\mathbf{X})]}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

To avoid any ambiguity, let us use $\mathbf{X}^s := (X_1^s, \dots, X_d^s)$ for a d -dimensional random vector of inputs that have $\rho^s(\mathbf{x})$ as PDF or mass function. Obviously, the random vector \mathbf{X}^s will lead to meet the safety objective, while \mathbf{X} will not in general. To generate random values of \mathbf{X}^s in general, the conditional sampling algorithm (see [38, 39, 15, 16] for more details) requires knowing the exact cumulative distribution function (CDF) of \mathbf{X}^s . By denoting with F_1, \dots, F_d the margins of the distribution of \mathbf{X} , the exact CDF of \mathbf{X}^s (i.e., F^s) is given by (see Corollary 2 in [13])

$$F^s(\mathbf{x}) := W(F_1(x_1), \dots, F_d(x_d)), \quad \forall \mathbf{x} \in \mathbb{R}^d, \quad (2)$$

where $W : [0, 1]^d \rightarrow [0, 1]$ is a distribution function, which is similar to a copula (see Equation (4)). If we use F_j^{\leftarrow} for the generalized inverse of F_j with $j = 1, \dots, d$, then $W(\cdot)$ is given by

$$W(u_1, \dots, u_d) = \frac{\mathbb{E}_{\mathbf{V}}[w_e(F_1^{\leftarrow}(V_1), \dots, F_d^{\leftarrow}(V_d))]}{\mathbb{E}_{\mathbf{U}}[w_e(F_1^{\leftarrow}(U_1), \dots, F_d^{\leftarrow}(U_d))]} \prod_{k=1}^d u_k, \quad (3)$$

where $V_j \sim \mathcal{U}(0, u_j)$; $U_j \sim \mathcal{U}(0, 1)$ and $\mathbb{E}_{\mathbf{V}}$ (resp. $\mathbb{E}_{\mathbf{U}}$) means that the expectation is taken w.r.t. \mathbf{V} (resp. \mathbf{U}). The weight function w_e used in (3) comes down to the desirability measure when \mathbf{X} is a random vector of independent variables. In

general, the weight function $w_e(\cdot)$ is defined as follows ([13]):

$$w_e(\mathbf{x}) = \begin{cases} D(\mathbf{x}) & \text{if } \rho(\mathbf{x}) = \prod_{j=1}^d \rho_j(x_j) \\ c(F_1(x_1); \dots, F_d(x_d)) D(\mathbf{x}) & \text{if } c(\cdot) \text{ is the copula PDF of } \mathbf{X} \text{ ,} \\ w(\mathbf{x}) D(\mathbf{x}) & \text{if } \rho(\mathbf{x}) = w(\mathbf{x}) \prod_{j=1}^d \rho_j(x_j) \end{cases}$$

with $w(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}_+$ and $\rho_j(\cdot)$ the marginal PDF or mass function of X_j . Proposition 1 shows that W given by (3) is the CDF of a known random vector.

Proposition 1. *Consider $\mathbf{X}^s = (X_1^s, \dots, X_d^s) \sim F^s$. Then, we have*

$$(F_1(X_1^s), \dots, F_d(X_d^s)) \sim W. \quad (4)$$

Proof. See Appendix A. □

In view of Proposition 1, for every desirability measure or weight function, there exists always a CDF with margins supported on $[0, 1]$ such that the random vector $(F_1(X_1^s), \dots, F_d(X_d^s))$ follows that distribution function. Contrary to copulas, the marginal CDFs of W are not necessarily the uniform distribution. Using the CDF of \mathbf{X}^s given by Equation (2), we are able to derive its marginal CDFs (i.e., $X_j^s \sim F_j^s$) and the generalized inverse of F_j^s (i.e., $F_j^{s\leftarrow}$) that rely on $W(\cdot)$. The following propositions deal with such issues.

Proposition 2. *For any $j \in \{1, \dots, d\}$, the j^{th} margin of the joint CDF W is*

$$W_j(u_j) := W(1, \dots, u_{j-1} = 1, u_j, u_{j+1} = 1, \dots, 1), \quad \forall u_j \in [0, 1].$$

The proof is straightforward keeping in mind that W is a probability (see Appendix B for instance). Denote with $W_j^{s\leftarrow}$ the generalized inverse of W_j .

Proposition 3. *Consider $U_j \sim \mathcal{U}(0, 1)$ and $x_j \in \mathbb{R}$. Then, we have*

$$F_j^s(x_j) = W_j(F_j(x_j)); \quad X_j^s = F_j^{s\leftarrow}(U_j) = F_j^{s\leftarrow}(W_j^{s\leftarrow}(U_j)) \text{ a.s.} \quad (5)$$

Proof. See Appendix B. □

Based on the above marginal CDFs, a new expression of the copula of $\mathbf{X}^s \sim F^s$ is derived in the following corollary. To that end, let $\text{Ran}F_j^s := \{F_j^s(x_j) : x_j \in \mathbb{R}\}$ denote the range of F_j^s for any $j \in \{1, \dots, d\}$.

Corollary 1. *Consider $u_j \in \text{Ran}F_j^s$ with $j = 1, \dots, d$. Then, the copula of \mathbf{X}^s is*

$$C(u_1, \dots, u_d) := W(W_1^{s\leftarrow}(u_1), \dots, W_d^{s\leftarrow}(u_d)). \quad (6)$$

Proof. See Appendix C. □

Note that W becomes a copula when $W_j(u_j) = u_j$ for any $j \in \{1, \dots, d\}$. Using the known CDFs W and C , the fact that $(F_1(X_1^s), \dots, F_d(X_d^s)) \sim W$ (see Equation (4)), $(F_1^s(X_1^s), \dots, F_d^s(X_d^s)) \sim C$, and bearing in mind the conditional sampling algorithm (see [38, 39, 15, 16] for more details), we are able to derive the dependency models (DMs) of $\mathbf{X}^s \sim F^s$ in order to sample the random values of \mathbf{X}^s (see Corollaries 2-3). To derive such DMs for the purpose of SA, denote with $(\pi_1, \dots, \pi_{d-1})$ an arbitrary permutation of $(\sim j) := \{1, \dots, d\} \setminus \{j\}, \forall j \in \{1, \dots, d\}$, and define the permuted random vectors

$$\left(F_j(X_j^s), F_{\pi_1}(X_{\pi_1}^s), \dots, F_{\pi_{d-1}}(X_{\pi_{d-1}}^s) \right) \sim \widetilde{W}; \quad \left(F_j^s(X_j^s), F_{\pi_1}^s(X_{\pi_1}^s), \dots, F_{\pi_{d-1}}^s(X_{\pi_{d-1}}^s) \right) \sim \widetilde{C}.$$

It is worth noting that W and \widetilde{W} (resp. C and \widetilde{C}) share the same class of distribution functions, but the parameters of such functions may be different. To provide the first-type of DMs, consider a d -dimensional random vector $(Z_j, Z_{\pi_1}, \dots, Z_{\pi_{d-1}}) \sim \widetilde{W}$, and denote with $\widetilde{W}_{\pi_1|j}$ (resp. $\widetilde{W}_{\pi_2|j, \pi_1}$) the CDF of Z_{π_2} conditional on Z_j (resp. Z_j, Z_{π_1}), and $\widetilde{W}_{\pi_1|j}^{\leftarrow}$ (resp. $\widetilde{W}_{\pi_2|j, \pi_1}^{\leftarrow}$) the associated generalized inverse.

Corollary 2. *Assume that \mathbf{X} is continuous and $(Z_j, Z_{\pi_1}, \dots, Z_{\pi_{d-1}}) \sim \widetilde{W}$. Then, there exists a dependency function $r_j^w : \mathbb{R}^d \rightarrow \mathbb{R}^{d-1}$ and a random vector $\mathbf{U}_{\sim j} \sim \mathcal{U}(0, 1)^{d-1}$ which is independent of $X_j^s \sim F_j^s$ such that*

$$\mathbf{X}_{\sim j}^s = r_j^w(X_j^s, \mathbf{U}_{\sim j}) := \left(F_{\pi_1}^{\leftarrow}(Z_{\pi_1}), \dots, F_{\pi_{d-1}}^{\leftarrow}(Z_{\pi_{d-1}}) \right), \quad (7)$$

where

$$\begin{bmatrix} Z_{\pi_1} := \widetilde{W}_{\pi_1|j}^{\leftarrow}(U_{\pi_1} | F_j(X_j^s)) \\ Z_{\pi_2} := \widetilde{W}_{\pi_2|j, \pi_1}^{\leftarrow}(U_{\pi_2} | F_j(X_j^s), Z_{\pi_1}) \\ \vdots \\ Z_{\pi_{d-1}} := \widetilde{W}_{\pi_{d-1}|j, \pi_1, \dots, \pi_{d-2}}^{\leftarrow}(U_{\pi_{d-1}} | F_j(X_j^s), Z_{\pi_1}, \dots, Z_{\pi_{d-2}}) \end{bmatrix}.$$

Proof. See Appendix D. □

To give the copula-based DMs of $\mathbf{X}^s \sim F^s$ in the following corollary, consider now $(Z_j, Z_{\pi_1}, \dots, Z_{\pi_{d-1}}) \sim \widetilde{C}$, and denote with $\widetilde{C}_{\pi_1|j}$ (resp. $\widetilde{C}_{\pi_2|j, \pi_1}$) the CDF of Z_{π_2} conditional on Z_j (resp. Z_j, Z_{π_1}), and $\widetilde{C}_{\pi_1|j}^{\leftarrow}$ (resp. $\widetilde{C}_{\pi_2|j, \pi_1}^{\leftarrow}$) the associated generalized inverse.

Corollary 3. *Consider independent variables $\mathbf{U}_{\sim j} \sim \mathcal{U}(0, 1)^{d-1}$ and $X_j^s \sim F_j^s$, and*

assume \mathbf{X} is continuous and $(Z_j, Z_{\pi_1}, \dots, Z_{\pi_{d-1}}) \sim \tilde{C}$. Then, a DM of \mathbf{X}^s is

$$\mathbf{X}_{\sim j}^s = r_j^c(X_j^s, \mathbf{U}_{\sim j}) := \left(F_{\pi_1}^{\leftarrow} \left(W_{\pi_1}^{\leftarrow} (Z_{\pi_1}) \right), \dots, F_{\pi_{d-1}}^{\leftarrow} \left(W_{\pi_{d-1}}^{\leftarrow} (Z_{\pi_{d-1}}) \right) \right), \quad (8)$$

where

$$\begin{bmatrix} Z_{\pi_1} := \tilde{C}_{\pi_1|j}^{\leftarrow} (U_{\pi_1} | F_j^s(X_j^s)) \\ Z_{\pi_2} := \tilde{C}_{\pi_2|j, \pi_1}^{\leftarrow} (U_{\pi_2} | F_j^s(X_j^s), Z_{\pi_1}) \\ \vdots \\ Z_{\pi_{d-1}} := \tilde{C}_{\pi_{d-1}|j, \pi_1, \dots, \pi_{d-2}}^{\leftarrow} (U_{\pi_{d-1}} | F_j^s(X_j^s), Z_{\pi_1}, \dots, Z_{\pi_{d-2}}) \end{bmatrix},$$

provided that F_j^s is continuous.

Proof. See Appendix E. □

The dependency models provided in Corollaries 2-3 are going to be used for i) generating the random values of \mathbf{X}^s and ii) deriving the sensitivity functionals (see Section 3). It is to be noted that such sampling designs rely directly on X_j^s and are different from the one used in [13], which makes use of the weighted distribution of X_j . Thus, we are able to work with \mathbf{X}^s using W or \tilde{W} and the prescribed CDFs F_1, \dots, F_d without keeping in mind the desirability measure.

Remark 1. *The above dependency models require computing the inverse of conditional distributions of the CDF W or the copula C . When W and C are smooth enough and are not time-demanding for the computations, the differential approach can be used for computing the conditional distributions (see [40, 39]). For time-demanding desirability measures or equivalently W or C , it is common to approximate the copula given by Equation (6) using fundamental copulas such as elliptical, Frank, Clayton or other class of copulas. For instance, Equation (6) can be used for easily computing the Spearman rho in general, and the Kendall tau for elliptical copulas ([39]).*

3. Testing the effects of dependent inputs on dynamic outputs

In general the components of the random vector \mathbf{X}^s are not independent (correlated or dependent), and Equations (7)-(8) provide the statistical representations of such dependency structures for every $j \in \{1, \dots, d\}$. Combining the time-dependent safe outputs (i.e., $f(\mathbf{X}^s, t)$) with any DM (i.e., $r_j(X_j^s, \mathbf{U}_{\sim j})$) given by Equations (7)-(8) yields

$$Y_s(X_j^s, \mathbf{U}_{\sim j}, t) := f(X_j^s, \mathbf{X}_{\sim j}^s, t) = f(X_j^s, r_j(X_j^s, \mathbf{U}_{\sim j}), t), \quad \forall t \in [0, T]. \quad (9)$$

Note that $Y_s(X_j^s, \mathbf{U}_{\sim j}, t)$ includes only independent variables as inputs thanks to the dependency functions.

3.1. Dynamic sensitivity functionals for dependent variables

Using the output $Y_s(X_j^s, \mathbf{U}_{\sim j}, t)$, the SFs associated with the input X_j^s with $j \in \{1, \dots, d\}$ are formally defined below. The first-order and total SFs of X_j^s are respectively given by ([41, 25, 28, 14])

$$f_j^{fo}(X_j^s, t) := \mathbb{E} [Y_s(X_j^s, \mathbf{U}_{\sim j}, t) | X_j^s] - \mathbb{E} [Y_s(X_j^s, \mathbf{U}_{\sim j}, t)], \quad \forall t \in [0, T];$$

$$f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t) := Y_s(X_j^s, \mathbf{U}_{\sim j}, t) - \mathbb{E} [Y_s(X_j^s, \mathbf{U}_{\sim j}, t) | \mathbf{U}_{\sim j}], \quad \forall t \in [0, T].$$

We also define the centered output by

$$Y_s^c(X_j^s, \mathbf{U}_{\sim j}, t) := Y_s(X_j^s, \mathbf{U}_{\sim j}, t) - \mathbb{E} [Y_s(X_j^s, \mathbf{U}_{\sim j}, t)].$$

We can see that SFs are zero-mean time-dependent stochastic-processes, and such SFs contain the primary information about the inputs effects on the whole safe outputs. Such SFs and the centered outputs may follow the Gaussian processes or other processes. The Karhunen-Loeve expansion (e.g., [42, 43]), which relies on the theory of RKHS ([44, 35, 45, 46]), allows for concentrating the second-order moments of such random processes in a few uncorrelated components.

3.2. Test hypotheses

3.2.1. Initial null hypothesis for the independence test

Since $Y_s(X_j^s, \mathbf{U}_{\sim j}, t) \in \mathbb{R}$ includes only independent inputs, it is known in [47] (Proposition 1) that $Y_s(X_j^s, \mathbf{U}_{\sim j}, t)$ does not depend on X_j^s at time $t \in [0, T]$ if and only if the total SF $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t) = 0$ almost surely (a.s.). Thus, our initial null hypothesis of the independence test between X_j^s and the dynamic outputs is

$$f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t) = 0, \quad \forall t \in [0, T] \quad a.s.;$$

because, it can happen that some inputs contribute to the model outputs at the beginning of the process and others at the end ([48, 49]). The above null hypothesis is equivalent to $\forall [f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t)] = 0, \quad \forall t \in [0, T]$, which is well-suited for the total SFs following the Gaussian processes. Since SFs may follow skewed and heavy-tailed distributions and/or may be non-stationary random processes ([34]), adequate measures of association between time-dependent outputs and inputs are going to be considered so as to account for the main properties of such random processes.

3.2.2. Kernel-based test hypotheses

To account for the necessary or desirable moments of the stochastic processes such as SFs, we are going to embed the total SFs into the appropriate RKHS or feature spaces ([44, 35, 45, 46]), which offer an interesting framework for statistical independence tests of two random vectors ([45, 46, 50]).

Formally, consider an arbitrary space \mathcal{X} and an Hilbert space \mathcal{H} endowed with the inner product $\langle \cdot, \cdot \rangle$. For a feature map $\phi : \mathcal{X} \rightarrow \mathcal{H}$, a symmetric and positive-definite kernel is defined as follows ([44]): $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ with $k(x, x') = \langle \phi(x), \phi(x') \rangle$. Given a random variable such as the total SFs at time $t_0 \in [0, T]$, that is, $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0)$, its mean element is defined by ([45, 46, 51, 50])

$$\mu_{F_0} (f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0)) := \mathbb{E}_{F_0} [k(\cdot, f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0))] ,$$

with F_0 the CDF of $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0)$. The symbol \mathbb{E}_{F_0} means that the expectation is taken w.r.t. $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0) \sim F_0$. We can see that this mean element is a linear statistic of the embedded random variable (i.e., $k(\cdot, f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0))$) lying in the RKHS induced by the kernel $k(\cdot)$, and it accounts for all the moments or desirable moments of $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0)$, depending on the kernels used. It is used for entirely characterizing the distributions of $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_0)$ in the case of characteristic kernels. A characteristic kernel is defined as follows ([45, 46, 51, 50]): given two CDFs G_1, G_2 and two random variables $Y_1 \sim G_1$ and $Y_2 \sim G_2$,

$$\mu_{G_1}(Y_1) = \mu_{G_2}(Y_2) \implies G_1 = G_2 ,$$

with $\mu_{G_1}(Y_1) = \mathbb{E}_{G_1} [k(\cdot, Y_1)]$.

To introduce the kernel-based test hypotheses, the concept of equivalent kernels is needed (see Definition 1 and [14]). In what follows, we assume that

$$(A1) : \left(\int_0^T \mathbb{E} \sqrt{k(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t), f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t))} dt < +\infty, \quad \forall j \in \{1, \dots, d\} \right).$$

Definition 1. Let $(X_j^{s'}, \mathbf{U}'_{\sim j})$ be an i.i.d. copy of $(X_j^s, \mathbf{U}_{\sim j})$ and assume (A1) holds. A kernel $k(\cdot)$ is said to be an equivalent kernel for the independence criterion between X_j^s and $(f(\mathbf{X}^s, t), \forall t \in [0, T])$ whenever $\forall t \in [0, T]$

$$\mathbb{E} \left[k \left(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t) \right) \right] - k(0, 0) = 0 \implies f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t) = 0 \text{ a.s. .}$$

For instance, we can check that the quadratic kernel given by $k_q(y, y') := y^2(y')^2$

and the absolute kernel given by $k_a(y, y') := |y||y'|$ are equivalent kernels for the independence criterion between X_j^s and $f(\mathbf{X}^s, t)$. For performing the independence tests in what follows, we use \mathcal{K}_E for the set of equivalent kernels. Thus, the kernel-based test hypotheses are defined as follows: $\forall k(\cdot) \in \mathcal{K}_E$ and $\forall q > 0$,

$$\begin{cases} H_0 : \int_0^T |\mathbb{E}[k(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t))] - k(0, 0)|^q dt = 0 \\ H_1 : \int_0^T |\mathbb{E}[k(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t))] - k(0, 0)|^q dt \neq 0 \end{cases}, \quad (10)$$

where here $(X_j^{s'}, \mathbf{U}'_{\sim j})$ and $(X_j^s, \mathbf{U}_{\sim j})$ are two independent random vectors.

3.3. Test statistics

To build the test statistic, we are given two independent samples of the total SFs, that is,

$$\left\{ f_j^{tot}(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t) \right\}_{i=1}^m; \quad \left\{ f_j^{tot}(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t) \right\}_{i=1}^m,$$

and two independent samples of the centered outputs, that is,

$$\left\{ Y_s^c(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t) \right\}_{i=1}^M; \quad \left\{ Y_s^c(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t) \right\}_{i=1}^M,$$

for integers m, M with $m \leq M$. Based on such samples, the usual, unbiased and consistent estimator of $\mu_k^{tot}(t) := \mathbb{E}[k(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t))]$ and of $\sigma_k^{tot}(t) := \mathbb{V}[k(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t))]$ are respectively given by

$$\widehat{\mu}_k^{tot}(t) := \frac{1}{m} \sum_{i=1}^m k\left(f_j^{tot}(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t), f_j^{tot}(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t)\right),$$

$$\widehat{\sigma}_{k,H_0}^{tot}(t) := \frac{1}{m} \sum_{i=1}^m \left(k\left(f_j^{tot}(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t), f_j^{tot}(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t)\right) - k(0, 0) \right)^2,$$

for all $t \in [0, T]$. It is to be noted that $\widehat{\sigma}_{k,H_0}^{tot}(t)$ is an unbiased estimator only under H_0 . Moreover, given the centered outputs, consider the unbiased and consistent estimator of $\mu_k^c(t) := \mathbb{E}[k(Y_s^c(X_j^s, \mathbf{U}_{\sim j}, t), Y_s^c(X_j^{s'}, \mathbf{U}'_{\sim j}, t))]$ given by

$$\widehat{\mu}_k^c(t) := \frac{1}{M} \sum_{i=1}^M k\left(Y_s^c(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t), Y_s^c(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t)\right).$$

Under H_0 , the central limit theorem ensures that $\forall t \in [0, T]$,

$$\sqrt{m} \left(\widehat{\mu}_k^{tot}(t) - k(0, 0) \right) \xrightarrow{D} \mathcal{N}(0, \sigma_k^{tot}(t)). \quad (11)$$

Now, let us consider the zero-mean Gaussian random process given by $G(t) \sim \mathcal{N}(0, \sigma_k^{tot}(t))$ with $t \in [0, T]$, and assume that the time-dependent outputs have been observed at time $t_\ell \in [0, T]$ with $\ell = 1, \dots, L$. The test statistic under H_0 and its asymptotic distribution are given below. To that end, denote with $\{Z_r \sim \mathcal{N}(0, 1)\}_{r=1}^\infty$ i.i.d. variables; $\{\lambda_r\}_{r=1}^\infty$ the eigenvalues related to the Karhunen-Loeve expansion of $(G(t), t \in [0, T])$, and assume that

(A2) : the kernel $k(\cdot) \in \mathcal{K}_E$.

Corollary 4. *If $q = 2$ and (A1)-(A2) hold, then the test statistic is given by*

$$T_{k, H_0}^{tot} := m \sum_{\ell=1}^L \left| \widehat{\mu_k^{tot}}(t_\ell) - k(0, 0) \right|^2 \xrightarrow{D} \sum_{r=1}^{+\infty} \lambda_r Z_r^2 \quad \text{if } m, L \rightarrow +\infty. \quad (12)$$

Proof. See Appendix F. □

Obviously, the test statistic T_{k, H_0}^{tot} vanishes in the case of independence between inputs and outputs. The critical values (i.e., T_α) can be computed as the empirical quantile of $\sum_{r=1}^R \lambda_r Z_r^2$ at the value $\alpha \in [0, 1]$, where $\lambda_1 > \lambda_2 > \dots > \lambda_R$ and $\lambda_r \approx 0$ for all $r > R$. In the case H_0 is rejected (i.e., $T_{k, H_0}^{tot} > T_\alpha$), the power of this test (i.e., the probability of the rejection region under H_1 ([52])) gives the probability to detect the effect of X_j^s when such an effect effectively exists. This power depends on kernels.

Remark 2. *Under H_0 and for any real $q > 0$, Slutsky's theorem implies*

$$m^{q/2} \left| \widehat{\mu_k^{tot}}(t) - k(0, 0) \right|^q \xrightarrow{D} |G(t)|^q.$$

This equation is relevant for generalizing the test statistic given by (12).

Remark 3. *The expressions $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t)$ and $Y_s^c(X_j^s, \mathbf{U}_{\sim j}, t)$ need to be computed in general. Thus, the following consistent estimators are going to be used:*

$$\widehat{\mu}(\mathbf{U}_{\sim j}, t) := \frac{1}{m} \sum_{i=1}^m Y_s(X_{i,j}^s, \mathbf{U}_{\sim j}, t) \xrightarrow{P} \mathbb{E}_{X_j^s} [Y_s(X_j^s, \mathbf{U}_{\sim j}, t)],$$

$$\widehat{\mu}(X_j^s, t) := \frac{1}{m} \sum_{i=1}^m Y_s(X_j^s, \mathbf{U}_{i, \sim j}, t) \xrightarrow{P} \mathbb{E}_{\mathbf{U}_{\sim j}} [Y_s(X_j^s, \mathbf{U}_{\sim j}, t)],$$

$$\widehat{\mu}(t) := \frac{1}{m} \sum_{i=1}^m Y_s(X_{i,j}^s, \mathbf{U}_{i, \sim j}, t) \xrightarrow{P} \mathbb{E} [Y_s(X_j^s, \mathbf{U}_{\sim j}, t)].$$

Thus, the plug-in estimators of $\mu_k^{tot}(t)$ and $\mu_k^c(t)$ are respectively given by

$$\widehat{\mu_k^{tot}}(t) := \frac{1}{m} \sum_{i=1}^m k \left(Y_s(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t) - \widehat{\mu}(\mathbf{U}_{i,\sim j}, t), Y_s(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t) - \widehat{\mu}(\mathbf{U}'_{i,\sim j}, t) \right),$$

$$\widehat{\mu_k^c}(t) := \frac{1}{M} \sum_{i=1}^M k \left(Y_s(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t) - \widehat{\mu}(t), Y_s(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t) - \widehat{\mu}(t) \right).$$

4. Dependence measures: kernel-based sensitivity indices

This section aims at introducing new dependence measures between time-dependent outputs and non-independent inputs that satisfy Renyi' axioms, that is, each measure takes its values within $[0, 1]$. The null value means that the involved quantities are independent, while the value one suggests a strict dependence between both quantities (see [37] for more details).

4.1. Deviation measures

It comes out from Corollary 4 that when X_j^s contributes to dynamic outputs, the theoretical expression of the generalized test statistic up to a normalized constant, that is,

$$\mathcal{D}_k^{1,q}(f_j^{tot}) := \int_0^T \left| \mathbb{E} \left[k \left(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t) \right) \right] - k(0,0) \right|^q dt, \quad \forall q > 0,$$

can serve as a reasonable measure of the deviation from independence. To account for the auto-correlations in particular and the cross-components in general, the above deviation measure is extended as follows: $\forall t_1, t_2 \in [0, T]$,

$$\mathcal{D}_k^{2,q}(f_j^{tot}) := \int_0^T \int_0^T \left| \mathbb{E} \left[k \left(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_1), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t_2) \right) \right] - k(0,0) \right|^q dt_1 dt_2.$$

Both measures can be included in a wide class of the deviation measures from independence between X_j^s and $Y_s(X_j^s, \mathbf{U}_{\sim j})$ defined below.

Definition 2. Let $(X_j^s, \mathbf{U}_{\sim j})$ and $(X_j^{s(1)}, \mathbf{U}_{\sim j}^{(1)})$ be two random vectors that are identically distributed. Assume that (A1)-(A2) hold. The generalized deviation-measure is defined by

$$\mathcal{D}_k^q(f_j^{tot}) := \int_0^T \int_0^T \left| \mathbb{E} \left[k \left(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_1), f_j^{tot}(X_j^{s(1)}, \mathbf{U}_{\sim j}^{(1)}, t_2) \right) \right] - k(0,0) \right|^q dt_1 dt_2.$$

Remark 4. The deviation measure $\mathcal{D}_k^q(f_j^{tot})$ comes down to $\mathcal{D}_k^{2,q}(f_j^{tot})$ when $(X_j^{s(1)}, \mathbf{U}_{\sim j}^{(1)})$ is an i.i.d. copy of $(X_j^s, \mathbf{U}_{\sim j})$, and $\mathcal{D}_k^{2,q}(f_j^{tot})$ comes down to $\mathcal{D}_k^{1,q}(f_j^{tot})$ when $t_2 = t_1$.

While these measures of deviation rely on the total SF of X_j^s , such measures can be applied to the first-order SF of X_j^s as well (i.e., $\mathcal{D}_k^{2,q}(f_j^{fo})$). As $f_j^{fo}(X_j^s, t) = \mathbb{E}_{\mathbf{U}_{\sim j}} [f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t)]$ with $t \in [0, T]$, $f_j^{fo}(X_j^s, t)$ contains partial information compared to $f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t)$. Therefore, a reasonable kernel or the associated measure of deviation should be able to incorporate such a kind of information, and this leads to the following definition.

Definition 3. Assume that (A1)-(A2) hold. A kernel $k(\cdot)$ is said to be sensitivity-compatible whenever $\forall t_1, t_2 \in [0, T]$

$$\begin{aligned} & \left| \mathbb{E} \left[k \left(f_j^{fo}(X_j^s, t_1), f_j^{fo}(X_j^{s'}, t_2) \right) \right] - k(0, 0) \right| \leq \\ & \left| \mathbb{E} \left[k \left(f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t_1), f_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t_2) \right) \right] - k(0, 0) \right|. \end{aligned}$$

When $(X_j^s, \mathbf{U}_{\sim j})$ and $(X_j^{s'}, \mathbf{U}'_{\sim j})$ are independent, it is to be noted that the quadratic kernel $k_q(y, y') := y^2(y')^2$ and the absolute kernel $k_a(y, y') := |y||y'|$ are sensitivity-compatible kernels according to Lemma 1. Also, such kernels are not characteristic kernels in general (see [45, 46, 50]). However, we can check that the quadratic kernel is a characteristic kernel for the class of Gaussian SFs.

Lemma 1. Assume that (A1)-(A2) hold. If $k(\cdot, x)$ is convex and $k(\cdot, 0) = 0$, then $\mathcal{D}_k^{1,q}(\bullet)$ and $\mathcal{D}_k^{2,q}(\bullet)$ are sensitivity-compatible measures of deviation.

Proof. The results hold by applying Jensen's inequality, bearing in mind that $f_j^{fo}(X_j^s, t) = \mathbb{E}_{\mathbf{U}_{\sim j}} [f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t)]$, $\forall t \in [0, T]$. \square

Note that it is sufficient to require the kernels to be convex only on the support of the centered outputs.

4.2. Kernel-based sensitivity indices for dynamic models

Using the different measures of deviation and Lemma 1, this section aims at introducing two types of dependence measures. Formally, the first-type dependence measures are defined as follows: the total kernel-based SI (Kb-SI) of X_j^s is given by

$$S_{T_j}^{1,k,q} := \frac{\mathcal{D}_k^{1,q}(f_j^{tot})}{\mathcal{D}_k^{1,q}(Y_s^c)}, \quad (13)$$

and the first-order Kb-SI of X_j^s is defined as

$$S_j^{1,k,q} := \frac{\mathcal{D}_k^{1,q}(f_j^{fo})}{\mathcal{D}_k^{1,q}(Y_s^c)}. \quad (14)$$

In the same sense, the second-type Kb-SIs of X_j^s are defined as follows:

$$S_{T_j}^{2,k,q} := \frac{\mathcal{D}_k^{2,q}(f_j^{tot})}{\mathcal{D}_k^{2,q}(Y_s^c)}; \quad S_j^{2,k,q} := \frac{\mathcal{D}_k^{2,q}(f_j^{fo})}{\mathcal{D}_k^{2,q}(Y_s^c)}. \quad (15)$$

It is to be noted that the second-type Kb-SIs account for the cross-components of the time-dependent SFs, while the first-type Kb-SIs treat independently the SFs at time t_1 and t_2 with $t_1 \neq t_2$. The main properties of such measures are provided in Theorem 1. To provide such results, we assume that

(A3) : the kernel $k(\cdot)$ is a sensitivity-compatible kernel.

Theorem 1. *Consider a symmetric and positive definite kernel $k(\cdot)$, and assume that (A1)-(A3) hold. Then, we have*

$$\begin{aligned} S_{T_j}^{1,k,q} = S_{T_j}^{2,k,q} = 0 &\iff f(\mathbf{X}) \text{ does not depend on } X_j; \\ 0 \leq S_j^{1,k,q} \leq S_{T_j}^{1,k,q} \leq 1; &\quad 0 \leq S_j^{2,k,q} \leq S_{T_j}^{2,k,q} \leq 1. \end{aligned} \quad (16)$$

Proof. The proofs are straightforward using Definition 3 and the conditional Jensen's inequality. \square

Obviously, the choice of the type of the dependence measures and/or the kernels should rely on the elements or information one wants to account for. In general, characteristic kernels such as the Gaussian kernels and the Laplacian kernels allow for capturing all the moments of SFs provided that such kernels are sensitivity-compatible on the support of SFs.

4.3. Extension to spatio-temporel models

For a spatio-temporel model given by $f(\cdot, t) : \mathbb{R}^d \rightarrow \mathbb{R}^n$ with $t \in [0, T]$, we can see that the SFs at a given time t are n -dimensional random vectors. Thus, the kernel-based sensitivity indices of Section 4.2 can be extended to cope with spatio-temporel models by using the multivariate kernels such as $k_2(\mathbf{y}, \mathbf{y}') := \langle \mathbf{y}, \mathbf{y}' \rangle^2$, $k_p(\mathbf{y}, \mathbf{y}') := \|\mathbf{y}\|_p \|\mathbf{y}'\|_p$ with $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^n$.

5. Computational issues: empirical kernel-based sensitivity indices

For computing the Kb-SIs in general, we are given four independent samples from $(X_j^s, \mathbf{U}_{\sim j})$, that is,

$$\left\{ (X_{i,j}^s, \mathbf{U}_{i,\sim j}) \right\}_{i=1}^m; \quad \left\{ (X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}) \right\}_{i=1}^m; \quad \left\{ (X_{i,j}^s, \mathbf{U}_{i,\sim j}) \right\}_{i=1}^M; \quad \left\{ (X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}) \right\}_{i=1}^M,$$

with integers $m \leq M$. Using the first two samples and the consistent estimators provided in Remark 3 in one hand, and keeping in mind the method of moments and the Slutsky theorem in the other hand, the consistent estimators of SFs are given as follows: $\forall t \in [0, T]$,

$$\widehat{f}_j^{fo}(X_j^s, t) := \widehat{\mu}(X_j^s, t) - \widehat{\mu}(t) \xrightarrow{P} f_j^{fo}(X_j^s, t),$$

$$\widehat{f}_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t) := Y_s(X_j^s, \mathbf{U}_{\sim j}, t) - \widehat{\mu}(\mathbf{U}_{\sim j}, t) \xrightarrow{P} f_j^{tot}(X_j^s, \mathbf{U}_{\sim j}, t),$$

$$\widehat{Y}_s^c(X_j^s, \mathbf{U}_{\sim j}, t) := Y_s(X_j^s, \mathbf{U}_{\sim j}, t) - \widehat{\mu}(t) \xrightarrow{P} Y_s^c(X_j^s, \mathbf{U}_{\sim j}, t).$$

In the same sense, we use the second sample for computing $\widehat{f}_j^{fo}(X_j^{s'}, t)$, $\widehat{f}_j^{tot}(X_j^{s'}, \mathbf{U}'_{\sim j}, t)$ and $\widehat{Y}_s^c(X_j^{s'}, \mathbf{U}'_{\sim j}, t)$. Based on these estimators, the plug-in approach leads to the derivation of the consistent estimators of Kb-SIs (see Corollaries 5-6). To that end, assume that we have observed the model output $Y_s(X_j^s, \mathbf{U}_{\sim j}, t)$ at time $t_\ell \in [0, T]$ with $\ell = 1, \dots, L$, and consider now the two last samples.

Corollary 5. *For the first-type Kb-SIs, assume that (A1)-(A3) hold and $m, M, L \rightarrow +\infty$. Then,*

$$\widehat{S}_j^{1,k,q} := \frac{\sum_{\ell=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{f}_j^{fo}(X_{i,j}^s, t_\ell), \widehat{f}_j^{fo}(X_{i,j}^{s'}, t_\ell) \right) - k(0, 0) \right|^q}{\sum_{\ell=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{Y}_s^c(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t_\ell), \widehat{Y}_s^c(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t_\ell) \right) - k(0, 0) \right|^q} \xrightarrow{P} S_j^{1,k,q};$$

$$\widehat{S}_{T_j}^{1,k,q} := \frac{\sum_{\ell=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{f}_j^{tot}(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t_\ell), \widehat{f}_j^{tot}(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t_\ell) \right) - k(0, 0) \right|^q}{\sum_{\ell=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{Y}_s^c(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t_\ell), \widehat{Y}_s^c(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t_\ell) \right) - k(0, 0) \right|^q} \xrightarrow{P} S_{T_j}^{1,k,q}.$$

Likewise, Corollary 6 provides the estimators of the second-type Kb-SIs.

Corollary 6. *Assume that (A1)-(A3) hold and $m, M, L \rightarrow +\infty$. Then, we have*

$$\widehat{S}_j^{2,k,q} := \frac{\sum_{\ell_1=1}^L \sum_{\ell_2=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{f}_j^{fo}(X_{i,j}^s, t_{\ell_1}), \widehat{f}_j^{fo}(X_{i,j}^{s'}, t_{\ell_2}) \right) - k(0, 0) \right|^q}{\sum_{\ell_1=1}^L \sum_{\ell_2=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{Y}_s^c(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t_{\ell_1}), \widehat{Y}_s^c(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t_{\ell_2}) \right) - k(0, 0) \right|^q} \xrightarrow{P} S_j^{2,k,q};$$

$$\widehat{S}_{T_j}^{2,k,q} := \frac{\sum_{\ell_1=1}^L \sum_{\ell_2=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{f}_j^{tot}(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t_{\ell_1}), \widehat{f}_j^{tot}(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t_{\ell_2}) \right) - k(0, 0) \right|^q}{\sum_{\ell_1=1}^L \sum_{\ell_2=1}^L \left| \frac{1}{M} \sum_{i=1}^M k \left(\widehat{Y}_s^c(X_{i,j}^s, \mathbf{U}_{i,\sim j}, t_{\ell_1}), \widehat{Y}_s^c(X_{i,j}^{s'}, \mathbf{U}'_{i,\sim j}, t_{\ell_2}) \right) - k(0, 0) \right|^q} \xrightarrow{P} S_{T_j}^{2,k,q}.$$

The proofs of Corollaries 5-6 are straightforward using $\widehat{\mu}_k^{tot}(t)$, $\widehat{\mu}_k^c(t)$ and bearing in mind the method of moments and Slutsky's theorem. It is worth noting that such

generic estimators can be improved for some specific kernels, such as the quadratic kernel by making use of the theory of U-statistics (see [41, 25] for instance).

6. Real case study

6.1. Model description

This section deals with the application of Kb-SIs to the mean component of the dynamic and stochastic risk-model used in [5] (see Appendix G). Such an academic model simulates the concentration of *L.Monocytogenes* in a cheese at an hourly time-step under some crude simplifications. During the processing times of a cheese, the Codex Alimentarius encourages controlling the chemical and process parameters in such a way that the final products meet a food safety objective (FSO). For *L.Monocytogenes*, the FSO is about 100 CFU ([5, 4]). Thus, the cheese processing is going to be said safe when the concentration of *L.Monocytogenes* in a 25-g portion of that cheese is less than 100 CFU or equivalently the log-concentration at time t (i.e., $Y(t)$) satisfies: $Y(t) \leq 2$. Only the outputs at days 1, 2, ..., 28 are considered in this paper.

This hazard model includes four independent input variables \mathbf{X} , that is, the temperatures during the three main periods of cheeses processing and the initial concentration of *L.Monocytogenes*. The initial distributions of such inputs used in this paper are listed in Table 1. To work with the safe outputs, the following desirability measure is applied ([5]). Given a sample point \mathbf{x} of the model inputs \mathbf{X} , its ability to meet a FSO is measured by

$$D(\mathbf{x}) := \int_0^T |Y(\mathbf{x}, t) - 2| dt, \quad \text{if } Y(\mathbf{x}, t) \leq 2 \text{ for every } t \in [0, T] \text{ and } 0 \text{ otherwise.}$$

The above desirability measure gives different chances to the sample points of being included in SA. Therefore, the new distribution of the model inputs (i.e, the distribution of \mathbf{X}^s) is no more that of \mathbf{X} .

Inputs	Description	Unit	Probability distribution
T_1	Temperature of first period	$^{\circ}C$	Truncated normal $\mathcal{N}(8, 5)$ on $[-1.72, 45.5]$
T_2	Temperature of second period	$^{\circ}C$	Truncated normal $\mathcal{N}(8, 5)$ on $[-1.72, 45.5]$
T_3	Temperature of third period	$^{\circ}C$	Truncated normal $\mathcal{N}((8, 5)$ on $[-1.72, 45.5]$
C_0	Initial concentration	log CFU	Uniform $\mathcal{U}(-4, 1)$

Table 1: Initial distributions of inputs. In $\mathcal{N}(5, 10)$, we have mean=5, standard deviation=10.

6.2. Numerical results

For computing the Kb-SIs, the copula given by (6) is approximated by a fundamental copula since its evaluation is time-demanding (see Remark 1). But, the exact margins are used for running the dependency models. The Spearman rhos (i.e., $\rho_{12} := 12 \mathbb{E}[C_{12}(U_1, U_2)] - 3$) are computed using the exact copula with 500 sample points for each pair of the copula inputs. The estimates are

$$\rho_{12} = 0.086, \rho_{13} = 0.085, \rho_{14} = 0.078, \rho_{23} = 0.085, \rho_{24} = 0.078, \rho_{34} = -0.079 .$$

Using the estimated Spearman rhos, the Gaussian and Student-t copulas are considered for approximating the exact copula. Figure 1 compares the exact-copula values with the approximated values using the Gaussian copula and Student-t copula with 7 degrees of freedom. Taking the absolute value of the difference between the exact copula and each approximated copula, and computing the maximum and the sum yield Table 2.

Copula	Maximun	Sum
Gaussian	0.053	4.947
Student-t	0.043	4.848

Table 2: Statistics of the absolute difference between the exact copula and approximated copulas.

Thus, the student-t copula with 7 degrees of freedom is used in what follows. The conditional distributions of the approximated copula are computed using the R-package copula, and all random values are generated using Sobol' sequence (R-package randtoolbox). For computing the Kb-SIs, we used $m = M = 1000$ as the sample size and considered the absolute kernel $k_a(y, y') := |y||y'|$, the quadratic kernel $k_q(y, y') := y^2(y')^2$, the Laplacian kernel $k_l(y, y') := \exp(-0.05|y - y'|)$ and the Gaussian kernel $k_g(y, y') := \exp(-0.05(y - y')^2)$. The estimated Kb-SIs associated with such kernels are provided in Table 3.

It comes out from Table 3 that C_0 is the most influential input followed by T_3 . Thus, the initial concentration and T_3 are the important drivers for monitoring and controlling *L.Monocytogenes* in the cheese processing.

7. Conclusion

New kernel-based dependence measures between time-dependent model outputs and inputs under the safety objectives have been developed. Such dependence measures are based on the normalized test statistics for the independence test between

Indices	First-order Kb-SIs				Total Kb-SIs			
	T_1	T_2	T_3	C_0	T_1	T_2	T_3	C_0
$S_{\bullet}^{1,k_a,1/2}$	0.000	0.000	0.352	0.877	0.000	0.000	0.353	0.877
$S_{\bullet}^{1,k_q,1/2}$	0.000	0.000	0.383	0.665	0.000	0.000	0.392	0.666
$S_{\bullet}^{1,k_g,1/2}$	0.000	0.000	0.358	0.909	0.000	0.000	0.358	0.909
$S_{\bullet}^{1,k_l,1/2}$	0.000	0.000	0.378	0.843	0.000	0.000	0.381	0.840
$S_{\bullet}^{2,k_a,1/2}$	0.000	0.000	0.352	0.877	0.000	0.000	0.353	0.877
$S_{\bullet}^{2,k_q,1/2}$	0.000	0.000	0.383	0.665	0.000	0.000	0.392	0.666
$S_{\bullet}^{2,k_g,1/2}$	0.000	0.000	0.475	0.894	0.000	0.000	0.474	0.895
$S_{\bullet}^{2,k_l,1/2}$	0.000	0.000	0.468	0.831	0.000	0.000	0.472	0.828

Table 3: Kernel-based sensitivity indices for the dynamic risk-model.

dynamic outputs and inputs, which comply with the safety objectives. Such measures lead to a statistical foundation of Kb-SIs for time-dependent models or processes using the asymmetric role of inputs and outputs. Mainly, two types of Kb-SIs are provided so as to account for the auto-correlations among SFs. Thanks to kernel methods, our approaches are flexible enough so that the proposed indices extend de-

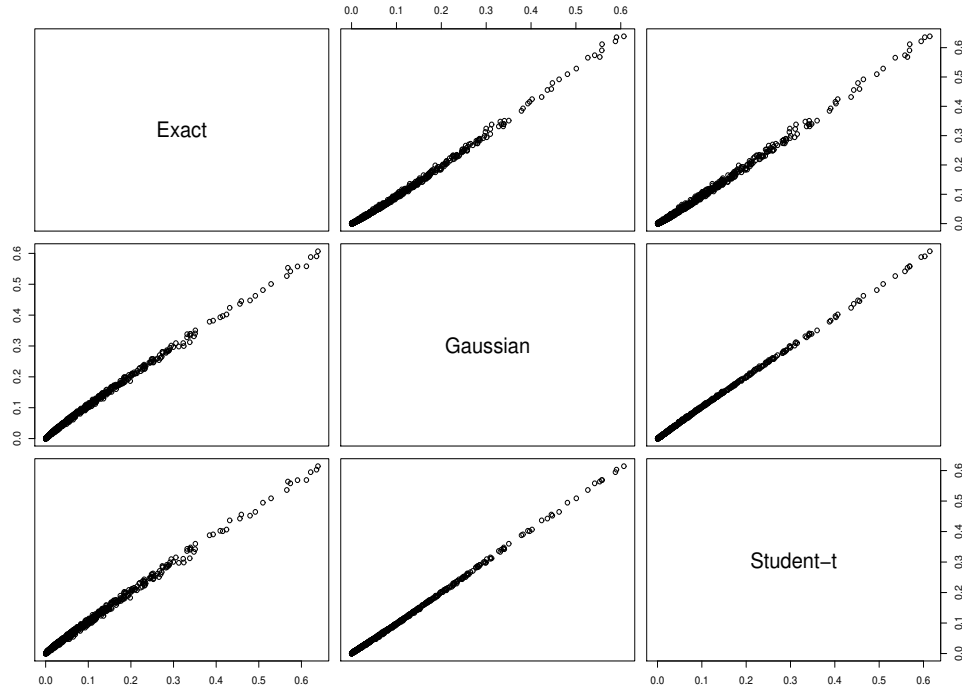


Figure 1: Comparison of the exact, Gaussian and Student-t copulas using 1000 sample points.

pendent generalized sensitivity indices ([15, 16]) to cope with i) desirable or sufficient moments of SFs, ii) every distribution of SFs, including skewed and heavy-tailed distributions, and iii) non-stationary SFs.

For computing such indices, we have derived a new CDF W (similar to copula) and the copula (i.e., C) of the inputs that takes account of the desirability measures. For the risk-model considered in this paper, the exact copula is approximated by the Student-t copula since its evaluation is time-demanding. But, the exact margins of the distribution of inputs (derived in this paper) are used. Numerical results have shown that different values of the sensitivity indices can be obtained using different kernels. Thus, among the set of equivalent kernels for the independence tests that are also sensitivity-compatible, the choice of a kernel should be based on the information we want to account for (e.g., statistical properties of SFs). Further investigations are needed for more precision about the choice of kernels, such as kernels that may lead to powerful tests.

The computations of Kb-SIs require more model runs compared to those of dependent generalized sensitivity indices since the SFs are unknown in general. Moreover, when the desirability measures involve the hazard risk-model, the number of model runs still increases with respect to the number of inputs. In future work, it is worth investigating low-cost computations of such indices.

Acknowledgments

We would like to thank the reviewers for their suggestions. Special thanks to one reviewer for its detailed comments that have helped improving our manuscript. We also thank the associate editor for its interest in our manuscript.

Appendix A Proof of Proposition 1

Consider $u_j \in \{F_j(x_j) : x_j \in \mathbb{R}\}$ with $j = 1, \dots, d$. Using Equation (2) and the fact that F_j^{\leftarrow} is an increasing function, we can write the probability

$$\begin{aligned} \mathbb{P}(F_1(X_1^s) \leq u_1, \dots, F_d(X_d^s) \leq u_d) &= \mathbb{P}(X_1^s \leq F_1^{\leftarrow}(u_1), \dots, X_d^s \leq F_d^{\leftarrow}(u_d)) \\ &\stackrel{(2)}{=} W(F_1(F_1^{\leftarrow}(u_1)), \dots, F_d(F_d^{\leftarrow}(u_d))) \\ &= W(u_1, \dots, u_d) . \end{aligned}$$

Appendix B Proof of Proposition 3

For the first result, we have by definition and Equation (2)

$$\begin{aligned} F_j^s(x_j) &= \mathbb{P}(X_1^s \leq +\infty, \dots, X_j^s \leq x_j, \dots, X_d^s \leq +\infty) \\ &\stackrel{(2)}{=} W(1, \dots, 1, F_j(x_j), 1, \dots, 1) = W_j(F_j(x_j)), \end{aligned}$$

because $\lim_{x \rightarrow +\infty} F_j(x) = 1$ for any $j \in \{1, \dots, d\}$. The second result becomes obvious.

Appendix C Proof of Corollary 1

It is known in [13] (Corollary 2) that the copula of \mathbf{X}^s is

$$C(u_1, \dots, u_d) = W(F_1(F_1^{s\leftarrow}(u_1)), \dots, F_d(F_d^{s\leftarrow}(u_d))).$$

Combining the above copula with the fact that $F_j^{s\leftarrow}(U_j) = F_j^{\leftarrow}(W_j^{\leftarrow}(U_j))$ (see Proposition 3) yields the result because $F_j(F_j^{s\leftarrow}(u_j)) = F_j(F_j^{\leftarrow}(W_j^{\leftarrow}(u_j))) = W_j^{\leftarrow}(u_j)$.

Appendix D Proof of Corollary 2

The general conditional sampling algorithm ([38, 39, 15, 16]) allows to write

$$\left[\begin{array}{l} Z_j := \widetilde{W}_j^{\leftarrow}(U_j) \\ Z_{\pi_1} := \widetilde{W}_{\pi_1|j}^{\leftarrow}(U_{\pi_1} | Z_j) \\ Z_{\pi_2} := \widetilde{W}_{\pi_2|j, \pi_1}^{\leftarrow}(U_{\pi_2} | Z_j, Z_{\pi_1}) \\ \vdots \\ Z_{\pi_{d-1}} := \widetilde{W}_{\pi_{d-1}|j, \pi_1, \dots, \pi_{d-2}}^{\leftarrow}(U_{\pi_{d-1}} | Z_j, Z_{\pi_1}, \dots, Z_{\pi_{d-2}}) \end{array} \right] \sim \widetilde{W}.$$

According to Proposition 3, $Z_j = F_j(X_j^s)$ for continuous distribution F_j , and the result holds.

Appendix E Proof of Corollary 3

It is an adaptation of the general conditional sampling algorithm ([38, 39, 15, 16] for more details). Indeed, we have

$$\mathbf{X}_{\sim j}^s \stackrel{d}{=} r_j^c(X_j^s, \mathbf{U}_{\sim j}) := \left(F_{\pi_1}^{s\leftarrow}(Z_{\pi_1}), \dots, F_{\pi_{d-1}}^{s\leftarrow}(Z_{\pi_{d-1}}) \right),$$

and the result holds using Proposition 3.

Appendix F Proof of Corollary 4

According to Equation (13), $(G(t), t \in [0, T])$ is a zero-mean Gaussian process. Therefore, the result holds using the Karhunen-Loeve expansion, that is, the squared norm of a zero-mean Gaussian process can be represented as $\sum_{r=1}^{+\infty} \lambda_r Z_r^2$ (see [42], Chapter 1).

Appendix G Dynamic risk model (see Section 6)

The time-dependent hazard model is based on the following equations that involve the characteristics of *L.Monocytogenes* and the process parameters. The model output at time t (i.e., $Y(t)$) is the log-concentration of *L.Monocytogenes* (CFU) in a cheese. It is given by

$$Y(t) = 9 - \log_{10} \left[1 + \left(\frac{10^9}{10^{C_0}} - 1 \right) \exp(-r(t) * t) \right],$$

where the growth rate (i.e., $r(t)$) is the product of the optimal growth rate (i.e., $\mu_{opt} = 1/4$), the factors of temperature (i.e., $\gamma(T)$), pH (i.e., $\gamma(pH)$) and water activity (i.e., $\gamma(aw)$). This rate is given by

$$r(t) = \frac{1}{4} \times \gamma(T(t)) \times \gamma(pH(t)) \times \gamma(aw),$$

where $T(t)$ and $pH(t)$ are the temperature and pH at time t , respectively. We use $T_1 = T(t)$ from days 0 to 7; $T_2 = T(t)$ from days 8 to 14; $T_3 = T(t)$ from days 15 to 28; and

$$\gamma(T(t)) = \frac{(T(t) - 45.5)(T(t) + 1.72)^2}{(38.72 [38.72(T(t) - 37) + 8.5(35.28 - 2T(t))])},$$

$$\gamma(pH(t)) = \begin{cases} \frac{(pH(t)-9.61)(pH(t)-4.71)}{2.59(pH(t)-7.1)+2.51(4.71-pH(t))} & \text{if } 4.71 < pH(t) < 9.61 \\ 0 & \text{otherwise} \end{cases},$$

with

$$pH(t) = -1.14 \times 10^{-9}t^3 + 3.04 \times 10^{-6}t^2 - 5.2 \times 10^{-4}t + 4.58, \quad \gamma(aw) = 0.91666.$$

References

- [1] J. Ghosh, J. E. Padgett, M. Sánchez-Silva, Seismic damage accumulation in highway bridges in earthquake-prone regions, *Earthquake Spectra* 31 (1) (2015) 115–135.

- [2] F. Turchetti, E. Tubaldi, E. Patelli, P. Castaldo, C. Málaga-Chuquitaype, Damage modelling of a bridge pier subjected to multiple earthquakes : a comparative study, *Bulletin of Earthquake Engineering* (2023).
- [3] T. Ross, T. A. McMeekin, Modeling microbial growth within food safety risk assessments, *Risk Analysis* 23 (1) (2003) 179–197.
- [4] F. Tenenhaus-Aziza, J.-J. Daudin, A. Maffre, M. Sanaa, Risk-based approach for microbiological food safety management in the dairy industry: The case of *listeria monocytogenes* in soft cheese made from pasteurized milk, *Risk Analysis* 34 (1) (2014) 56–74.
- [5] M. Lamboni, M. Sanaa, F. Tenenhaus-Aziza, 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) (2014) 751–764.
- [6] S. Elsworth, T. Filatova, A. Jakeman, A. Kettner, M. Zellner, I. Athanasiadis, S. Hamilton, R. Axtell, D. Brown, J. Gilligan, M. Janssen, D. Robinson, J. Rozenberg, I. Ullah, S. Lade, Eight grand challenges in socio-environmental systems modeling, *Socio-Environmental Systems Modelling* 2 (2020).
- [7] K. Beven, J. Freer, Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology, *Journal of Hydrology* 249 (2001) 11–29.
- [8] B. Iooss, Revue sur l’analyse de sensibilité globale de modèles numériques, *Journal de la Société Française de Statistique* 152 (2011) 1–23.
- [9] M. Lamboni, B. Iooss, A.-L. Popelin, F. Gamboa, Derivative-based global sensitivity measures: General links with Sobol’ indices and numerical tests, *Mathematics and Computers in Simulation* 87 (0) (2013) 45 – 54.
- [10] C. R. Rao, On discrete distributions arising out of methods of ascertainment, *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)* 27 (2/4) (1965) 311–324.
- [11] G. P. Patil, C. R. Rao, Weighted distributions and size-biased sampling with applications to wildlife populations and human families, *Biometrics* 34 (2) (1978) 179–189.
- [12] J. Navarro, J. M. Ruiz, Y. D. Aguila, Multivariate weighted distributions: a review and some extensions, *Statistics* 40 (1) (2006) 51–64.

- [13] M. Lamboni, On exact distribution for multivariate weighted distributions and classification, *Methodol Comput Appl Probab* 25 (2023) 1–41. doi:<https://doi.org/10.1007/s11009-023-09993-2>.
- [14] M. Lamboni, Kernel-based sensitivity indices for any model behavior and screening, *Socio-Environmental Systems Modelling* 5 (18566) (2023).
- [15] M. Lamboni, S. Kucherenko, Multivariate sensitivity analysis and derivative-based global sensitivity measures with dependent variables, *Reliability Engineering & System Safety* 212 (2021) 107519.
- [16] M. Lamboni, On dependent generalized sensitivity indices and asymptotic distributions, arXiv preprint arXiv:2104.12938 (2022).
- [17] M. Lamboni, Efficient dependency models: simulating dependent random variables, *Mathematics and Computers in Simulation* , submitted on 03/01/2021 (2021).
- [18] I. M. Sobol, Sensitivity analysis for non-linear mathematical models, *Mathematical Modelling and Computational Experiments* 1 (1993) 407–414.
- [19] T. A. Mara, S. Tarantola, Variance-based sensitivity indices for models with dependent inputs, *Reliability Engineering & System Safety* 107 (2012) 115 – 121.
- [20] M. Kubicek, E. Minisci, M. Cisternino, High dimensional sensitivity analysis using surrogate modeling and high dimensional model representation, *International Journal for Uncertainty Quantification* 5 (5) (2015) 393–414.
- [21] S. D. Veiga, Global sensitivity analysis with dependence measures, *Journal of Statistical Computation and Simulation* 85 (7) (2015) 1283–1305.
- [22] S. Tarantola, T. A. Mara, Variance-based sensitivity indices of computer models with dependent inputs: The fourier amplitude sensitivity test, *International Journal for Uncertainty Quantification* 7 (6) (2017) 511–523.
- [23] A. Alexanderian, P. A. Gremaud, R. C. Smith, Variance-based sensitivity analysis for time-dependent processes, *Reliability Engineering & System Safety* 196 (2020) 106722.
- [24] M. B. Heredia, C. Prieur, N. Eckert, Nonparametric estimation of aggregated sobol’ indices: Application to a depth averaged snow avalanche model, *Reliability Engineering & System Safety* 212 (2021) 107422.

- [25] M. Lamboni, Multivariate sensitivity analysis: Minimum variance unbiased estimators of the first-order and total-effect covariance matrices, *Reliability Engineering & System Safety* 187 (2019) 67 – 92.
- [26] B. Iooss, C. Prieur, Shapley effects for sensitivity analysis with correlated inputs: Comparisons with sobol’ indices, numerical estimation and applications, *International Journal for Uncertainty Quantification* 9 (5) (2019) 493–514.
- [27] S. Da-Veiga, F. Gamboa, B. Iooss, C. Prieur, *Basics and Trends in Sensitivity Analysis: Theory and Practice in R*, SIAM, 2021.
- [28] M. Lamboni, Weak derivative-based expansion of functions: Anova and some inequalities, *Mathematics and Computers in Simulation* 194 (2022) 691–718.
- [29] T. Perrin, O. Roustant, J. Rohmer, O. Alata, J. Naulin, D. Idier, R. Pedreros, D. Moncoulon, P. Tinard, Functional principal component analysis for global sensitivity analysis of model with spatial output, *Reliability Engineering & System Safety* 211 (2021) 107522.
- [30] J. Barr, H. Rabitz, A generalized kernel method for global sensitivity analysis, *SIAM/ASA Journal on Uncertainty Quantification* 10 (1) (2022) 27–54.
- [31] M. Lamboni, H. Monod, D. Makowski, Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models, *Reliability Engineering and System Safety* 96 (2011) 450–459.
- [32] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, Sensitivity indices for multivariate outputs, *Comptes Rendus Mathématique* 351 (7) (2013) 307–310.
- [33] S. Roux, S. Buis, F. Lafolie, M. Lamboni, Cluster-based gsa: Global sensitivity analysis of models with temporal or spatial outputs using clustering, *Environmental Modelling & Software* 140 (2021) 105046.
- [34] H. Cramér, On the theory of stationary random processes, *Annals of Mathematics* (1940) 215–230.
- [35] A. Berlinet, C. Thomas, T. A. Gnan, *Reproducing Kernel Hilbert Space in probability and statistics*, Kluwer Academic, 2004.
- [36] A. Gretton, R. Herbrich, A. Smola, O. Bousquet, B. Schölkopf, Kernel methods for measuring independence, *Journal of Machine Learning Research* 6 (2005) 2075–2129.

- [37] A. Renyi, On measures of dependence, *Acta Mathematica Academiae Scientiarum Hungarica* 10 (3-4) (1959) 441–451.
- [38] L. Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag, New York., 1986.
- [39] A. J. McNeil, R. Frey, P. Embrechts, *Quantitative Risk Management*, Princeton University Press, Princeton and Oxford, 2015.
- [40] R. Nelsen, *An Introduction to Copulas*, Springer-Verlag, New York., 2006.
- [41] M. Lamboni, Uncertainty quantification: a minimum variance unbiased (joint) estimator of the non-normalized Sobol’ indices, *Statistical Papers* (2018) –doi: <https://doi.org/10.1007/s00362-018-1010-4>.
- [42] H. H. Kuo, *Gaussian Measures in Banach Spaces*. *Lecture Notes in Mathematics* 463, Springer, Berlin, 1975.
- [43] J. O. Ramsay, B. W. Silverman, *Applied Functional Data Analysis: Methods and Case Studies*, Springer-Verlag, Berlin, 2002.
- [44] N. Aronszajn, Theory of reproducing kernels, *Transactions of the American Mathematical Society* 68 (1950) 337–404.
- [45] A. Gretton, K. Borgwardt, M. Rasch, B. Schölkopf, A. Smola, A kernel method for the two-sample-problem, in: B. Schölkopf, J. Platt, T. Hoffman (Eds.), *Advances in Neural Information Processing Systems*, Vol. 19, MIT Press, 2007.
- [46] A. Smola, A. Gretton, L. Song, B. Schölkopf, A hilbert space embedding for distributions, in: *International Conference on Algorithmic Learning Theory*, Springer, 2007, pp. 13–31.
- [47] M. Lamboni, Derivative-based integral equalities and inequality: A proxy-measure for sensitivity analysis, *Mathematics and Computers in Simulation* 179 (2021) 137 – 161.
- [48] K. Campbell, M. D. McKay, B. J. Williams, Sensitivity analysis when model outputs are functions, *Reliability Engineering & System Safety* 91 (10-11) (2006) 1468 – 1472.
- [49] M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, H. Monod, Multivariate global sensitivity analysis for dynamic crop models, *Fields Crop Research* 113 (2009) 312–320.

- [50] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, A. Smola, A kernel two-sample test, *J. Mach. Learn. Res.* 13 (2012) 723–773.
- [51] K. Fukumizu, A. Gretton, B. Schölkopf, B. K. Sriperumbudur, Characteristic kernels on groups and semigroups, in: D. Koller, D. Schuurmans, Y. Bengio, L. Bottou (Eds.), *Advances in Neural Information Processing Systems*, Vol. 21, Curran Associates, Inc., 2009.
- [52] A. Wald, Asymptotically most powerful tests of statistical hypotheses, *The Annals of Mathematical Statistics* 12 (1) (1941) 1–19.