



**HAL**  
open science

# Global sensitivity analysis with aggregated Shapley effects, application to avalanche hazard assessment

María Belén Heredia, Clémentine Prieur, Nicolas Eckert

## ► To cite this version:

María Belén Heredia, Clémentine Prieur, Nicolas Eckert. Global sensitivity analysis with aggregated Shapley effects, application to avalanche hazard assessment. *Reliability Engineering and System Safety*, 2022, 222 (108420), pp.1-11. 10.1016/j.ress.2022.108420 . hal-02908480v2

**HAL Id: hal-02908480**

**<https://hal.science/hal-02908480v2>**

Submitted on 28 Feb 2022

**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.

# Global sensitivity analysis with aggregated Shapley effects, application to avalanche hazard assessment

María Belén Heredia<sup>a</sup>, Clémentine Prieur<sup>b</sup>, Nicolas Eckert<sup>a</sup>

<sup>a</sup>*Univ. Grenoble Alpes, INRAE, UR ETNA, Grenoble, France*

<sup>b</sup>*Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, 38000 Grenoble, France*

---

## Abstract

Dynamic models are simplified representations of some real-world entities that change over time. They are essential analytical tools with significant applications, e.g., in environmental and social sciences. Due to physical constraints applied on the outputs, it happens that input parameters are confined to a non-rectangular domain. In order to perform sensitivity analysis in this setting, we introduce the notion of aggregated Shapley effects and we propose an algorithm to estimate them with associated bootstrap confidence intervals. Our procedure is applied to analyze the sensitivity of an avalanche flow dynamic model from an input/output sample obtained by considering only input combinations leading to avalanche events that are both realistic and of interest for risk purposes. More precisely, we analyze the sensitivity in two different settings: (i) little knowledge on the input parameter probability distribution, and (ii) well-calibrated input parameter distribution. This leads insightful results regarding avalanche dynamics and potential related hazard, which demonstrate the usefulness of our approach for practical problems.

*Keywords:* Global sensitivity analysis, dependent inputs, aggregated Shapley effects, bootstrap confidence intervals, snow avalanche propagation model, snow avalanche hazard assessment

---

## 1. Introduction

Dynamic models are simplified representations of some real-world entities that change over time, in equations or computer codes. These models are useful for the analysis of real-world phenomena, e.g., in environmental or social sciences. For a better understanding of a phenomenon or for forecasting purposes, it might be important to identify which input parameters entering in the formulation of such dynamic models are influential on the outputs of interest. Determining these influential parameters is one aim of global sensitivity analysis (GSA) (see, e.g., [1] or more recently [2, 3] for a review of global sensitivity measures).

A particularity of dynamic models is that the outputs they produce are functional, typically time and/or space dependent (see, e.g., [4, 5, 6]). More

specifically, the application that motivated our study is a snow avalanche flow dynamic model proposed by [7] and heavily used for hazard and risk within a stochastic setting [8, 9, 10]. In order to describe the variability of all avalanche events likely to occur on a given path and assess the related hazard, we do not focus on a single avalanche event, so that input parameters of the propagation model cannot be considered as stochastically independent. Indeed, even if input parameter combinations are first sampled independently, we only analyze in the GSA the ones leading to snow avalanche simulations that are both realistic and of interest for risk purposes. For example, an avalanche simulation is rejected if its volume is inferior to  $7000 \text{ m}^3$  as such small events do not represent a significant threat for downstream settlements. As mentioned in [11], this setting encompasses a wide range of situations encountered in the natural sciences, engineering, design, economics and finances where model parameters are subject to certain limitations imposed e.g. by conservation laws, geometry, costs, quality constraints etc. Our previous work in [5] was focused on the analysis of a single avalanche event, based on aggregated Sobol' indices whose interpretation is not clear in the present framework of dependent inputs. For that reason, we develop a GSA methodology capable to handle dependent inputs, with unknown dependence structure, as far as dynamic models through functional outputs (or high-dimensional multivariate outputs as functional outputs are most of the time represented on a discrete grid).

In this paper, we focus on sensitivity analysis based on Shapley effects introduced in [12]. The main advantage of such an approach is that it is possible to attribute a non negative sensitivity index to each parameter, and the sum of the indices is equal to one [13, 14], even if the inputs are not independent.

A first algorithm to estimate Shapley effects was proposed in [15]. Then Benoumechiara and Elie-Dit-Cosaque [16] proposed a block bootstrap procedure for quantifying the uncertainty of an estimate produced by the aforementioned algorithm. More recently, [13] proposed a subset aggregation procedure which leads to a significant reduction of the variance of Shapley effect estimation. Let us also cite the algorithm proposed in [17] based on the Möbius inverse, which offers a computationally efficient alternative for the estimation of Shapley effects, the simple Monte Carlo sampling-based algorithm proposed in [18] focused on independent inputs but which can be extended with a loss of efficiency to the dependent framework. The main advantage of the subset aggregation procedure introduced in [13] is that it has a version, based on nearest-neighbors, which does not require the ability to sample from the exact conditional distributions of the input parameters. Note that the authors in [13] do not quantify the uncertainty associated to their estimation procedure with confidence intervals.

In the present paper, we introduce the notion of aggregated Shapley effects for multivariate outputs with potentially dependent inputs, generalizing the notion of aggregated Sobol' indices [19, 20]. We also study their invariance properties and we propose a given data estimation procedure, extending the nearest neighbor approach introduced in [13] and the block bootstrap procedure introduced in [16]. By given data procedure, we mean a procedure which can be applied to any given finite set of input/output data. Eventually, we

demonstrate the usefulness of this methodology in our complex avalanche application where the set of output simulations is constrained in order to keep only meaningful snow avalanche events. More precisely, we perform the GSA of our avalanche flow dynamic model in two different settings: (i) little knowledge on the input parameter probability distribution, and (ii) well-calibrated input parameter distribution as function of available field observations [8].

The paper is organized as follows. In Section 2, we introduce aggregated Shapley effects and state their main properties. In Section 3, we propose a given data estimation procedure for aggregated Shapley effects by extending the nearest-neighbor procedure introduced in [13] for the estimation of scalar Shapley effects. In Section 4, we extend the block bootstrap algorithm introduced in [16] to the given data framework. Finally in Section 5, our GSA procedure is applied to an avalanche dynamic model. We discuss our conclusions and perspectives in Section 6. The proofs of the propositions stated in Section 2, as far as auxiliary statistics and graphs are postponed to Appendix A. Additional numerical experiments are provided as supplementary material.

## 2. Aggregated Shapley effects

Shapley effects are sensitivity measures to quantify input importance proposed by [12]. Several recent works have demonstrated the great interest of Shapley effects for analyzing sensitivity in the framework of dependent inputs (see, e.g., [15, 21, 14]). In this section, we introduce the notion of aggregated Shapley effects, a natural extension of Shapley effects to multivariate outputs, generalizing the notion of aggregated Sobol' indices [19, 22].

### 2.1. Definition

Let  $\mathbf{Y} = f(\mathbf{X})$  with  $\mathbf{Y} = (Y_1, \dots, Y_j, \dots, Y_p) \in \mathbb{R}^p$  and  $\mathbf{X} = (X_1, \dots, X_i, \dots, X_d) \in \mathbb{R}^d$ . The input vector is defined on some probability space  $(\Omega, \mathcal{F}, \mathbb{P}_{\mathbf{X}})$  and  $f \in \mathbb{L}^2(\mathbb{P}_{\mathbf{X}})$ . For any  $\mathbf{u} \subseteq \{1, \dots, d\}$ , let us define  $-\mathbf{u} = \{1, \dots, d\} \setminus \mathbf{u}$ . We set  $\mathbf{X}_{\mathbf{u}} = (X_i)_{i \in \mathbf{u}}$ . In our application to avalanche hazard assessment,  $\mathbf{Y} = (Y_1 = f(s_1, \mathbf{X}), \dots, Y_p = f(s_p, \mathbf{X}))$ , with  $s_1, \dots, s_p \in \mathbb{R}$  the  $p$  discretization points along the avalanche corridor.

For any  $j = 1, \dots, p$  and any  $i = 1, \dots, d$ , the Shapley effect of  $Y_j$  with respect to  $X_i$  is defined as in [12] by:

$$\text{Sh}_i^j = \frac{1}{d} \sum_{\mathbf{u} \subseteq -\{i\}} \binom{d-1}{|\mathbf{u}|}^{-1} (\text{val}_j(\mathbf{u} \cup \{i\}) - \text{val}_j(\mathbf{u})) \quad (1)$$

with

$$\text{val}_j(\mathbf{u}) = \frac{\text{Var}(\mathbb{E}(Y_j | \mathbf{X}_{\mathbf{u}}))}{\text{Var}(Y_j)}, \quad \mathbf{u} \subseteq \{1, \dots, d\}. \quad (2)$$

Thus  $\text{Sh}_i^j$  is the Shapley value of input  $X_i$  for characteristic function  $\mathbf{u} \mapsto \text{val}_j(\mathbf{u})$  (see [23] for more details).

Generalizing the notion of aggregated Sobol' indices in [19, 20], we introduce the notion of aggregated Shapley effects. For any  $i = 1, \dots, d$ , the aggregated Shapley effect of  $\mathbf{Y} \in \mathbb{R}^p$  with respect to  $X_i$  is defined as:

$$\text{GSh}_i = \frac{\sum_{j=1}^p \text{Var}(Y_j) \text{Sh}_i^j}{\sum_{j=1}^p \text{Var}(Y_j)}. \quad (3)$$

## 2.2. Properties

In Propositions 2.1 and 2.2 below, we state the main properties of aggregated Shapley effects. The proofs are postponed to Appendix A.1.

**Proposition 2.1.** *The aggregated Shapley effects  $\text{GSh}_i$ ,  $i \in \{1, \dots, d\}$ , correspond to the Shapley values for characteristic function defined as:*

$$\mathbf{u} \subseteq \{1, \dots, d\} \mapsto \text{val}(\mathbf{u}) = \frac{\sum_{j=1}^p \text{Var}(Y_j) \text{val}_j(\mathbf{u})}{\sum_{j=1}^p \text{Var}(Y_j)} = \frac{\text{tr}(\Sigma_{\mathbf{u}})}{\text{tr}(\Sigma)},$$

with  $\Sigma_{\mathbf{u}}$  the covariance matrix of  $\mathbb{E}(\mathbf{Y}|\mathbf{X}_{\mathbf{u}})$  and  $\Sigma$  the covariance matrix of  $\mathbf{Y}$ .

In the next proposition, we prove that aggregated Shapley effects accomplish the natural requirements for a sensitivity measure mentioned in [22, Proposition 3.1].

**Proposition 2.2.** *Let  $i \in \{1, \dots, d\}$ . The following items hold true.*

- i.  $0 \leq \text{GSh}_i \leq 1$ .
- ii.  $\text{GSh}_i$  is invariant by left-composition by any nonzero scaling of  $f$ , which means, for any  $\lambda \in \mathbb{R}$  with  $\lambda \neq 0$ , the aggregated Shapley effect  $\text{GSh}'_i$  of  $\lambda f(\mathbf{X})$  is  $\text{GSh}_i$ .
- iii.  $\text{GSh}_i$  is invariant by left-composition of  $f$  by any isometry of  $\mathbb{R}^p$ , which means, for any  $O \in \mathbb{R}^{p \times p}$  such that  $O^t O = I$ , the aggregated Shapley effect  $\text{GSh}'_i$  of  $O f(\mathbf{X})$  is  $\text{GSh}_i$  for all  $i \in \{1, \dots, d\}$ .

In practice, there does not exist any analytical formula for aggregated Shapley effects. Therefore we propose in the next section an estimator built on a given input/output sample  $(\mathbf{X}^{(k)}, \mathbf{Y}^{(k)})$ ,  $k = 1, \dots, n$ .

## 3. Estimation procedure for scalar and aggregated Shapley effects

In this section, we first recall the definition of the given data estimator of scalar Shapley effects introduced in [13, Section 6] (see Section 3.1). Then we define in Section 3.2 a natural extension of this estimator to aggregated Shapley effects. In the targeted application to snow avalanche modeling (see Section 5), the sample is obtained from an acceptance rejection algorithm. Therefore we need a so-called given data estimator, that is an estimator that can be computed with any given input/output sample.

3.1. *Given data estimator of scalar Shapley effects proposed in [13]*

Let  $\mathbf{u} \subseteq \{1, \dots, d\}$ . As noticed in [15, Theorem 1], replacing the characteristic function  $\mathbf{u} \mapsto \tilde{c}_j(\mathbf{u}) = \text{Var}(\mathbb{E}(Y_j | \mathbf{X}_{\mathbf{u}}))$  by  $\mathbf{u} \mapsto c_j(\mathbf{u}) = \mathbb{E}(\text{Var}(Y_j | \mathbf{X}_{-\mathbf{u}}))$  in Equation (1) does not change the definition of Shapley effects. The estimation of the Shapley effects  $\text{Sh}_i^j$ ,  $1 \leq i \leq d$ , then relies on the estimation of the  $c_j(\mathbf{u})$ 's for all  $\mathbf{u} \subseteq \{1, \dots, d\}$ . To estimate  $c_j(\mathbf{u}) = \mathbb{E}(\text{Var}(Y_j | \mathbf{X}_{-\mathbf{u}}))$ , it is usual to use a double Monte Carlo estimator: an inner loop to estimate the variance and an outer loop to estimate the expectation. In the following, we estimate scalar Shapley effects with the estimator introduced in [13] which combines a subset aggregation procedure with a given data adaptation of the double Monte Carlo estimator of  $c_j(\mathbf{u})$ .

The estimator introduced in [13, Section 6] works as follows: assume we are given a set of  $n$  independent random samples  $(\mathbf{X}^{(k)}, \mathbf{Y}^{(k)})_{1 \leq k \leq n}$  of  $(\mathbf{X}, \mathbf{Y})$ , then for  $1 \leq j \leq p$ ,  $\emptyset \subsetneq \mathbf{u} \subsetneq \{1, \dots, d\}$ , let  $1 \leq N_{\mathbf{u}} \leq n$  be an integer, let  $(s_{\ell})_{1 \leq \ell \leq N_{\mathbf{u}}}$  be a sample of uniformly distributed integers in  $\{1, \dots, n\}$  (with or without replacement). The inner Monte Carlo loop estimates the conditional variance  $\text{Var}(Y_j | \mathbf{X}_{-\mathbf{u}})$  using  $N_I$  nearest neighbors of each  $\mathbf{X}_{-\mathbf{u}}^{s_{\ell}}$ ,  $1 \leq \ell \leq N_{\mathbf{u}}$ :

$$\widehat{E}_{\mathbf{u}, s_{\ell}}^j = \frac{1}{N_I - 1} \sum_{i=1}^{N_I} \left( \sum_{v: \mathbf{X}_{-\mathbf{u}}^v \in \mathcal{B}_{-\mathbf{u}, \ell}} Y_j^v - \frac{1}{N_I} \bar{Y}_{s_{\ell}} \right)^2 \quad (4)$$

with  $\bar{Y}_{s_{\ell}} = \frac{1}{N_I} \sum_{v: \mathbf{X}_{-\mathbf{u}}^v \in \mathcal{B}_{-\mathbf{u}, \ell}} Y_j^v$  and  $\mathcal{B}_{-\mathbf{u}, \ell}$  the set of  $N_I$  closest neighbors of

$\mathbf{X}_{-\mathbf{u}}^{s_{\ell}}$ . If there are ex aequo among the neighbors, one draws randomly without replacement. The closest neighbors are computed using the Euclidean distance in  $\mathbb{R}^{k_{\mathbf{u}}}$  with  $k_{\mathbf{u}}$  denoting the cardinality of  $-\mathbf{u}$ . Then, we define the outer Monte Carlo loop to estimate the expectation as:

$$\widehat{c}_j(\mathbf{u}) = \frac{1}{N_{\mathbf{u}}^*} \sum_{\ell=1}^{N_{\mathbf{u}}^*} \widehat{E}_{\mathbf{u}, s_{\ell}}^j. \quad (5)$$

Note that  $c_j(\emptyset) = 0$  and  $c_j(\{1, \dots, d\}) = \text{Var}(Y_j)$ , which is assumed to be known in [13], and that is estimated by the empirical variance in the present paper. In [13, Theorem 6.6.], it is proven that under mild assumptions,  $\widehat{c}_j(\mathbf{u})$  converges in probability to  $c_j(\mathbf{u})$  as  $n$  and  $N_{\mathbf{u}}^*$  go to  $\infty$ . Finally, the subset aggregation procedure consists in estimating scalar Shapley effects by plugging (5) in Equation (1):

$$\widehat{\text{Sh}}_i^j = \frac{1}{d \widehat{\sigma}_j^2} \sum_{\mathbf{u} \subseteq -i} \binom{d-1}{|\mathbf{u}|}^{-1} (\widehat{c}_j(\mathbf{u} \cup \{i\}) - \widehat{c}_j(\mathbf{u})) \quad (6)$$

where  $\widehat{\sigma}_j^2$  is the empirical estimator of  $\text{Var}(Y_j)$ . Note that the choice of the Euclidean distance for the computation of closest neighbors has been done arbitrarily as the most simple choice, and that future work could expand the

framework to other distances and investigate to which extent results are sensitive to this choice.

**Remark 3.1.** *In this remark, we discuss the choice of the numbers  $N_{\mathbf{u}}$ ,  $\mathbf{u} \subseteq \{1, \dots, d\}$ . These integers are related to the estimation accuracy of the expectation of the conditional variances  $\mathbb{E}(\text{Var}(Y_j | \mathbf{X}_{-\mathbf{u}}))_{\emptyset \subsetneq \mathbf{u} \subsetneq \{1, \dots, d\}}$  and are referred to as the accuracies hereafter. If all the  $N_{\mathbf{u}}$  were chosen equal to  $n$ , the computation cost for the search of the nearest neighbors would be in  $\mathcal{O}(n(2^d - 2))$ . To reduce this cost to a prescribed computational budget  $N_{tot}$ , with  $2^d - 2 \leq N_{tot}$  much smaller than  $n(2^d - 2)$ , the authors in [13] suggested to choose the set of accuracies that minimize the estimation variance for a prescribed total cost  $N_{tot} = \sum_{\emptyset \subsetneq \mathbf{u} \subsetneq \{1, \dots, d\}} N_{\mathbf{u}}$ . Note that for a given data strategy, we do not measure the cost in terms of number of model evaluations, as the sample is already available. Here the cost is measured in terms of nearest-neighbor search  $N_{tot}$ . In the framework where it is possible to sample from the conditional distributions of the input vector, the optimal accuracies [13, Proposition 4.2., Remark 4.3.] are given by  $N_{\mathbf{u}}^* = \text{Round}\left(N_{tot} \binom{d}{|\mathbf{u}|}^{-1} (d-1)^{-1}\right)$ , for all  $\emptyset \subsetneq \mathbf{u} \subsetneq \{1, \dots, d\}$ , with for  $x \in \mathbb{R}$ ,  $\text{Round}(x)$  denoting the integer greater than one closest to  $x$ . In the given data framework, we compared this choice of accuracies with the uniform one,  $N_{\mathbf{u}} = \text{Round}(N_{tot}/(2^d - 2))$  for all  $\emptyset \subsetneq \mathbf{u} \subsetneq \{1, \dots, d\}$ , for a test case in dimension  $d = 10$ . The accuracies  $N_{\mathbf{u}}^*$  suggested in [13] led to better results, at least on that experiment. The results of these tests are available as supplementary material.*

**Remark 3.2.** *Regarding the number of neighbors, the authors in [15] proved that in the usual double loop setting, choosing  $N_I = 3$  is better than choosing  $N_I = 2$  and that the gain in increasing  $N_I$  from 3 decreases as  $N_I$  increases. In our experiments we chose  $N_I = 3$  although the best choice in the given data framework is an open question, that should deserve great attention in future research.*

### 3.2. Estimator of the aggregated Shapley effects

Given scalar Shapley effect estimators whose definition is recalled in the previous section, we propose to estimate aggregated Shapley effects by:

$$\begin{aligned} \widehat{\text{GSh}}_i &= \frac{\sum_{j=1}^p \widehat{\sigma}_j^2 \widehat{\text{Sh}}_i^j}{\sum_{j=1}^p \widehat{\sigma}_j^2} \\ &= \frac{1}{d \sum_{j=1}^p \widehat{\sigma}_j^2} \sum_{j=1}^p \sum_{\mathbf{u} \subseteq -i} \binom{d-1}{|\mathbf{u}|}^{-1} (\widehat{c}_j(\mathbf{u} \cup \{i\}) - \widehat{c}_j(\mathbf{u})) \end{aligned} \quad (7)$$

with  $\widehat{\sigma}_j^2$  the empirical estimator of  $\text{Var}(Y_j)$  and  $\widehat{c}_j(\mathbf{u})$  defined by (5).

### 3.3. Dimension reduction: functional principal component analysis

For space or time-dependent models, a preliminary step consists in performing a functional principal component analysis (fPCA). Then we compute the aggregated Shapley effects for the vectorial output composed with the first principal components. Most of the time, it allows easier interpretation for real applications. KL expansion is obtained by applying the principal component analysis through conditional expectation (PACE) method introduced in [24].

## 4. Block bootstrap confidence intervals

We propose in this section to construct confidence intervals with a block bootstrap procedure, extending to the given data framework the procedure proposed in [16]. Bootstrap by blocks is necessary to preserve the nearest-neighbor structure in Equation (4). Algorithm 4.1 describes more precisely the block bootstrap procedure.

---

**Algorithm 4.1**  $B$  bootstrap samples for  $\widehat{\text{Sh}}_i^j$  and  $\widehat{\text{GSh}}_i$

---

**Inputs:** (i) A  $n$  i.i.d. random sample  $(\mathbf{X}^k, \mathbf{Y}^k)_{k \in \{1, \dots, n\}}$  with  $\mathbf{X}^k \in \mathbb{R}^d$  and  $\mathbf{Y}^k \in \mathbb{R}^p$ . (ii) For each  $\emptyset \subsetneq \mathbf{u} \subsetneq \{1, \dots, d\}$ , a  $N_{\mathbf{u}}$  random sample  $(s_{\ell})_{1 \leq \ell \leq N_{\mathbf{u}}}$  from  $\{1, \dots, n\}$ .

**Outputs:**  $B$  bootstrap samples for  $\widehat{\text{Sh}}_i^j$  and  $\widehat{\text{GSh}}_i$ .

**for**  $b = 1$  to  $b = B$  **do**

1. Create a  $n$  bootstrap sample  $\mathbf{Y}^{(b)}$  by sampling with replacement from the rows of  $\mathbf{Y}$ .

2. Compute, for  $1 \leq j \leq p$ ,  $\widehat{\sigma}_j^{2,(b)}$  the empirical variance of  $\mathbf{Y}_j^{(b)}$ .

3. For each  $j \in \{1, \dots, p\}$ :

3.1. For all  $\mathbf{u}$  and for all  $(s_{\ell})_{1 \leq \ell \leq N_{\mathbf{u}}}$  compute  $\widehat{E}_{\mathbf{u}, s_{\ell}}^j$  using (4).

3.2. For all  $\mathbf{u}$ , create a  $N_{\mathbf{u}}$  bootstrap sample  $\widehat{E}_{\mathbf{u}, s_{\ell}}^{j,(b)}$  by sampling with replacement from  $(\widehat{E}_{\mathbf{u}, s_{\ell}}^j)_{1 \leq \ell \leq N_{\mathbf{u}}}$  computed in Step 3.1.

3.3. Compute  $\widehat{c}_j^{(b)}(\mathbf{u}) = 1/N_{\mathbf{u}} \sum_{\ell=1}^{N_{\mathbf{u}}} \widehat{E}_{\mathbf{u}, s_{\ell}}^{j,(b)}$  for all  $\mathbf{u}$  using (5).

3.4. Compute the  $b$  bootstrap sample of  $\widehat{\text{Sh}}_i^j$  according to (6):

$$\widehat{\text{Sh}}_i^{j,(b)} = \frac{1}{d \widehat{\sigma}_j^{2,(b)}} \sum_{\mathbf{u} \subseteq -i} \binom{d-1}{|\mathbf{u}|}^{-1} \left( \widehat{c}_j^{(b)}(\mathbf{u} \cup \{i\}) - \widehat{c}_j^{(b)}(\mathbf{u}) \right).$$

4. Compute the  $b$  bootstrap sample of  $\widehat{\text{GSh}}_i$  using (7):

$$\widehat{\text{GSh}}_i^{(b)} = \frac{1}{d \sum_{j=1}^p \widehat{\sigma}_j^{2,(b)}} \sum_{j=1}^p \sum_{\mathbf{u} \subseteq -i} \binom{d-1}{|\mathbf{u}|}^{-1} \left( \widehat{c}_j^{(b)}(\mathbf{u} \cup \{i\}) - \widehat{c}_j^{(b)}(\mathbf{u}) \right).$$

**end for**

---



It is the first time a block bootstrap procedure for given data estimation of (aggregated) Shapley effects is proposed. For a scalar output, only Steps 1 to 3 of Algorithm 4.1 should be used, with the block bootstrap procedure described by Steps 3.1 to 3.3.

## 5. Application to Snow avalanche modeling

In this section, we present the implementation of the whole procedure on our real case-study. However, preliminary to this implementation, we conducted numerical experiments on toy models to evaluate the impact of the total budget  $N_{tot}$  and of the numbers  $N_u^*$  on the accuracy of Shapley effect estimation as far as on the coverage probability of bootstrap confidence intervals. The results we obtained are provided as supplementary material.

### 5.1. Model

Our snow avalanche model is based on shallow-water (depth-averaged) Saint-Venant equations and considers the avalanche as a fluid in motion. In more detail, it considers only the dense layer of the avalanche. The flow depth is then small compared to its length. The model assumes the avalanche is flowing on a curvilinear profile  $z = l(x)$ , where  $z$  is the elevation and  $x$  is the projected distance measured from the top of the avalanche path. Under these assumptions, shallow-water approximations of the mass and momentum equations can be used:

$$\begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial hv}{\partial x} &= 0 \\ \frac{\partial hv}{\partial t} + \frac{\partial}{\partial x} \left( hv^2 + \frac{h^2}{2} \right) &= h(g \sin \theta - F) \end{aligned}$$

where  $v = \|\vec{v}\|$  is the flow velocity,  $h$  is the flow depth,  $\theta$  is the local angle,  $t$  is the time,  $g$  is the gravity constant and  $F = \|\vec{F}\|$  is a frictional force. The model uses the classical Voellmy (see, e.g., [25]) frictional force  $F = \mu g \cos \theta + g/(\xi h)v^2$ , where  $\mu$  and  $\xi$  are friction parameters. This formulation remains partially ad-hoc, and arguably oversimplified with regards to the complexity of real snow flows (see, e.g., [26]). However, it is still by far the most common choice in the snow and avalanche community (see, e.g., [27]), as it provides robust simulation results in many configurations.

The partial differential equation (PDE) system is solved numerically using a finite volume scheme on the path curvilinear profile with a 5m resolution. Hence, even if the PDEs are expressed in a two-dimensional frame, they in fact represent a one-dimensional flow on the curvilinear profile (see, e.g., [28]). For each avalanche simulation, standard boundary conditions detailed in the original paper [7] where used, and it has been established in [29] that the choice of boundary conditions generally does not have many influence on the simulation results.

The numerical model depends on six inputs: the friction parameters  $\mu$  and  $\xi$ , the length  $l_{\text{start}}$  of the avalanche release zone, the snow depth  $h_{\text{start}}$  within the release zone, the abscissa corresponding to the beginning of the release zone denoted by  $x_{\text{start}}$  and the discretized topography of the flow path, denoted by  $D = (\mathbf{x}, \mathbf{z}) \in \mathbb{R}^{N_s \times 2}$  where  $\mathbf{x} \in \mathbb{R}^{N_s}$  is the vector of projected abscissa positions and  $\mathbf{z} = l(\mathbf{x}) \in \mathbb{R}^{N_s}$  is the elevation vector.  $N_s$  is the number of points of the discretized path. The model outputs are the flow velocity, flow depth trajectories in the path  $D$  and runout distance of an avalanche. Note that the model has two functional and one scalar outputs and these three outputs are the objects of the GSA study. Specifically, for each avalanche, the model evaluates velocity and flow depth all along the path topography at each time step of the simulation, which leads, at the end of the simulation, a  $N_s \times T_{sim}$  flow depth and velocity field, with  $T_{sim}$  the number of time steps of the simulation. From this set of outputs, we retain and analyse, at each position of the path, the maximal velocity and maximal flow depth only (namely two vectors of size  $N_s$  for each simulation).

We use for  $D$  the topography of a path located in Bessans, France. This particular path already considered in other works [30, 8, 10, 31] is well documented in the French avalanche database [32]. We develop our GSA in two contexts or scenarios: (i) little knowledge on the input parameter probability distribution, and (ii) well-calibrated input parameter distributions.

Note eventually that we do not fully distinguish within both scenarios different types of avalanches (e.g. involving wet/dry snow). The reason is that from a hazard assessment perspective for settlements, one needs to consider all avalanches likely to occur, independently from the type of the flow/snow. As a consequence, the distributions of friction parameters used in Scenarios 1 and 2, notably of friction parameter  $\mu$  that represents in first approximation snow quality in Voellmy friction law, primarily correspond to the whole local avalanche activity altogether, either from expert considerations (Scenario 1) or according to the local data (Scenario 2). For both scenarios, our range of friction coefficient values is well in accordance with large data sets from the literature [33]. Also, for the case study, it is important to acknowledge that the data set used to obtain the posterior distribution of scenario 2 indeed includes both dry and wet snow avalanches. However, as dry snow avalanches with low values of friction parameter  $\mu$  generally have the longest runout distances and the highest velocity, scenario 2 that aims at investigating sensitivity in local hazard assessment gives more weight to them by excluding avalanches with high friction values.

## 5.2. Scenario 1

### 5.2.1. Principle

We first determine the most influential input parameters sampled from uniform distributions. We thus expect from the GSA a better understanding of the numerical model. Inputs  $\mu$  and  $\xi$  vary in their physical value ranges. Inputs  $l_{\text{start}}$  and  $h_{\text{start}}$  vary in their spectrum of reasonable values given the characteristics of the avalanche path. The  $x_{\text{start}}$  input distribution is determined by

Input	Description	Distribution
$\mu$	Static friction coefficient	$\mathcal{U}[0.05, 0.65]$
$\xi$	Turbulent friction [m/s <sup>2</sup> ]	$\mathcal{U}[400, 10000]$
$l_{\text{start}}$	Length of the release zone [m]	$\mathcal{U}[5, 300]$
$h_{\text{start}}$	Mean snow depth in the release zone [m]	$\mathcal{U}[0.05, 3]$
$x_{\text{start}}$	Release abscissa [m]	$\mathcal{U}[0, 1600]$

Table 1: Avalanche model, scenario 1: Input description and uncertainty intervals. In the GSA, we consider  $\text{vol}_{\text{start}} = l_{\text{start}} \times h_{\text{start}} \times 72.3 / \cos(35^\circ)$  instead of  $h_{\text{start}}$  and  $l_{\text{start}}$ .

computing the abscissa interval where the release zone average slope is superior to  $30^\circ$ . Indeed, the slope remains above  $30^\circ$  during the first 1600m of the path. A good approximation of avalanche release zones is commonly obtained this way. Since different studies [27, 34] suggest that the volume of snow is a critical quantity that controls flow dynamics, we consider  $\text{vol}_{\text{start}}$  as input of the GSA instead of  $h_{\text{start}}$  and  $l_{\text{start}}$ . The latter is evaluated as  $\text{vol}_{\text{start}} = l_{\text{start}} \times h_{\text{start}} \times 72.3 / \cos(35^\circ)$ . The mean width and slope of the release zone equal to 72.3m and  $35^\circ$ , respectively. All uncertainty intervals are summed-up in Table 1. The original input correlations are close to 0 since the original inputs were drawn independently. They are not exactly equal to zero as they are estimated from a sample.

For a given avalanche simulation, its functional velocity and flow depth outputs have a high number of zeros because they are null above the beginning of the release zone and after the runout position. Also, there might be some avalanche simulations that are meaningless in physical terms and/or not useful to assess the related risk. Therefore to perform GSA, we select simulations that accomplish the following rules:

- (i) avalanche simulation is flowing in the interval  $[1600m, 2412m]$ ,
- (ii) its volume is superior to  $7000 \text{ m}^3$  and,
- (iii) avalanche runout distance is inferior to 2500m which corresponds to the end of the path.

In the following, we use the terminology *constrained sample* to deal with the sample obtained after applying rules (i) to (iii). This set of constraints makes that we switch from a large sample of simulated avalanches to a much smaller one, and the constraints by the runout distance play a critical role in this reduction. Indeed, even if most of the original avalanches are able to flow rather far downwards in the path due to the slope, we analyse in this GSA only those that are able to travel beyond the flat area at the bottom of the path, and, hence, represent the highest risk for people. Specifically, for the studied path, the return period of avalanches in the interval  $[1600m, 2412m]$  varies from 1 to 10,000 years according to [8] and local data (not used in this scenario). Roughly speaking, a return period is the mean time in which a given runout distance is reached or exceeded at a given path's position. Hence, this GSA focuses on

the most extreme, yet possible, snow avalanches. In addition, the constraint on snow volumes makes that the GSA focuses on medium to very large snow avalanche sizes, which have a high damage potential.

### 5.2.2. Global sensitivity analysis results

We first ran  $n_0 = 100\,000$  avalanche simulations from an i.i.d. sample of input distributions described in Table 1. Then, by applying (i) to (iii) the size of the *constrained sample* was reduced to  $n_1 = 6152$  (Appendix A.2). Even if the initial sample did not present any significant correlation structure, *the constrained sample* does. For example, inputs  $\mu$  and  $\xi$  were independent for the initial sample but the correlation computed after applying the constraints is about 0.31. Note that the input parameter correlations induced by the constraints were the main motivation to compute Shapley effects and not Sobol' indices in this first scenario.

In Appendix A.2 are plotted highest density region (HDR) boxplots for the velocity and flow depth, obtained by using the R package `rainbow` developed by [35]. The HDR boxplot is a visualization tool for functional data based on kernel density estimation of the scores associated to the two first principal components of the functional data (see [36] for further details). In the data we consider in our study, the avalanche velocity ranges from  $0.1\text{ms}^{-1}$  to  $71.56\text{ms}^{-1}$  and avalanches are decelerating (see Figure A.8 a). Flow depths vary from 0.03m to 7.52m. The flow depth curves exhibit high fluctuations in the interval  $[2100\text{m}, 2300\text{m}]$  (see Figure A.8 b) which corresponds to a region where path's topography is mostly convex. Runout distances vary from 2409m to 2484m (see Figure A.8 c). The focus, for each simulation, on maximal velocity and flow depth at each position of the path explains the large velocity and flow depth values that are attained at many positions of the path, but with a strong variability from one retained avalanche to another. These high values and high variability between all avalanche events potentially occurring on the studied path are in full consistency with classical observations of real snow avalanches (e.g., [37]). By contrast, the used physical constraints define a specific population of retained snow avalanches in terms of runout distances, with a lower variability from one retained avalanche to another.

On Figure 1 panels a and b, ubiquitous (pointwise) Shapley effects of velocity and flow depth curves are shown, respectively. Depending on the output, results are quite different. For velocity,  $x_{start}$  is the most relevant during a large part of the path but its importance decreases along the path and, conversely, the importance of the other inputs increases. For the flow depth output, the most important input is  $vol_{start}$  since the corresponding Shapley effects vary from 0.4 to 0.2 along the path. Nevertheless, other inputs are not completely negligible. Input importance also varies according to the topography. In fact, the ubiquitous effect variation corresponds to local slope changes (see Figure 1 a and b). Correlations between ubiquitous effects and local slope have been computed and are rather high. For example, for the velocity, the absolute value of the correlation is higher than 0.51 for all input parameters. This implies that local slope changes play an important role on the input contribution to output

variations, a nice result showing the relevance of the GSA analysis to understand the dynamical properties of the flow. Eventually, for runout distance, the four inputs appear as relevant, but the uncertainty regarding Shapley values is very high. This is presumably a consequence of the constraints we use, since the latter considerably reduce the variance of runout distances with regards to the original sample. This may preclude investigating in details to which extent the different inputs control this specific output.

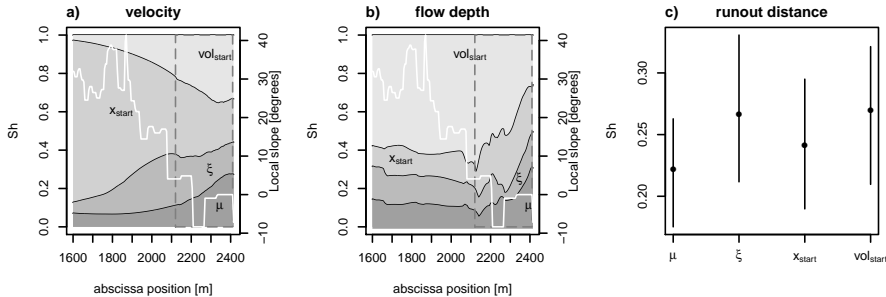


Figure 1: Avalanche model, scenario 1: a) and b) ubiquitous Shapley effects of velocity and flow depth curves, resp. and, c) runout distance Shapley effects. Shapley effects are estimated with a sample of size 6152 and  $N_{tot}=2002$ . The local slope is displayed with a white line. A gray dotted rectangle box is displayed at interval [2017, 2412] where snow avalanche return periods vary from 10 to 10 000 years. The number of bootstrap samples is fixed to  $B = 500$ .

Figure 2 (resp., 3) shows the correlations between the fPCs and the velocity (resp., flow depth) output for scenario 1. The first fPC has a positive correlation with the velocity, always higher than 0.57, however the correlation decreases drastically from the top to the bottom of the path. By contrast, the second fPC is less correlated to the velocity, the correlation is positive at the top of the path, then the correlation decays and even becomes highly negative for  $x \geq 2300$  m. Hence, these two components identify well two regions of the path where velocity behaves differently: the first fPC corresponds to the release and transit area of the path, where avalanches accelerate and maintain their velocity while the second fPC identifies the runout area where avalanches are decelerating and eventually stop. The behavior is similar for the flow depth, with a sharp peak in the second fPC corresponding to the location in the runout area where slope becomes null and then negative. To strengthen our physical interpretation, we estimated the scalar Shapley effects for each fPC with  $n = 6152$  and  $N_{tot} = 2002$ . As far as velocity is concerned, for the first fPC,  $x_{start}$  is the most important input and for the second fPC,  $vol_{start}$ ,  $x_{start}$  and  $\mu$  are the most relevant. For the flow depth, for the first fPC,  $vol_{start}$  is the most relevant. For the second fPC, the inputs  $x_{start}$ ,  $vol_{start}$  and  $\mu$  are the most important inputs.

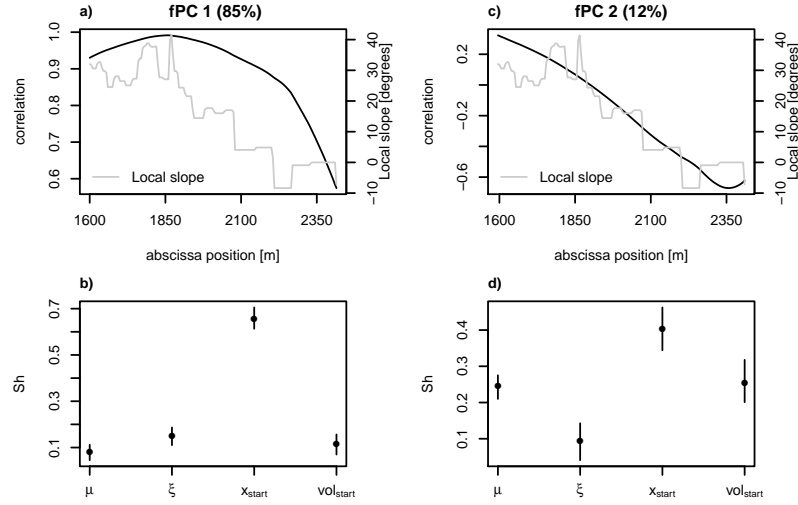


Figure 2: Avalanche model, scenario 1: a) and c) correlations between the first and second fPCs and the velocity. Scalar Shapley effects for b) fPC 1 and d) fPC 2 are estimated with  $n = 6152$  and  $N_{tot} = 2002$ . For the confidence intervals,  $B = 500$  is used. The local slope is displayed with a gray line.

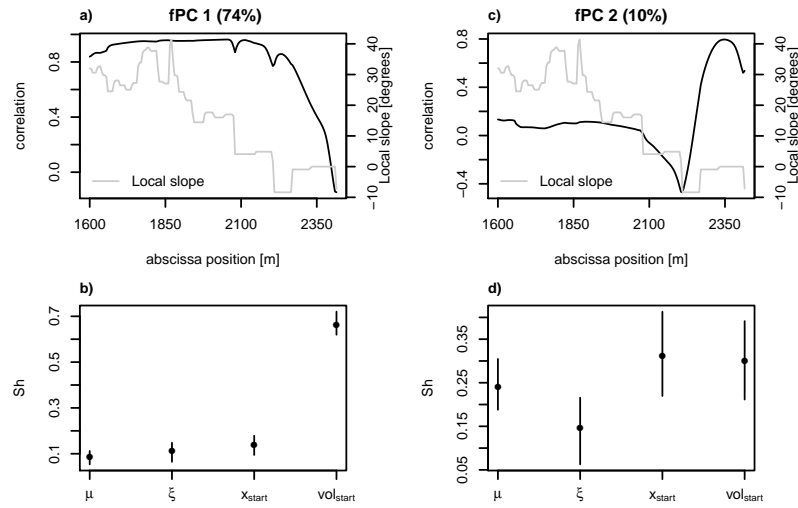


Figure 3: Avalanche model, scenario 1: a) and c) correlations between the first and second fPCs and the flow depth output. Scalar Shapley effects for b) fPC 1 and d) fPC 2 are estimated with  $n = 6152$  and  $N_{tot} = 2002$ . For the confidence intervals,  $B = 500$  is used. The local slope is displayed with a gray line.

Figure 4 shows aggregated Shapley effects and 90% confidence intervals computed over space intervals  $[x, 2412]$  where  $x \in \{1600, 1700, \dots, 2412\}$ . The aggregated effects are computed in the first fPCs explaining more than 95% of the output variance. Aggregated effects seem more robust than ubiquitous effects, specially in zones where local slope shows high variations (see Figure 4 compared to Figure 1). For explaining more than 95% of the velocity output variance, 2 fPCs are required, while, for explaining more than 95% of the flow depth output variance, at most 4 fPCs are required, depending on  $x$ . For the velocity output, the most important input is  $x_{\text{start}}$  in the interval  $[1600m, 2100m]$  but its importance decreases along the path. In the interval  $[2017m, 2412m]$  where return periods are non trivial,  $x_{\text{start}}$  and  $\text{vol}_{\text{start}}$  are the most important followed by  $\mu$  and  $\xi$ . For the flow depth output,  $\text{vol}_{\text{start}}$  is the most relevant but its importance decreases along the path. At the end of the path from 2300m to 2412m where return periods are high (between 100 to 10 000 years), confidence intervals intersect. It seems thus difficult to deduce a clear ranking of the inputs for these last portions of the path. Nevertheless, it seems that none of the inputs is negligible, even at the very end of the path. In summary, to estimate velocities with accuracy, the release zone and volume are the most important parameters and, for the flow depth, a good approximation of the volume released is essential.

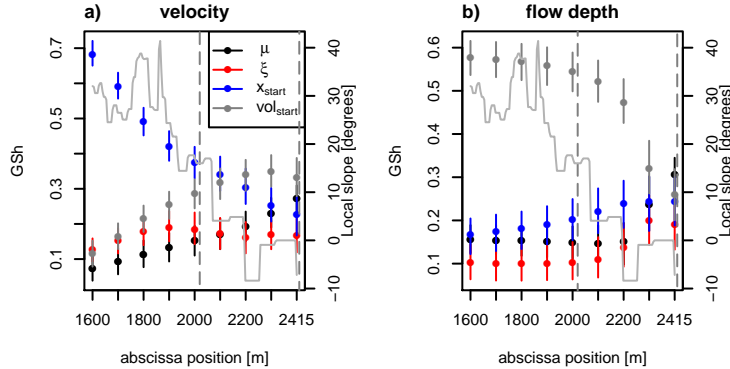


Figure 4: Avalanche model, scenario 1: a) and b) aggregated Shapley effects of velocity and flow depth curves calculated over space intervals  $[x, 2412m]$  where  $x \in \{1600m, 1700m, \dots, 2412m\}$ . Shapley effects are estimated with a set of independent samples of size 6152 and  $N_{\text{tot}}=2002$ . Effects are estimated using the first fPCs explaining more than 95% of the output variance. The local slope is displayed with a gray line. A gray dotted rectangle is displayed at  $[2017m, 2412m]$  where snow avalanche return periods vary from 10 to 10 000 years. The number of bootstrap samples is fixed to  $B = 500$ .

### 5.3. Scenario 2

#### 5.3.1. Principle

The aim is now to determine the most influential inputs in the context of strong knowledge regarding input distributions. In [8], the authors developed

a Bayesian framework to estimate input distributions from available avalanche observations. The objective is long-term avalanche hazard assessment in order to later assess the related risk for buildings and people. In the avalanche literature, it is assumed that  $\xi$  depends on the path topography, so it is a parameter and not a variable varying from one avalanche to another in [8]’s model. The input  $\xi$  is therefore fixed to its posterior estimate, 1300. Other input variables in this scenario are dependent. The dependence between  $h_{\text{start}}$  and  $l_{\text{start}}$  is modeled with a linear function  $l_{\text{start}} = 31.25 + 87.5h_{\text{start}}$ , and similarly as in scenario 1, we consider  $\text{vol}_{\text{start}}$  as input of the GSA instead of  $h_{\text{start}}$  and  $l_{\text{start}}$ . The complete input distribution resulting from the Bayesian inference on the studied path is described in Table 2. Input correlations have been computed. As an example, the correlation between  $\mu$  and  $\text{vol}_{\text{start}}$  is 0.8. Then, as a preliminary step to GSA, we apply the following constraints:

- (i) avalanche is flowing in the interval [1600m, 2204m] where snow avalanche return periods vary from 10 to 300 years, according to the work of [8] and local data (implicitly used in this scenario for model calibration),
- (ii) avalanche volume is superior to 7 000 m<sup>3</sup> and,
- (iii)  $\mu$  coefficient is inferior to 0.39 so as to focus on dry snow avalanches with long runouts and high velocities.

Under these conditions, we sample a large set of dry snow avalanches that could cause material or human damages on the studied site according to the available local data, namely avalanches fully relevant for land-use planning purposes.

Input	Distribution
$x_{\text{nstart}} = \frac{x_{\text{start}}}{1600}$	Beta(1.38, 2.49)
$h_{\text{start}}   x_{\text{nstart}}$	Gamma( $\frac{1}{0.45^2}(1.52 + 0.03x_{\text{nstart}})^2$ , $\frac{1}{0.45^2}(1.52 + 0.03x_{\text{nstart}})$ )
$l_{\text{start}}$	$31.25 + 87.5h_{\text{start}}$
$\mu   h_{\text{start}}, x_{\text{nstart}}$	$\mathcal{N}(0.449 - 0.013x_{\text{nstart}} + 0.025h_{\text{start}}, 0.11^2)$

Table 2: Avalanche model: Scenario 2. Input description and uncertainty intervals.  $x_{\text{nstart}}$  is a normalization of  $x_{\text{start}}$ . A linear relationship between  $h_{\text{start}}$  and  $l_{\text{start}}$  inferred from the local data is used [8]. Gamma distribution parameters are its shape and rate. In the the GSA, we consider  $\text{vol}_{\text{start}} = l_{\text{start}} \times h_{\text{start}} \times 72.3 / \cos(35^\circ)$  instead of  $h_{\text{start}}$  and  $l_{\text{start}}$ .

### 5.3.2. Global sensitivity analysis results

We first ran  $n_0 = 100\,000$  avalanches from an i.i.d. sample of input distribution following Table 2. After applying the constraints, the sample size was reduced to  $n_2 = 1284$  and the input distribution was modified. For example,  $\mu$  and  $\text{vol}_{\text{start}}$  correlation changes from 0.8 to 0.2 which is still non negligible. Ubiquitous Shapley effects are displayed on Figure 5 panels a and b. For the velocity, the three inputs have a similar importance till 1900m, then  $\text{vol}_{\text{start}}$  importance decreases and  $\mu$  and  $x_{\text{start}}$  importance increases (see Figure 5 a). Similarly, as in scenario 1, the effects show fluctuations which correspond to changes in local slope, which clearly appears on fPCs (Appendix A.3). In particular, for the flow depth, input effects suffer radical changes when the local



slope decreases from  $20^\circ$  to  $10^\circ$  (see Figure 5 b). For the runout distance, all inputs appear as relevant, or, at least, difficult to discriminate (see Figure 5 c). Again, this may be linked to the set of constraints that targets a specific population of snow avalanches, with a rather low variability in terms of runout distances from one avalanche to another.

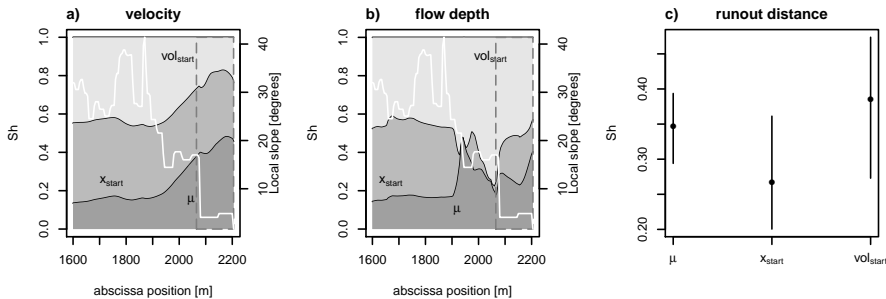


Figure 5: Avalanche model, scenario 2: a) and b) ubiquitous Shapley effects of velocity and flow depth curves, c) runout distance Shapley effects. Shapley effects are estimated with a set of independent samples of size 1284 and  $N_{tot}=800$ . The local slope is displayed with a white line. A gray dotted rectangle shows the interval [2064, 2204] where return periods vary from 10 to 300 years. The number of bootstrap samples is fixed to  $B = 500$ .

Aggregated effects (see Figure 6) present less fluctuations and are easier to interpret (see Figure 5). In summary, under this second scenario, it is fundamental to have a good approximation of the released volume and abscissa for velocity forecasting, while for flow depth forecasting, a good approximation of released volume is desirable. Nevertheless, none of the other inputs are negligible. Note that the uncertainty associated to the estimation of Shapley effects at 2204m is high (see the width of the corresponding confidence intervals on Figure 6). To outperform the estimation accuracy at the end of the path generating a larger initial sample of avalanches is possible, but the computational burden is prohibitive.

#### 5.4. Outcomes for avalanche modelling, related hazard, assessment and outlooks

For both scenarios, our GSA results are in overall accordance with previous works documenting the sensitivity of snow avalanche models to their input conditions using less formalised methods [38]. They also relate to more general results regarding snow avalanche dynamics. For instance, it is well known that in the transit area of the path, an avalanche accelerates up to a maximal velocity and this limit velocity can be explicitly computed on a constant slope under the assumption of a Voellmy friction law [39, 40, 41]. More generally, with depth-averaged flow equations, it is known that initial conditions corresponding to the avalanche release are important only in the early phases of the flow and are forgotten later, especially in the runout area where velocity and flow depths become primarily constrained by local slope. It is therefore all the more logical that the influence of  $x_{start}$  on velocity decreases as the avalanche flows downslope.

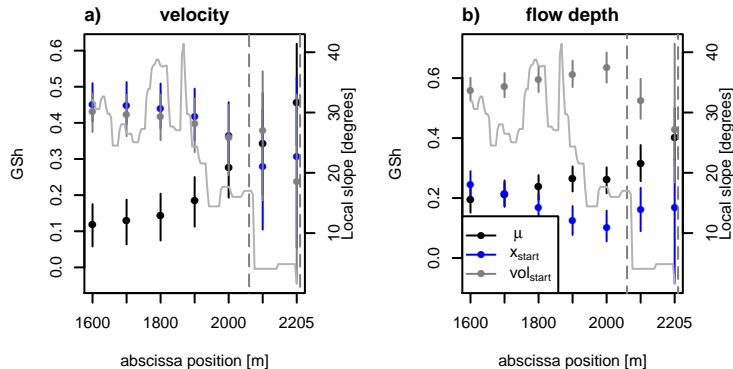


Figure 6: Avalanche model, scenario 2: a) and b) aggregated Shapley effects of velocity and flow depth curves calculated over space intervals  $[x, 2204]$  where  $x \in \{1600, 1700, \dots, 2204\}$  and using the first fPCs which have 95% of output variance. Shapley effects are estimated with a set of independent samples of size 1284 and  $N_{tot}=800$ . The local slope is displayed with a gray line. A gray dotted rectangle is displayed at  $[2017\text{m}, 2204\text{m}]$  where return periods vary from 10 to 300 years. The number of bootstrap samples is fixed to  $B = 500$ .

However, in the snow avalanche literature, it is thought that runout distances is the variable that is primarily controlled by the friction parameters, and the velocity/flow depths much less [42, 27]. By contrast, our results suggest that things are not that clear, with the different outputs being controlled by friction parameters and other avalanche model input parameters with rather similar importance. Also, the fPCs convey new useful information about the avalanche model and its sensitivity, as demonstrated by the probative linkages between local slope and sensitivity indices we highlighted. This provides first concrete illustration of the interest of our approach for practitioners/physicists of the field. On this basis, our approach could/should now be applied to other avalanche paths, and to different avalanche models with different friction laws so as to generalize and widen the results obtained and, more widely, to elucidate new physics related to avalanche flows. For instance, the control of friction parameters by the different input variables, may indicate complex physics which is not well understood so far, and the Voellmy friction law we used is robust and widely used in the community but still oversimplified with regards to reality. Hence, conducting an in-depth sensitivity analysis within our framework with different, more complex, friction laws may well help to understand which terms of the friction model critically influence avalanche flows.

## 6. Conclusions and further perspectives

In this work, we introduced the notion of aggregated Shapley effects and we stated their main properties. We extended the subset aggregation procedure with double Monte Carlo given data estimator introduced in [13] to the estimation of aggregated Shapley effects. Also, we proposed an algorithm to

construct bootstrap confidence intervals for scalar and aggregated Shapley effects, extending the block bootstrap procedure introduced in [16] to the given data framework. Many interesting questions as the choice of the number of nearest neighbors  $N_I$  or the one of the accuracy numbers  $N_u$  are still open and represent challenging theoretical perspectives. From the applicative point of view, first insightful results regarding avalanche dynamics and potential related hazard, were obtained. Our analysis was focused on the computation of the complete multivariate distribution of all potential avalanches at the study site, namely on the hazard quantification step of risk assessment. However, avalanche risk estimates within runout areas can be directly evaluated from the output of the avalanche dynamic model [43, 44], combining them with vulnerability relations for buildings and people (e.g., [45]) so as to evaluate destruction and lethality rates. Our framework could be in the future used to conduct a complete sensitivity analysis in avalanche risk assessment taking into account both the hazard and vulnerability/exposure components of risk. This would complement existing results currently limited to separate investigations of avalanche risk sensitivity to vulnerability relations and runout distributions [10, 45]. More widely, global sensitivity analyses remain so far seldom and rough in the snow avalanche community and in the wider community of people working on gravitational hazards. Such analyses are yet urgently needed to better quantify and mitigate the related risks, notably to focus the model developments and simulation efforts on the most influential inputs. Our work i) proposes an approach having the theoretical properties required to handle the complexity of dependent functional inputs in full rigour, and ii) demonstrates its applicability and potential usefulness. It should therefore be seen as an important step forward likely to be used in the future by a wide range of practitioners of gravitational hazard domain (snow avalanches, landslides, rockfall, etc.).

## Funding

M.B. Heredia was granted by OSUG@2020 labex. Within the CDP Trajectories, this work was supported by the “Investissements d’avenir” program (ANR-15-IDEX-02) of the French National Research Agency. Computations were performed using Froggy platform of the CIMENT infrastructure <https://ciment.ujf-grenoble.fr> supported by Rhône-Alpes region (GRANT CPER07 13 CIRA), OSUG@2020 labex (ANR10 LABX56) and Equip@Meso project (ANR-10-EQPX-29-01) of the program “Investissements d’avenir”.

## Acknowledgments

We thank S. Da Veiga for fruitful discussions on nearest-neighbor search algorithms, E. Parent, T. Faug and P. Hagenmuller for advices on the avalanche application. We also thank the associate editor and anonymous reviewers for their fruitful comments on the paper.

## References

- [1] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, *Global sensitivity analysis: the primer*, John Wiley & Sons, 2008.
- [2] S. Da Veiga, F. Gamboa, B. Iooss, C. Prieur, *Basics and Trends in Sensitivity Analysis: Theory and Practice in R*, SIAM, 2021.
- [3] S. Razavi, A. Jakeman, A. Saltelli, C. Prieur, B. Iooss, E. Borgonovo, E. Plischke, S. L. Piano, T. Iwanaga, W. Becker, et al., The future of sensitivity analysis: an essential discipline for systems modeling and policy support, *Environmental Modelling & Software* 137 (2021) 104954.
- [4] M. Lamboni, H. Monod, D. Makowski, Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models, *Reliability Engineering & System Safety* 96 (4) (2011) 450 – 459.
- [5] 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.
- [6] A. Alexanderian, P. A. Gremaud, R. C. Smith, Variance-based sensitivity analysis for time-dependent processes, *Reliability Engineering & System Safety* 196 (2020) 106722.
- [7] M. Naaim, F. Naaim-Bouvet, T. Faug, A. Bouchet, Dense snow avalanche modeling: flow, erosion, deposition and obstacle effects, *Cold Regions Science and Technology* 39 (2) (2004) 193 – 204.
- [8] N. Eckert, M. Naaim, E. Parent, Long-term avalanche hazard assessment with a bayesian depth averaged propagation model, *Journal of Glaciology* 56 (2010) 563–586.
- [9] R. Schläppy, N. Eckert, V. Jomelli, M. Stoffel, D. Grancher, D. Brunstein, M. Naaim, M. Deschatres, Validation of extreme snow avalanches and related return periods derived from a statistical-dynamical model using tree-ring techniques, *Cold Regions Science and Technology* 99 (2014) 12–26.
- [10] P. Favier, N. Eckert, D. Bertrand, M. Naaim, Sensitivity of avalanche risk to vulnerability relations, *Cold Regions Science and Technology* 108 (2014) 163 – 177.
- [11] S. Kucherenko, O. V. Klymenko, N. Shah, Sobol’indices for problems defined in non-rectangular domains, *Reliability Engineering & System Safety* 167 (2017) 218–231.
- [12] A. Owen, Sobol’ indices and shapley value, *SIAM/ASA Journal on Uncertainty Quantification* 2 (1) (2014) 245–251.

- [13] B. Broto, F. Bachoc, M. Depecker, Variance Reduction for Estimation of Shapley Effects and Adaptation to Unknown Input Distribution, *SIAM/ASA Journal on Uncertainty Quantification* 8 (2) (2020) 693–716.
- [14] 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.
- [15] E. Song, B. Nelson, J. Staum, Shapley effects for global sensitivity analysis: Theory and computation, *SIAM/ASA Journal on Uncertainty Quantification* 4 (1) (2016) 1060–1083.
- [16] N. Benoumechiara, K. Elie-Dit-Cosaque, Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms, *ESAIM: Proceedings and Surveys* 65 (2019) 266–293.
- [17] E. Plischke, G. Rabitti, E. Borgonovo, Computing shapley effects for sensitivity analysis, *SIAM/ASA Journal on Uncertainty Quantification* 9 (4) (2021) 1411–1437.
- [18] T. Goda, A simple algorithm for global sensitivity analysis with shapley effects, *Reliability Engineering & System Safety* 213 (2021) 107702.
- [19] M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, H. Monod, Multivariate global sensitivity analysis for dynamic crop models, *Field Crops Research* 113 (3) (2009) 312–320.
- [20] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, Sensitivity indices for multivariate outputs, *Comptes Rendus Mathématique* 351 (7) (2013) 307 – 310.
- [21] A. B. Owen, C. Prieur, On shapley value for measuring importance of dependent inputs, *SIAM/ASA Journal on Uncertainty Quantification* 5 (1) (2017) 986–1002.
- [22] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, Sensitivity analysis for multidimensional and functional outputs, *Electronic Journal of Statistics* 8 (1) (2014) 575–603.
- [23] L. S. Shapley, A value for n-person games, in: H. W. Kuhn, A. W. Tucker (Eds.), *Contribution to the Theory of Games II* (*Annals of Mathematics Studies* 28), Princeton University Press, Princeton, NJ, 1953, pp. 307–317.
- [24] F. Yao, H.-G. Müller, J.-L. Wang, Functional data analysis for sparse longitudinal data, *Journal of the American Statistical Association* 100 (470) (2005) 577–590.
- [25] O. Buser, H. Frutiger, Observed maximum run-out distance of snow avalanches and the determination of the friction coefficients  $\mu$  and  $\xi$ , *Journal of Glaciology* 26 (94) (1980) 121–130.

- [26] A. Köhler, J. McElwaine, B. Sovilla, M. Ash, P. Brennan, The dynamics of surges in the 3 february 2015 avalanches in vallée de la sionne, *Journal of Geophysical Research: Earth Surface* 121 (11) (2016) 2192–2210.
- [27] P. Bartelt, Y. Bühler, O. Buser, M. Christen, L. Meier, Modeling mass-dependent flow regime transitions to predict the stopping and depositional behavior of snow avalanches, *Journal of Geophysical Research: Earth Surface* 117 (F1).
- [28] R. Greve, T. Koch, K. Hutter, Unconfined flow of granular avalanches along a partly curved surface. i. theory, *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 445 (1924) (1994) 399–413.
- [29] M. Naaim, Dense avalanche numerical modeling: interaction between avalanche and structures, in: *25 years of snow avalanche research*, Voss, NOR, 12-16 May 1998, 1998, pp. 187–191.
- [30] N. Eckert, É. Parent, T. Faug, M. Naaim, Optimal design under uncertainty of a passive defense structure against snow avalanches: from a general bayesian framework to a simple analytical model, *Natural Hazards and Earth System Sciences* 8 (5) (2008) 1067–1081.
- [31] N. Eckert, M. Naaim, F. Giacona, P. Favier, A. Lavigne, D. Richard, F. Bourrier, E. Parent, Repenser les fondements du zonage règlementaire des risques en montagne «récurrents», *La Houille Blanche* 2 (2018) 38–67.
- [32] E. Bourova, E. Maldonado, J.-B. Leroy, R. Alouani, N. Eckert, M. Bonnefoy-Demongeot, M. Deschatres, A new web-based system to improve the monitoring of snow avalanche hazard in france, *Natural Hazards and Earth System Sciences* 16 (5) (2016) 1205–1216.
- [33] M. Naaim, Y. Durand, N. Eckert, G. Chambon, Dense avalanche friction coefficients: influence of physical properties of snow, *Journal of Glaciology* 59 (216) (2013) 771–782. doi:10.3189/2013JoG12J205.
- [34] W. Brian Dade, H. E. Huppert, Long-runout rockfalls, *Geological Society of America* 26 (9) (1998) 803–806.
- [35] R. J. Hyndman, H. L. Shang, Rainbow plots, bagplots, and boxplots for functional data, *Journal of Computational and Graphical Statistics* 19 (1) (2010) 29–45.
- [36] R. J. Hyndman, Computing and graphing highest density regions, *The American Statistician* 50 (2) (1996) 120–126.
- [37] E. Thibert, H. Bellot, X. Ravanat, F. Ousset, G. Pulfer, M. Naaim, P. Hagenmuller, F. Naaim-Bouvet, T. Faug, K. Nishimura, Y. Ito, B. Djebbar, A. Prokop, P. Schön, A. Soruco, C. Vincent, A. Limam, R. Héno, The full-scale avalanche test-site at lautaret pass (french alps), *Cold Regions Science and Technology* 115 (2015) 30–41.

- [38] C. P. Borstad, D. McClung, Sensitivity analyses in snow avalanche dynamics modeling and implications when modeling extreme events, *Canadian Geotechnical Journal* 46 (9) (2009) 1024–1033.
- [39] B. Salm, A. Burkard, H. Gubler, Berechnung von Fliesslawinen: eine Anleitung für Praktiker mit Beispielen, Eidgenössisches Institut für Schnee- und Lawinenforschung, Weissfluhjoch/Davos, 1990.
- [40] P. Gauer, Comparison of avalanche front velocity measurements and implications for avalanche models, *Cold Regions Science and Technology* 97 (2014) 132 – 150.
- [41] D. McClung, P. Gauer, Maximum frontal speeds, alpha angles and deposit volumes of flowing snow avalanches, *Cold Regions Science and Technology* 153 (2018) 78–85.
- [42] J. D. Dent, T. E. Lang, Modeling of snow flow, *Journal of Glaciology* 26 (94) (1980) 131–140.
- [43] C. J. Keylock, D. M. McClung, M. M. Magnússon, Avalanche risk mapping by simulation, *Journal of Glaciology* 45 (150) (1999) 303–314.
- [44] N. Eckert, C. Keylock, D. Bertrand, E. Parent, T. Faug, P. Favier, M. Naaim, Quantitative risk and optimal design approaches in the snow avalanche field: Review and extensions, *Cold Regions Science and Technology* 79 (2012) 1–19.
- [45] P. Favier, D. Bertrand, N. Eckert, M. Naaim, A reliability assessment of physical vulnerability of reinforced concrete walls loaded by snow avalanches, *Natural Hazards and Earth System Sciences* 14 (3) (2014) 689–704.

## Appendix A.

In this appendix, we provide the proofs of Propositions 2.1 and 2.2, some descriptive statistics of avalanche simulations for scenario 1 and the results we obtained for the functional principal component analysis for Scenario 2.

### *Appendix A.1. Proofs of Propositions 2.1 and 2.2*

*Proof of Proposition 2.1.* The proof is straightforward. It is a direct consequence of the linearity and additivity properties of the Shapley value. Let  $i \in \{1, \dots, d\}$  and  $j \in \{1, \dots, p\}$ . The characteristic function  $\text{val}_j$  (see (2)) has Shapley value  $\text{Sh}_i^j$ ,  $i \in \{1, \dots, d\}$ . Thanks to the linearity and additivity properties (see properties 4. and 5. of the Shapley value), the characteristic

function  $\mathbf{u} \subseteq \{1, \dots, d\} \mapsto \text{val}(\mathbf{u}) = \frac{\sum_{j=1}^p \text{Var}(Y_j) \text{val}_j(\mathbf{u})}{\sum_{i=1}^p \text{Var}(Y_i)}$  leads to the Shapley value  $\frac{\sum_{j=1}^p \text{Var}(Y_j) \text{Sh}_i^j}{\sum_{i=1}^p \text{Var}(Y_i)}$ . Moreover,

$$\text{val}(\mathbf{u}) = \frac{\sum_{j=1}^p \text{Var}(Y_j) \text{val}_j(\mathbf{u})}{\sum_{j=1}^p \text{Var}(Y_j)} = \frac{\sum_{j=1}^p \text{Var}(\mathbb{E}(Y_j|X_{\mathbf{u}}))}{\sum_{j=1}^p \text{Var}(Y_j)} = \frac{\text{tr}(\Sigma_{\mathbf{u}})}{\text{tr}(\Sigma)} \quad (\text{A.1})$$

where  $\Sigma_{\mathbf{u}}$  is the covariance matrix of  $\mathbb{E}(\mathbf{Y}|X_{\mathbf{u}})$  and  $\Sigma$  is the covariance matrix of  $\mathbf{Y}$ . Thus for any  $\mathbf{u} \subseteq \{1, \dots, d\}$ ,  $\text{val}(\mathbf{u})$  is equal to the aggregated Sobol' index associated to  $\mathbf{u}$ .  $\square$

*Proof of Proposition 2.2. i.* As for all  $j \in \{1, \dots, p\}$   $0 \leq \text{Sh}_i^j \leq 1$  and as the sum of the non negative weights  $\text{Var}(Y_j)/\sum_{\ell=1}^p \text{Var}(Y_{\ell})$  is one, we deduce that  $0 \leq \text{GSh}_i \leq 1$ .

*ii.* Note that  $\text{GSh}'_i$  can be written as  $\text{GSh}'_i = \sum_{j=1}^p \text{Var}(\lambda Y_j) \text{Sh}_i^{\prime j} / \sum_{j=1}^p \text{Var}(\lambda Y_j)$ , where  $\text{Sh}_i^{\prime j}$  is the Shapley effect associated to the characteristic function  $\text{val}'_j$ . Note that  $\text{val}'_j(\mathbf{u}) = \text{Var}(\mathbb{E}(\lambda Y_j|X_{\mathbf{u}}))/\text{Var}(\lambda Y_j) = \text{val}_j(\mathbf{u})$  for all  $\mathbf{u} \subseteq \{1, \dots, d\}$ . Thus,  $\text{Sh}_i^{\prime j} = \text{Sh}_i^j$  from where  $\text{GSh}'_i = \text{GSh}_i$  which means the aggregated Shapley effect is invariant by any nonzero scaling of  $f$ .

*iii.* Let us write  $g(\mathbf{X}) = Of(\mathbf{X}) = O\mathbf{Y} = \mathbf{U}$ . The characteristic function associated to the aggregated Shapley effect  $\text{GSh}'_i$  of  $\mathbf{U}$  is then (see Equation (A.1))  $\text{val}'(\mathbf{u}) = \text{tr}(\Sigma_{\mathbf{u}}^{\mathbf{U}})/\text{tr}(\Sigma^{\mathbf{U}})$  where  $\Sigma_{\mathbf{u}}^{\mathbf{U}}$  is the covariance matrix of  $\mathbb{E}(\mathbf{U}|X_{\mathbf{u}})$  and  $\Sigma^{\mathbf{U}}$  is the covariance matrix of  $\mathbf{U}$ . Then, for all  $\mathbf{u} \subseteq \{1, \dots, d\}$

$$\text{val}'(\mathbf{u}) = \frac{\text{tr}(\Sigma_{\mathbf{u}}^{\mathbf{U}})}{\text{tr}(\Sigma^{\mathbf{U}})} = \frac{\text{tr}(O\Sigma_{\mathbf{u}}^{\mathbf{Y}}O^t)}{\text{tr}(O\Sigma^{\mathbf{Y}}O^t)} = \frac{\text{tr}(\Sigma_{\mathbf{u}}^{\mathbf{Y}})}{\text{tr}(\Sigma^{\mathbf{Y}})} = \text{val}(\mathbf{u}).$$

As  $\text{val}(\mathbf{u})$  has a unique Shapley value  $\text{GSh}_i$ ,  $\text{val}'(\mathbf{u})$  has Shapley value  $\text{GSh}_i$  which proves that  $\text{GSh}'_i = \text{GSh}_i$  for all  $i \in \{1, \dots, d\}$ .  $\square$



Appendix A.2. Avalanche modelling, scenario 1: data description

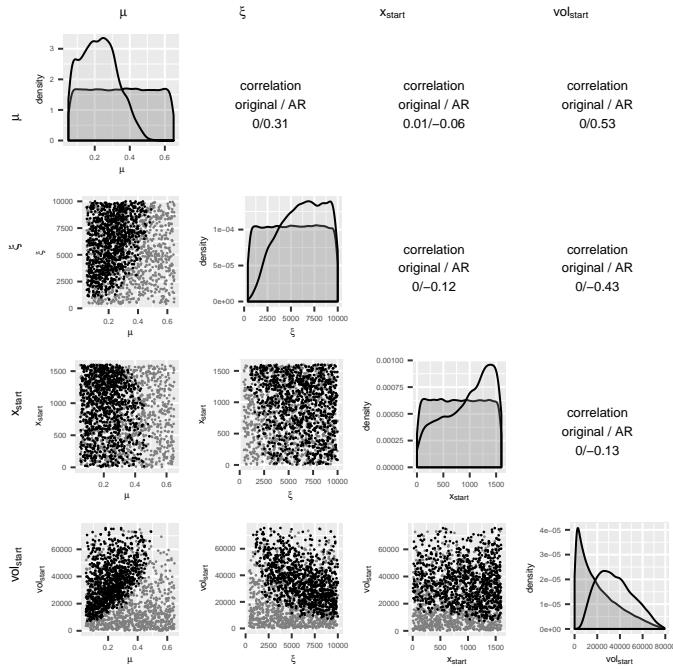


Figure A.7: Avalanche model, scenario 1: scatter-plots of initial (black points) and constrained (gray points) samples. In the figure's diagonal, the density function of the initial (gray color) and constrained (transparent) samples are displayed. Input correlations of the original and constrained (denoted by AR as obtained from an acceptance rejection algorithm) samples are shown. 1000 subsamples of original and constrained samples are used for illustration purpose.

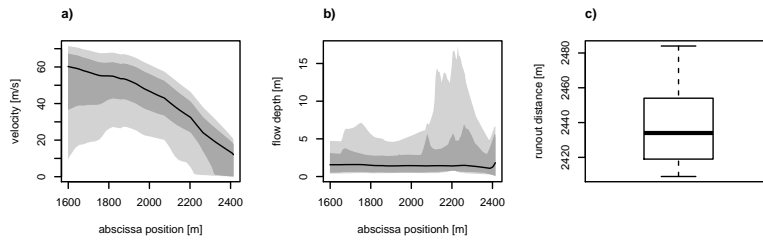


Figure A.8: Avalanche model, scenario 1: a) and b) functional HDR boxplots of velocity and flow depth curves, resp. It is shown 50% HDR (dark gray), 100% HDR (light gray) and modal curve, the curve in the sample with the highest density (black line). c) runout distance boxplot. The constrained sample size is  $n_1 = 6152$ .

Appendix A.3. Functional principal components obtained for Scenario 2

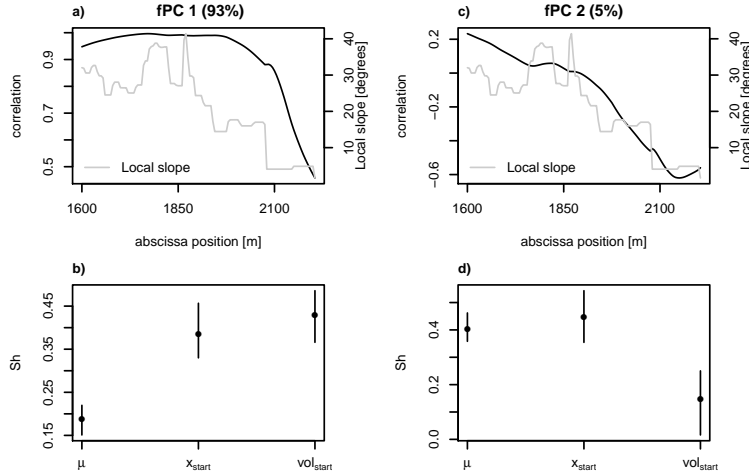


Figure A.9: Avalanche model, scenario 2: a) correlation between the velocity and the first fPC, c) correlation between the velocity and the second fPC. Scalar Shapley effects for b) fPC 1 and d) fPC 2 are estimated with  $n = 1284$ ,  $N_{tot} = 800$ . Confidence intervals are built with  $B = 500$ . The local slope is displayed with a gray line.

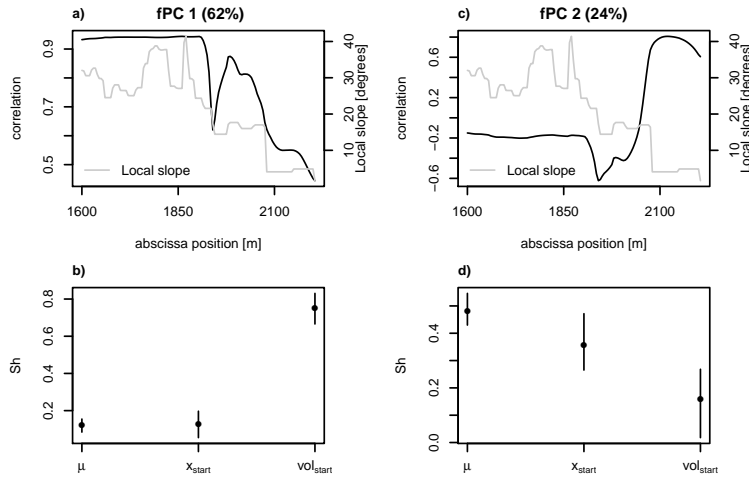


Figure A.10: Avalanche model, scenario 2: a) correlation between the flow depth and the first fPC, c) correlation between the flow depth and the second fPC. Scalar Shapley effects for b) fPC 1 and d) fPC 2 are estimated with  $n = 1284$ ,  $N_{tot} = 800$ . Confidence intervals are built with  $B = 500$ . Local slope is displayed with a gray line.

# Supplementary material

## Global sensitivity analysis with aggregated Shapley effects, application to avalanche hazard assessment

María Belén Heredia<sup>a</sup>, Clémentine Prieur<sup>b</sup>, Nicolas Eckert<sup>a</sup>

<sup>a</sup>Univ. Grenoble Alpes, INRAE, UR ETNA, Grenoble, France

<sup>b</sup>Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, 38000 Grenoble, France

*Keywords:* Global sensitivity analysis, dependent inputs, aggregated Shapley effects, bootstrap confidence intervals, snow avalanche propagation model, snow avalanche hazard assessment

In this supplementary material, we provide numerical experiments conducted on toy models to validate the GSA method we introduced in the paper

### 1. Comparison between two sets of accuracy numbers

In this appendix, we compare on a toy model the uniform allocation of the budget  $N_{tot}$ , namely  $N_u = \text{Round}(N_{tot}/(2^d - 2))$ , with the allocation introduced in [1], defined by  $N_u^* = \text{Round}\left(N_{tot} \binom{d}{|u|}^{-1} (d - 1)^{-1}\right)$ . We consider the

Gaussian linear model example from [2],  $f(\mathbf{X}) = \sum_{i=1}^d \beta_i X_i$ , where the  $X_i$ s are independent centered Gaussian variables with standard deviation  $\sigma_i = i^2$  and,  $\beta_i = 1$  for all  $i \in \{1, \dots, d\}$ . The scalar Shapley effects can be computed analytically [2]:  $\text{Sh}_i = \beta_i^2 \sigma_i^2 / \sum_{i=1}^d \sigma_i^2$ . We used  $N = 100$  independent sets of samples of size  $n = 10\,000$  and  $N_{tot} = 54\,000$  to draw the boxplots shown in Figure 1. Table 1 shows the accuracy to estimate each conditional element  $\mathbb{E}(\text{Var}(Y|\mathbf{X}_{-u}))$  as a function of  $|u|$  in dimension  $d = 10$ . On Figure 1, we see that, at least for this toy example, the allocation proposed in [1] leads to a better accuracy than the uniform allocation. When dimension is low (for example  $d = 4$  as in the mass-spring model), taking  $N_u^*$  or  $N_u$  give similar results because the two options are quite similar.

$ u $	1	2	3	4	5	6	7	8	9
$N_u^*$	600	133	50	29	24	29	50	133	600
$N_u = \text{Round}(N_{tot}/(2^d - 2))$	53	53	53	53	53	53	53	53	53

Table 1: Input dimension  $d = 10$ :  $N_u$  values for both strategies (i) and (ii) for  $n = 10\,000$  (number of simulations) and  $N_{tot} = 54\,000$  (total cost).

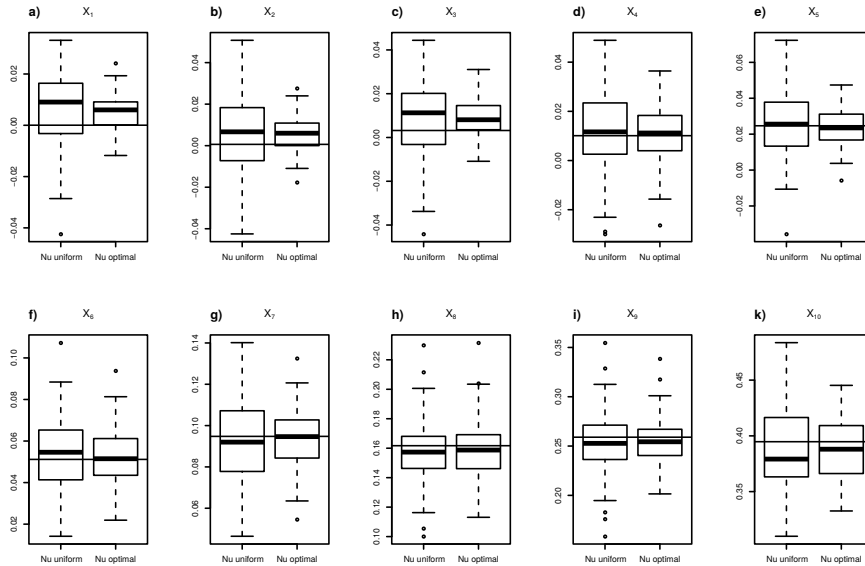


Figure 1: Linear Gaussian model with  $d = 10$  inputs: comparison between  $N_u = \text{Round}(N_{tot}/(2^d - 2))$  for all  $u$  (called  $N_u$  uniform) and  $N_u = N_u^*$  for all  $u$  with  $n = 10000$  and  $N_{tot} = 54000$ . The true scalar Shapley effects are displayed with black lines.

## 2. Estimation accuracy and coverage probability of bootstrap confidence intervals on toy models

In this section, we numerically study the accuracy of the estimation procedure as far as the coverage probability of bootstrap confidence intervals. We consider two toy models: a multivariate linear Gaussian model and the functional mass-spring model proposed in [3]. To estimate scalar Shapley effects, we use the function `shapleySubsetMc` of the R package `sensitivity` [4] corresponding to the estimation procedure defined by Equations (4), (5) and (6) in the main document. Functional PCA is performed using the R package `FPCA` [5].

### 2.1. Multivariate linear Gaussian model in dimension $d = 2$

We consider a multivariate linear model with two Gaussian inputs for which the analytical formulation of the scalar and aggregated Shapley effects is provided in [6].

The model  $f$  is defined as  $\mathbf{Y} = f(\mathbf{X}) = \mathbf{B}^T \mathbf{X}$  with  $\mathbf{X} \sim \mathcal{N}(\mu, \Gamma)$ ,  $\Gamma \in \mathbb{R}^{d \times d}$  a positive-definite matrix and  $\mathbf{B} \in \mathbb{R}^{d \times p}$ . In this example, we consider  $d = 2$  and  $p = 3$  which means  $\mathbf{Y} = (Y_1, Y_2, Y_3)$ . The variance of the centered random variables  $X_1$  and  $X_2$  are equal to  $\sigma_1^2 = 1$  and  $\sigma_2^2 = 3$ , respectively and their correlation  $\rho = 0.4$ . Thus the covariance matrix  $\Gamma$  of  $\mathbf{X}$  and the coefficients of

$B = (\beta_{ij}) \in \mathbb{R}^{2 \times 3}$  are given by:

$$\Gamma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 1 & 0.69 \\ 0.69 & 3 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 4 & 0.1 \\ 1 & 3 & 0.9 \end{bmatrix}.$$

For  $j = 1, 2, 3$ , we then get  $\text{Var}(Y_j) = \sigma_{Y_j}^2 = \beta_{1j}^2\sigma_1^2 + 2\rho\beta_{1j}\beta_{2j}\sigma_1\sigma_2 + \beta_{2j}^2\sigma_2^2$ . The scalar Shapley effects can be computed and we get:

$$\begin{aligned} \sigma_{Y_j}^2\phi_1^j &= \beta_{1j}^2\sigma_1^2 \left(1 - \frac{\rho^2}{2}\right) + \rho\beta_{1j}\beta_{2j}\sigma_1\sigma_2 + \beta_{2j}^2\sigma_2^2\frac{\rho^2}{2}, \\ \sigma_{Y_j}^2\phi_2^j &= \beta_{2j}^2\sigma_2^2 \left(1 - \frac{\rho^2}{2}\right) + \rho\beta_{1j}\beta_{2j}\sigma_1\sigma_2 + \beta_{1j}^2\sigma_1^2\frac{\rho^2}{2}. \end{aligned}$$

Then, the aggregated Shapley effects for  $i \in \{1, 2\}$  are computed according to Equation (3) in the main document. Let us first focus on scalar Shapley effect estimation and the associated confidence intervals, for example scalar Shapley effects for the first component of the output,  $Y_1$ . For  $Y_1$ , the most important input is  $X_2$ , with a Shapley effect around 0.66. In Figure 2, we analyze the evolution of the estimation accuracy through the mean absolute error (MAE) and the probability of coverage (POC) as a function of  $n$  and  $N_{tot}$ . The value of  $N_{tot}$  varies from 100 to  $n(2^d - 2) = 2n$ . For each combination of  $n$  and  $N_{tot}$ ,  $N = 300$  independent sets of random samples are used for estimating the MAE and the POC. The number of bootstrap samples is fixed to  $B = 500$ . The evolution of the 5% and 95% quantiles of the MAE is also drawn. As expected, the estimation accuracy increases with  $n$  and  $N_{tot}$ . Also, for fixed  $n$ , it increases with  $N_{tot}$ . As  $N_{tot}$  reaches  $n$ , which means  $N_{\mathbf{u}} = n/2$  for all  $\mathbf{u}$  as  $d = 2$ , the decrease of the MAE slows down from exponential to linear. This behavior seems coherent with the result stated in [1, Corollary 6.8] which asserts that there exists  $0 < C < +\infty$  such that, for  $N_{\mathbf{u}} \geq Cn^{1/(d-|\mathbf{u}|)}$  and  $\delta > 0$ :

$$|\hat{c}_1(\mathbf{u}) - c_1(\mathbf{u})| = o_p \left( \frac{1}{n^{1/(2(d-|\mathbf{u} |) - \delta)}} \right). \quad (1)$$

The bound in (1), although suffering from the curse of dimensionality, does not depend on  $N_{tot}$  anymore. To the best of our knowledge, solving efficiently the dimensionality issue in the estimation of Shapley effects is still an open challenge. Note that very recently, an algorithm based on random forests which improves the computational cost of Shapley effects was proposed in [7] (see also [8] for a recent review on the use of random forests for GSA). Finally, the behavior of the POC is as expected, around 0.9 whatever  $n$  and  $N_{tot}$ . The evolution of the MAE and POC for the estimation of aggregated Shapley effects as function of  $n$  and  $N_{tot}$  is displayed in Figure 3. Similarly as for the estimation of scalar effects, the accuracy increases with  $n$  and  $N_{tot}$ . However for fixed  $n$ , the POC deteriorates

for large  $N_{tot}$  (POC is around 0.8 for  $N_u = n$ ). The reason for this deterioration is not well understood yet. However, these numerical experiments suggest that it is reasonable to keep all the  $N_u$  strictly below  $n$  to guarantee the POC accuary. We also estimated Shapley effects in the case where input correlation  $\rho$  set equal to 0.9. The conclusions were similar both for the estimation accuracy and for the coverage probability, thus we decided not to present the figures here.

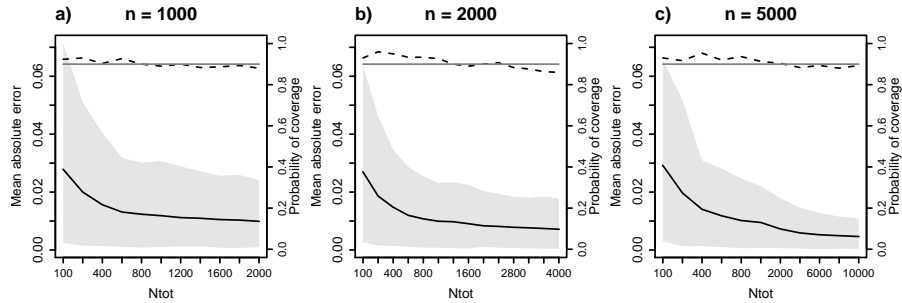


Figure 2: Linear Gaussian model,  $d = 2$ : MAE for scalar Shapley effects of output  $Y_1$  estimated from  $N=300$  sets of independent samples of size a)  $n = 1000$ , b)  $n = 2000$  and c)  $n = 5000$ , as a function of  $N_{tot}$ . The 0.05 and 0.95 pointwise quantiles of the absolute error are drawn with gray polygons. Empirical coverage probability of the 90% bootstrap simultaneous confidence intervals is displayed with dotted lines. Theoretical coverage probability 0.9 is shown with a plain gray line. The number of bootstrap samples is fixed to  $B = 500$ .

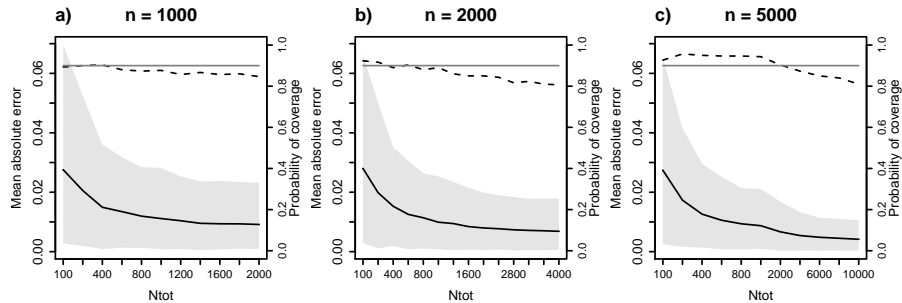


Figure 3: Linear Gaussian model,  $d = 2$ : MAE for aggregated Shapley effects estimated from  $N=300$  sets of independent samples of size a)  $n = 1000$ , b)  $n = 2000$  and c)  $n = 5000$ , as a function of  $N_{tot}$ . The 0.05 and 0.95 pointwise quantiles of the absolute error are drawn with gray polygons. Empirical coverage probability of the 90% bootstrap simultaneous confidence intervals is displayed with dotted lines. Theoretical coverage probability 0.9 is also shown with a gray plain line. The number of bootstrap samples is fixed to  $B = 500$ .

## 2.2. Mass-spring model

In this section, we consider a test case with discretized functional output: the functional mass-spring model considered in [3], where the displacement of a

mass connected to a spring is considered:

$$m\ell''(t) + c\ell'(t) + k\ell(t) = 0, \quad (2)$$

with initial conditions  $\ell(0) = l$ ,  $\ell'(0) = 0$ , and  $t \in [0, 40]$ . There exists an analytical solution to Equation (2). This model has four inputs (see more details in Table 2). The model output is the vector  $\mathbf{Y} = f(\mathbf{X}) = (\ell(t_1), \dots, \ell(t_{800}))$ ,  $t_i = 0.05i$  with  $i \in \{1, \dots, 800\}$ .

Input	Description	Distribution
$m$	mass (kg)	$\mathcal{U}[10, 12]$
$c$	damping constant ( $\text{Nm}^{-1}\text{s}$ )	$\mathcal{U}[0.4, 0.8]$
$k$	spring constant ( $\text{Nm}^{-1}$ )	$\mathcal{U}[70, 90]$
$l$	initial elongation (m)	$\mathcal{U}[-1, -0.25]$

Table 2: Mass-spring model: Input description and uncertainty intervals.  $\mathcal{U}$  denotes the uniform distribution. The inputs are independent from each other.

Inputs are considered independent. The true aggregated Shapley effects are unknown but they are approximated using a high sample size  $n = 25000$  and  $N_{tot} = 10000$ . Then, the Shapley effects estimated are  $\widehat{GS}_m = 0.38$ ,  $\widehat{GS}_c = 0.01$ ,  $\widehat{GS}_k = 0.51$  and,  $\widehat{GS}_l = 0.09$ . Given these results, inputs ranking is:  $k$ ,  $m$ ,  $l$  and  $c$  which corresponds to the same ranking obtained using Sobol' indices (see Table 3 of [3]). The discretized output is high-dimensional ( $p = 800$ ). We perform fPCA (see Section 3.3 in the main document) to estimate the effects using the first  $q \ll p$  fPCs. Figure 4 shows the POC and bias evolution if different values in  $n$  and  $N_{tot}$  are used for the aggregated effect estimation. We use the first 6 fPCs which explain 95% of the output variance (see Figure 4 a). For each  $n$  and  $N_{tot}$  combination, the aggregated Shapley effects are estimated for a set of  $N = 100$  independent random samples and confidence intervals are estimated with a set of  $B = 500$  bootstrap samples. The accuracy is not very good if sample size is small  $n = 1000$  (see Figure 4 b). However, it reduces drastically when the sample size  $n$  increases as expected. On our experiments, the smallest MAE is achieved for  $n = 5000$  and  $N_{tot} = 2002$  (see Figure 4 d). The confidence interval reaches the expected POC 0.9 as soon as  $n$  and  $N_{tot}$  are large enough (see Figure 4 b).

## References

- [1] B. Broto, F. Bachoc, M. Depecker, Variance Reduction for Estimation of Shapley Effects and Adaptation to Unknown Input Distribution, SIAM/ASA Journal on Uncertainty Quantification 8 (2) (2020) 693–716.
- [2] A. B. Owen, C. Prieur, On shapley value for measuring importance of dependent inputs, SIAM/ASA Journal on Uncertainty Quantification 5 (1) (2017) 986–1002.
- [3] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, Sensitivity analysis for multidimensional and functional outputs, Electronic Journal of Statistics 8 (1) (2014) 575–603.

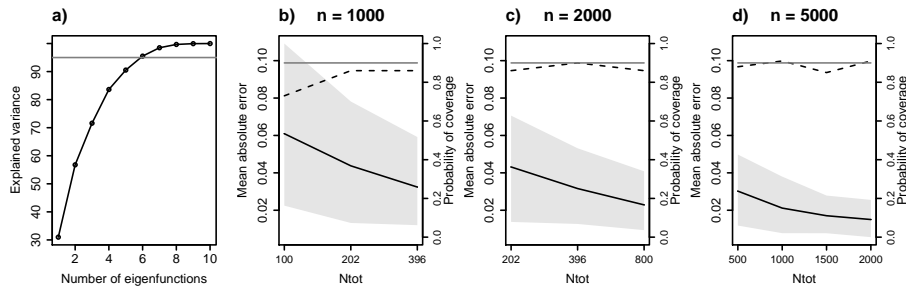


Figure 4: Mass-spring model: a) Explained variance as a function of the number of principal components. The gray line is displayed at 95% of the variance explained which corresponds to 6 components. The MAE computed from a set of  $N = 100$  independent realizations of the aggregated Shapley effect estimates, using the first 6 principal components, is drawn as a function of  $N_{tot}$  with b)  $n = 1000$ , c)  $n = 2000$  and d)  $n = 5000$ . The 0.05 and 0.95 pointwise quantiles of the MAE are drawn with gray polygons. The probability of coverage of the 90% bootstrap simultaneous intervals is displayed with a dotted line. The theoretical POC 0.9 is highlighted with a plain gray line. The number of bootstrap samples is fixed to  $B = 500$ .

- [4] G. Pujol, B. Iooss, A. Janon, sensitivity package, version 1.11, The Comprehensive R Archive Network, 2015.  
URL <http://www.cran.r-project.org/web/packages/sensitivity/>
- [5] Y. Chen, C. Carroll, X. Dai, J. Fan, P. Z. Hadjipantelis, K. Han, H. Ji, H.-G. Mueller, J.-L. Wang, fdapace: Functional Data Analysis and Empirical Dynamics, r package version 0.5.1 (2020).  
URL <https://CRAN.R-project.org/package=fdapace>
- [6] 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.
- [7] C. B enard, G. Biau, S. Da Veiga, E. Scornet, Shaff: Fast and consistent shapley effect estimates via random forests, arXiv preprint arXiv:2105.11724.
- [8] A. Antoniadis, S. Lambert-Lacroix, J.-M. Poggi, Random forests for global sensitivity analysis: A selective review, Reliability Engineering & System Safety 206 (2021) 107312.