Aggregated Shapley effects: nearest-neighbor estimation procedure and confidence intervals. Application to avalanche long term forecasting. *

María Belén Heredia†, Clémentine Prieur‡, and Nicolas Eckert†

Abstract. Dynamic models are simplified representations of some real-world entity that change over time, in equations or computer code. The outputs produced by dynamic models are typically time and/or space dependent and due to physical constraints the parameters that are part of the formulation of such models cannot be considered as independent from each others. Dynamic models provide essential analytical tools with significant applications, e.g., in environmental and social sciences. The outputs produced by dynamic models can be significantly sensitive to variations of parameters entering in their formulation (input parameters), and identifying influential input parameters is one aim of sensitivity analysis. A global sensitivity analysis (GSA) consists in modeling unknown input parameters by a probability distribution which propagates through the model to the outputs. Then, input parameters are ordered according to their contribution on the model outputs by computing sensitivity measures. In this paper, we extend Shapley effects, a sensitivity measure well suited for dependent input parameters, to the framework of dynamic models. We also propose an algorithm to estimate the so-called aggregated Shapley effects and to construct bootstrap confidence intervals for the estimation of scalar and aggregated Shapley effects. We measure the performances of the estimation procedure and the accuracy of the probability of coverage of the bootstrap confidence intervals on toy models. Finally, our procedure is applied to perform a GSA of an avalanche flow dynamic model, for which the input/output sample we have was obtained from an acceptance-rejection algorithm. More precisely, we analyze the sensitivity in two different settings. In the first setting, we consider that we have little knowledge on the input parameter probability distribution. The second setting focuses on an avalanche corridor already documented by anterior avalanche risk studies.

Key words. Global sensitivity analysis, dependent inputs, aggregated Shapley effects, bootstrap confidence intervals, avalanche flow dynamic model

AMS subject classifications. 62J10, 62F40, 62P12

1. Introduction.

Dynamic models are simplified representations of some real-world entity that change over time, in equations or computer code. These models are useful for the analysis of real-world phenomena, e.g., in environmental or social sciences [32]. For a better understanding of a phenomenon or for long term forecasting, it might be important to identify input parameters entering in the formulation of such dynamic models, particularly the ones which are influential on the outputs of interest. Determining these influential parameters is one aim of global sensitivity analysis (GSA). A global sensitivity analysis (GSA) consists in modeling unknown input parameters by a probability distribution which propagates through the model to the outputs. Then, input parameters are ordered according to their contribution on the model outputs by computing sensitivity measures. In the literature, there exists different global sensitivity measures, e.g., variance based measures such as Sobol' indices [56, 46],

*Submitted to the editors DATE.

Funding: This work was funded by OSUG@2020 and the French National Research Agency.

†Univ. Grenoble Alpes, INRAE, UR ETNA, Grenoble, France (maria-belen.heredia@inrae.fr).

‡Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, 38000 Grenoble, France.

density based measures [6, 7, 60], entropy measures [4], etc. A review of global sensitivity measures can be found in, e.g., [8] or [30].

Due to modeling constraints inherent to many applications, model input parameters might be dependent. It happens indeed that input parameters are interrelated by physical constraints, as for example it is the case for the model presented in [52] modeling the response of a nuclear reactor. In [40], the input parameters of a natural gas transmission model are sampled from an acceptance-rejection algorithm thus can not be considered as independent (see also [35]). A particularity of dynamic models considered in this paper is that the output they produce are typically time and/or space dependent (see e.g., [1, 38]). More specifically, the application that motivated our study is an avalanche flow dynamic model which produces three outputs: the functional flow velocity and depth and the scalar runout distance, which corresponds to the distance traveled by the avalanche. Samples are obtained from an acceptation-rejection algorithm thus (i) input parameters are dependent, (ii) input parameters are not necessarily confined in a rectangular region and (iii) input parameters have unknown probability distribution. For these reasons, we develop a GSA which can handle complex input parameter probability distribution and functional outputs (or multivariate outputs if we discretize functional ones).

Although the independence assumption on input parameters is unrealistic in many applications, it is traditionally required to interpret or to compute sensitivity measures. In other words, if input parameters are dependent, some sensitivity measures are difficult to interpret. E.g., if input parameters are dependent, the functional ANOVA decomposition used for the interpretation of Sobol' indices is not unique and Sobol' indices can actually sum to greater than one. Some authors have proposed strategies to estimate variance based sensitivity measures if input parameters are dependent (cite, e.g., [62, 39, 11, 41, 36, 42, 64, 61, 27]). However, these papers do not provide an univocal way of partitioning the influence of input parameters on the output. In [33], grouped Sobol' indices are introduced. Grouped Sobol' indices can be defined if the input parameters can be splitted in independent groups of dependent parameters, then a Sobol' index is attributed to each group, but not to each input parameter. Other authors have proposed alternative sensitivity measures such as moment independent sensitivity measures (see, e.g., [6]) or have adapted existing procedures to the framework of dependent input parameters (see, e.g., the screening procedure presented in [26]). A more complete review of this literature can be found in [31].

The Shapley effects are a variance based sensitivity measure proposed by [46], which are still meaningful in the framework of dependent input parameters [47]. This measure is based on the Shapley value which is a cooperative game theory concept. Briefly speaking, Shapley value ensures a fair distribution of a gain among team players according to their individual contributions. As a sensitivity measure, [46] adapted the Shapley value into the Shapley effects by considering model input parameters as players and the gain function as the output variance. 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 [9, 31].

Regarding the estimation of the Shapley effects, [58], [9] and [50] proposed estimation algorithms. [58] proposed two estimators for Shapley effects. [5] proposed bootstrap confidence intervals for [58] estimators. [50] proposed an estimation algorithm based on the Möbious inverse to reduce estimation computational cost. In fact, it is well known that Shapley effects estimation is costly. In the algorithm proposed in [58], it is assumed that it is possible to sample from the distribution of a subset of the input parameters conditionally to the complementary set of input parameters. In [9], the authors proposed given data estimators based on nearest-neighbor, which can be computed from a i.i.d. sample of input parameters, which is in general more convenient for real applications. It is worth to mention that give data estimators of Sobol' indices have also been proposed in the literature: we can cite the EASI spectral method of [48], [49] which relies on the notion of class-conditional densities, the nonparametric estimation methods of [13] or [57], the fully Bayesian given data procedure proposed by [3], and more recently in [23] estimators based on rank statistics. But even if Sobol' indices estimation is available when input parameters are dependent, their interpretation is still difficult. Shapley effects have been studied in other works, e.g., [31] analyzed the effect of linear correlation between Gaussian inputs on the Shapley effects. Shapley effects have been also used in real application e.g., in a nuclear application where inputs are correlated [52], and in the multi-physic coupling modeling of a rod ejection accident in a pressurized water reaction [14]. Finally, [51] extended Shapley effects to also provide information about input interactions.

In this work, we extend Shapley effects to multivariate or functional outputs in the framework of dependent input parameters. When outputs are multivariate or functional, it is possible to compute a sensitivity Shapley effect for each component of the output, however this approach leads to results that are difficult to interpret [1] or particularly redundant if we consider the case of discretized functional outputs [37]. [37] and [25] extended Sobol' indices to multivariate or functional outputs. [1] extended Sobol' indices to time-dependent outputs. Following these papers, we introduce and study the properties of what we call aggregated Shapley effects. If the output dimension is high (as it is the case, e.g., when considering the discretization of a functional output), a dimension reduction can be applied as a preliminary step to estimate efficiently aggregated Shapley effects. We use the Karhunen-Love (KL) expansion as in [37, 1]. More precisely to perform KL expansion, we use the functional principal component analysis proposed by [63]. The extension of Shapley effects to multivariate outputs has been early studied in [14], but here we analyze more deeply its definition, properties and estimation. We also provide a bootstrap algorithm to estimate confidence intervals for scalar and aggregated Shapley effects motivated by [5].

Our method is motivated by the study of an avalanche flow dynamic model which depends on some poorly known inputs [17]. This model is employed for elaborating land-use maps or for designing defense structures [44, 22]. Many of the input parameters entering in the formulation of the model are uncertain. Understanding the influence of these parameters on the model outputs is important for the a better comprehension of avalanche phenomenon, but also for determining the most influential parameter on which effort should be concentrated to provide more accurate long term forecasting. In our application, the input/output sample is obtained from an acceptance-rejection algorithm. We analyze the sensitivity in two different settings. In the first setting, we consider that we have little knowledge on the input parameter probability distribution. The second setting focuses on an avalanche corridor already documented by anterior avalanche risk studies [15].

In summary, the main contributions of this work are: (i) to extend Shapley effects to models with multivariate or functional outputs, (ii) to provide an algorithm to construct bootstrap confidence intervals for scalar and aggregated Shapley effect estimation (iii) and, to apply our GSA procedure to a complex avalanche application where samples are obtained from an acceptance-rejection algorithm. The paper is organized as follows. In Section 2, aggregated Shapley effects and their main properties are described. In Section 3, we propose an estimator for aggregated Shapley effects in a given data framework by extending the Monte-Carlo nearest-neighbor estimator of scalar Shapley effects introduced in [9]. At the end of the section, we describe the functional principal components analysis algorithm to perform model dimension reduction proposed by [63]. In Section 4, we propose a bootstrap algorithm to construct confidence intervals of the scalar and aggregated Shapley effect estimations based on [5]. In Section 5, we test our estimation procedure on two toy models: a multivariate linear Gaussian model and the mass-spring model. Finally in Section 6, our GSA procedure is applied to an avalanche model. We discuss our conclusions and perspectives in Section 7.

2. Aggregated Shapley effects.

Shapley effects are sensitivity measures to quantify input importance proposed by [46]. These measures are particularly useful when inputs are dependent. Shapley effects are based in the concept of Shapley value, introduced in the framework of game theory [55], which consists into dividing a game gain among a group of players in an equitable way. As sensitivity measures, Shapley effects consider model inputs as players and output variance as game function. Shapley effects can be naturally extended to multivariate output models by following the ideas presented in [24] and [37] to generalize Sobol' indices to multivariate output models (see also [1] for time-dependent models). We call these new sensitivity measures aggregated Shapley effects.

2.1. Definition.

Let us define $\mathbf{Y} = (Y_1, \ldots, Y_j, \ldots, Y_p) = f(\mathbf{X})$ the p multivariate output of a model f that depends on d random inputs $\mathbf{X} = (X_1, \ldots, X_d)$. The inputs are defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P}_{\mathbf{X}})$ and $f \in \mathbb{L}^2(\mathbb{P}_{\mathbf{X}})$. For any $\mathfrak{u} \subseteq \{1, \ldots, d\}$, let us define $-\mathfrak{u} = \{1, \ldots, d\} \setminus \mathfrak{u}$ its complement. We set $\mathbf{X}_{\mathfrak{u}} = (X_i)_{i \in \mathfrak{u}}$. Note that the inputs are not necessary independent.

In the framework of our application to avalanche long term forecasting, the model produces outputs of the form $\mathbf{Y} = (Y_1 = f(s_1, \mathbf{X}), \ldots, Y_p = f(s_p, \mathbf{X}))$, with $s_1, \ldots, s_p \in \mathbb{R}$ the p discretization points along the avalanche corridor.

In this section we recall the definition and main properties of the Shapley value, on which the definition of Shapley effects is based. Given a set of d players in a coalitional game and a charateristic function val : $2^d \to \mathbb{R}$, $\text{val}(\emptyset) = 0$, the Shapley value $(\phi_1, \ldots, \phi_d)$ is the only distribution of the total gains $\text{val}(\{1, \ldots, d\})$ to the players satisfying the desirable properties listed below:

1. (Efficiency) $\sum_{i=1}^d \phi_i = \text{val}(\{1, \ldots, d\})$.
2. (Symmetry) If $\text{val}(\mathfrak{u} \cup \{i\}) = \text{val}(\mathfrak{u} \cup \{\ell\})$ for all $\mathfrak{u} \subseteq \{1, \ldots, d\} - \{i, j\}$, then $\phi_i = \phi_\ell$.
3. (Dummy) If $\text{val}(\mathfrak{u} \cup \{i\}) = \text{val}(\mathfrak{u})$ for all $\mathfrak{u} \subseteq \{1, \ldots, d\}$, then $\phi_i = 0$.
4. (Additivity) If val and val' have Shapley values ϕ and ϕ' respectively, then the game with characteristic function val + val' has Shapley value $\phi_i + \phi'_i$ for $i \in \{1, \ldots, d\}$.

It is proved in [55] that according to the Shapley value, the amount that player i gets given a coalitional game (val, d) is:

$$\phi_i = \frac{1}{d} \sum_{\mathfrak{u} \subseteq -\{i\}} \binom{d-1}{|\mathfrak{u}|}^{-1} (\mathrm{val}(\mathfrak{u} \cup \{i\}) - \mathrm{val}(\mathfrak{u})) \;\; \forall i \in \{1, \ldots, d\}. \tag{2.1}$$

The Shapley value also satisfies the linearity property:

5. (Linearity) Let $\lambda \in \mathbb{R}$, if $\lambda\mathrm{val}$ and val have Shapley values ϕ' and ϕ, then $\phi'_i = \lambda\phi_i$ for all $i \in \{1, \ldots, d\}$.

The linearity property is used to prove some of the nice properties of aggregated Shapley effects (see Propositions 2.1 and 2.2 further).

The Shapley effects are defined by considering the characteristic function of the game as:

$$\mathrm{val}_j(\mathfrak{u}) = \frac{\mathrm{Var}\left(\mathbb{E}(Y_j|\mathbf{X}_\mathfrak{u})\right)}{\mathrm{Var}(Y_j)}, \; \mathfrak{u} \subseteq \{1, \ldots, d\} \tag{2.2}$$

in Equation (2.1). Thus, the scalar Shapley effect of input i in output j is defined as:

$$Sh_i^j = \frac{1}{d\,\mathrm{Var}(Y_j)} \sum_{\mathfrak{u} \subseteq -\{i\}} \binom{d-1}{|\mathfrak{u}|}^{-1} \left(\mathrm{Var}\left(\mathbb{E}(Y_j|\mathbf{X}_{\mathfrak{u}\cup i})\right) - \mathrm{Var}\left(\mathbb{E}(Y_j|\mathbf{X}_\mathfrak{u})\right)\right). \tag{2.3}$$

Shapley effects can be naturally extended to models with multivariate outputs following ideas from [24] and [37] where authors proposed to extend Sobol' indices to multivariate outputs. Aggregated Shapley effect of an input i is then defined as:

$$GSh_i = \frac{\sum_{j=1}^p \mathrm{Var}(Y_j) Sh_i^j}{\sum_{j=1}^p \mathrm{Var}(Y_j)}, \tag{2.4}$$

where Sh_i^j is the scalar Shapley effect of input X_i in output Y_j. This sensitivity measure is a weighted sum of the scalar Shapley effects where weights correspond to the proportion of the variance of each output over the sum of all individual variances.

2.2. Properties. In this section, we prove some nice properties of aggregated Shapley effects.

Proposition 2.1. *The aggregated Shapley effects GSh_i, $i \in \{1, \ldots, d\}$, correspond to the Shapley value with characteristic function defined as:*

$$val(i) = \frac{\sum_{j=1}^p Var(Y_j) val_j(i)}{\sum_{j=1}^p Var(Y_j)}. \tag{2.5}$$

Proof. The proof is straightforward. It is a direct consequence of the linearity and additivity properties of the Shapley value. Let $i \in \{1, \ldots, d\}$ and $j \in \{1, \ldots, p\}$. The characteristic function val_j (see Equation 2.2) has Shapley value 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 $\frac{\sum_{j=1}^p \mathrm{Var}(Y_j)\mathrm{val}_j(i)}{\sum_{i=1}^p \mathrm{Var}(Y_j)}$ leads to the Shapley value $\frac{\sum_{j=1}^p \mathrm{Var}(Y_j) Sh_i^j}{\sum_{i=1}^p \mathrm{Var}(Y_j)}$. ■

The characteristic function (2.5) can be written in matricial form:

$$\mathrm{val}(i) = \frac{\sum_{j=1}^p \mathrm{Var}(Y_j)\mathrm{val}_j(i)}{\sum_{i=1}^p \mathrm{Var}(Y_j)} = \frac{\sum_{j=1}^p \mathrm{Var}(\mathbb{E}(Y_j|X_i))}{\sum_{i=1}^p \mathrm{Var}(Y_j)} = \frac{tr(\Sigma_i)}{tr(\Sigma)} \tag{2.6}$$

where Σ_i is the covariance matrix of $\mathbb{E}(\mathbf{Y}|X_i)$ and Σ is the covariance matrix of $\mathbf{Y}$. Note that the characteristic function val of aggregated Shapley effects corresponds to the definition of the aggregated Sobol' indices introduced in [37, 24]. In the next proposition, we prove that aggregated Shapley effects accomplish the natural requirements for a sensitivity measure mentioned in Proposition 3.1 in [24].

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

i. $0 \le GSh_i \le 1$.

ii. *GSh_i is invariant by left-composition by any nonzero scaling of f, which means, for any $\lambda \in \mathbb{R}$, the aggregated Shapley effect GSh'_i of $\lambda f(\boldsymbol{X})$ is GSh_i.*

iii. *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 GSh'_i of $Of(\boldsymbol{X})$ is GSh_i for all $i \in \{1, \dots, d\}$.*

Proof. *i.* As for all $j \in \{1, \dots, p\}$ $0 \le Sh_i^j \le 1$ and as the sum of the non negative weights $\mathrm{Var}(Y_j)/\sum_{\ell=1}^p \mathrm{Var}(Y_\ell)$ is one, we deduce that $0 \le GSh_i \le 1$. *ii.* Note that GSh'_i can be written as $GSh'_i = \frac{\sum_{j=1}^p \mathrm{Var}(\lambda Y_j) Sh_i'^j}{\sum_{j=1}^p \mathrm{Var}(\lambda Y_j)}$, where $Sh_i'^j$ is the Shapley effect associated to the characteristic function val'_j. Notice that $\mathrm{val}'_j(i) = \frac{\mathrm{Var}(\mathbb{E}(\lambda Y_j|X_i))}{\mathrm{Var}(\lambda Y_j)} = \mathrm{val}_j(i)$. Thus, $Sh_i'^j = Sh_i^j$ from where $GSh'_i = GSh_i$ 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 GSh'_i of $\mathbf{U}$ is then (see Equation (2.6)) $\mathrm{val'}(i) = tr(\Sigma_i^{\mathbf{U}})/tr(\Sigma^{\mathbf{U}})$ where $\Sigma_i^{\mathbf{U}}$ is the covariance matrix of $\mathbb{E}(\mathbf{U}|X_i)$ and $\Sigma^{\mathbf{U}}$ is the covariance matrix of $\mathbf{U}$. Then,

$$\mathrm{val'}(i) = \frac{tr(\Sigma_i^{\mathbf{U}})}{tr(\Sigma^{\mathbf{U}})} = \frac{tr(O\Sigma_i^{\mathbf{Y}}O^t)}{tr(O\Sigma^{\mathbf{Y}}O^t)} = \frac{tr(\Sigma_i^{\mathbf{Y}})}{tr(\Sigma^{\mathbf{Y}})} = \mathrm{val}(i).$$

As $\mathrm{val}(i)$ has an unique Shapley value GSh_i, $\mathrm{val}'(i)$ has Shapley value GSh_i which proves that $GSh'_i = GSh_i$ for all $i \in \{1, \dots, d\}$. ■

In this section, we have proven that aggregated Shapley effects are sensitivity measures. In the next section, we describe the estimation procedure we propose for aggregated Shapley effects, based the estimation procedure of scalar Shapley effects proposed in [9, Section 6] when observing an i.i.d. sample of $(\mathbf{X}, \mathbf{Y})$. Such a procedure, which does not require a specific form for the design of experiments is also called given data procedure.

3. Estimation procedure for scalar and aggregated Shapley effects.

The aggregated Shapley effect estimation procedure we propose in this section is based on the given data estimation procedure of the scalar Shapley effects introduced in [9, Section 6.1.1.]. In the application we consider in Section 6, samples are constructed using acceptance-rejection rules. Therefore the standard pick-freeze estimation procedure (see, e.g., [34]) can not be used as it is based on a specific pick-freeze type design of experiments. It is the reason why we turn to the given data estimation procedure of scalar Shapley effects introduced in [9, Section 6.1.1.]. For sake of clarity, we first present the estimation procedure for scalar Shapley effects in Subsection 3.1 before extending it to the estimation of aggregated Shapley effects in Subsection 3.2.

3.1. Double Monte Carlo given data estimation of scalar Shapley effects.

As noticed in [58, Theorem 1], replacing the characteristic function $\tilde{c}_j(\mathfrak{u}) = \mathrm{Var}(\mathbb{E}(Y_j|\mathbf{X}_\mathfrak{u}))$ by the characteristic function $c_j(\mathfrak{u}) = \mathbb{E}(\mathrm{Var}(Y_j|\mathbf{X}_{-\mathfrak{u}}))$ with $\mathfrak{u} \subseteq \{1,\dots,d\}$ in Equation (2.3) does not change the definition of Shapley effects. Moreover, as pointed in [58] (based on the work in [59]), the double Monte Carlo estimator of $\tilde{c}_j(\mathfrak{u})$ can suffer from a non neglectable bias if the inner loop sample is small, while in contrast the double Monte Carlo estimator of $c_j(\mathfrak{u})$ is unbiased for any sample size. For that reason, we turn to the double Monte Carlo estimator of $c_j(\mathfrak{u})$. To estimate the scalar Shapley effects from the estimates of $c_j(\mathfrak{u})$, $\mathfrak{u} \subseteq \{1,\dots,d\}$, the two aggregation procedures are discussed in [9, Section 4], the random permutation aggregation procedure, and the subset aggregation procedure. We focus in this work on the subset aggregation procedure as it allows a variance reduction. Note that $c_j(\emptyset) = 0$ and that $c_j(\{1,\dots,d\}) = \mathrm{Var}(Y_j)$, which is assumed to be known in [9], and that is estimated by the empirical variance in the present paper. As already mentioned, we consider the given data version for the subset aggregation procedure with double Monte Carlo introduced in [9, Section 6.1.1.] for the estimation of scalar Shapley effects. More precisely, given a n sample $(\mathbf{X}^{(i)}, \mathbf{Y}^{(i)})$, $1 \le i \le n$ of $(\mathbf{X}, \mathbf{Y})$, we define:

$$\widehat{c_j}(\mathfrak{u}) = \frac{1}{N_\mathfrak{u}} \sum_{\ell=1}^{N_\mathfrak{u}} \widehat{E}^j_{\mathfrak{u},s_\ell} \quad \text{with} \tag{3.1}$$

$$\widehat{E}^j_{\mathfrak{u},s_\ell} = \frac{1}{N_I - 1} \sum_{i=1}^{N_I} \left(f_j\left(\mathbf{X}^{(k_n^{-\mathfrak{u}}(s_\ell,i))}\right) - \frac{1}{N_I} \sum_{h=1}^{N_I} f_j\left(\mathbf{X}^{(k_n^{-\mathfrak{u}}(s_\ell,h))}\right)\right)^2 \tag{3.2}$$

with the notation $f_j(\mathbf{X}) = Y_j$. For $\emptyset \subsetneq \mathfrak{v} \subsetneq \{1,\dots,d\}$, the index $k_n^{\mathfrak{v}}(l,m)$ denotes as in [9, Section 6] the index such that $\mathbf{X}_\mathfrak{v}^{k_n^{\mathfrak{v}}(l,m)}$ is the (or one of the) m-th closest element to $\mathbf{X}_\mathfrak{v}^{(l)}$ in $(\mathbf{X}_\mathfrak{v}^{(i)})_{1\le i\le n}$ and such that $(k_n^{\mathfrak{v}}(l,m))_{1\le m\le N_I}$ are two by two distinct and $(s_\ell)_{1\le \ell\le N_\mathfrak{u}}$ is a sample of uniformly distributed integers without replacement in $\{1,\dots,n\}$. N_I and N_u are respectively the Monte-Carlo sample sizes for the conditional variance and expectation. The choice of these two parameters is discussed further. In [9, Theorem 6.6.], it is proved that under theoretical assumptions, $\widehat{c_j}(\mathfrak{u})$ converges in probability to $c_j(\mathfrak{u})$ when n and $N_\mathfrak{u}$ go to ∞.

The algorithm that consists in estimating scalar Shapley effects by plugging (3.1) in Equation (2.3) is called subset aggregation procedure as:

$$\widehat{Sh}_i^j = \frac{1}{d\,\hat{\sigma}_j^2} \sum_{\mathfrak{u} \subseteq -i} \binom{d-1}{|\mathfrak{u}|}^{-1} \left(\widehat{c_j}(\mathfrak{u} \cup \{i\}) - \widehat{c_j}(\mathfrak{u})\right) \tag{3.3}$$

where $\hat{\sigma}_j^2$ is the empirical estimator of $\mathrm{Var}(Y_j)$. Note that, in the subset aggregation procedure, $N_{\mathfrak{u}}$ depends on each $\emptyset \subsetneq \mathfrak{u} \subsetneq \{1,\ldots,d\}$.

Finally, we discuss the choice of N_I and $N_{\mathfrak{u}}$ for all $\emptyset \subsetneq \mathfrak{u} \subsetneq \{1,\ldots,d\}$. We set as in [9] $N_I = 3$ and we choose $N_{\mathfrak{u}}$ according to the rule proposed in [9, Proposition 4.2.] which aims at minimizing $\sum_{i=1}^d \mathrm{Var}(\widehat{Sh}_i^j)$ for a fixed total cost $\kappa \sum_{\emptyset \subsetneq \mathfrak{u} \subsetneq \{1,\ldots,d\}} N_{\mathfrak{u}} = N_{tot}$ fixed by the user. Note that the optimal values $N_{\mathfrak{u}}^* = \left\lfloor N_{tot} \binom{d}{|\mathfrak{u}|}^{-1} (d-1)^{-1} \right\rfloor$, $\emptyset \subsetneq \mathfrak{u} \subsetneq \{1,\ldots,d\}$, do not depend on $1 \le j \le p$. The optimal values $N_{\mathfrak{u}}^*$ are computed under theoretical assumptions that are not satisfied for the given data version of the estimators. However, numerical experiments in [9] show that this choice performs well in practice. Note that the estimator cost in terms of number of model evaluations is n while the cost in terms of nearest-neighbors search is N_{tot}. In [9, Proposition 6.12.], it is proved that under theoretical assumptions the scalar Shapley effect estimators $\widehat{Sh}_i^j$ converge to the scalar Shapley effects in probability when n and N_{tot} go to ∞. Once more, although theoretical assumptions for the convergence are not guaranteed in the applications, numerical performance of the estimators have been demonstrated in [9].

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 the aggregated Shapley effects by:

$$\widehat{GSh}_i = \frac{\sum_{j=1}^p \hat{\sigma}_j^2 \widehat{Sh}_i^j}{\sum_{j=1}^p \hat{\sigma}_j^2} = \frac{1}{d \sum_{j=1}^p \hat{\sigma}_j^2} \sum_{j=1}^p \sum_{\mathfrak{u} \subseteq -i} \binom{d-1}{|\mathfrak{u}|}^{-1} \left(\widehat{c_j}(\mathfrak{u} \cup \{i\}) - \widehat{c_j}(\mathfrak{u})\right), \tag{3.4}$$

with $\hat{\sigma}_j^2$ the empirical estimator of $\mathrm{Var}(Y_j)$ and with $\widehat{c_j}(\mathfrak{u})$ defined by (3.1).

3.3. Dimension reduction: functional principal component analysis. If model f is space or time-dependent, inspired by [1] and [37], we perform a Karhunen-Loève (KL) expansion to obtain a low-rank model representation. In fact, aggregated Shapley effects might be computed more effectively in a low-rank representation. To perform KL expansion, we use the principal component analysis through conditional expectation (PACE) method proposed by [63] (see also [2] for an illustration of its application). More precisely, we have a collection of n independent trajectories of a smooth random function $f(., \mathbf{X})$ with unknown mean $\mu(s) = \mathbb{E}(f(s, \mathbf{X}))$, $s \in \tau$, where τ is a bounded and closed interval in $\mathbb{R}$, and covariance function $G(s_1, s_2) = \mathrm{Cov}(f(s_1, \mathbf{X}), f(s_2, \mathbf{X}))$, $s_1, s_2 \in \tau$. We assume that G has a L^2 orthogonal expansion in terms of eigenfunction ξ_k and non increasing eigenvalues λ_k such that:

$$G(s_1, s_2) = \sum_{k \ge 1} \lambda_k \xi_k(s_1, \mathbf{X}) \xi_k(s_2, \mathbf{X}),\, s_1, s_2 \in \tau.$$

The KL orthogonal expansion of $f(s, \mathbf{X})$ is:

$$f(s,\mathbf{X}) = \mu(s) + \sum_{k\geq 1} \alpha_k(\mathbf{X})\xi_k(s) \approx \mu(s) + \sum_{k=1}^{q} \alpha_k(\mathbf{X})\xi_k(s), s \in \tau, \tag{3.5}$$

where $\alpha_k(\mathbf{X}) = \int_\tau f(s,\mathbf{X})\xi_k(s)ds$ is the k-th functional principal component (fPC) and q is a truncation level. For fPCs estimation, the authors in [63] propose first to estimate $\hat{\mu}(s)$ using local linear smoothers and to estimate $\widehat{G}(s_1, s_2)$ using local linear surface smoothers ([21]). The estimates of eigenfunctions and eigenvalues correspond then to the solutions of the following integral equations:

$$\int_\tau \widehat{G}(s_1, s)\widehat{\xi}_k(s_1)ds_1 = \widehat{\lambda}_k\, \widehat{\xi}_k(s), s \in \tau,$$

with $\int_\tau \widehat{\xi}(s)ds = 1$ and $\int_\tau \widehat{\xi}_k(s)\widehat{\xi}_m(s) = 0$ for all $m \neq k \leq q$. The problem is solved by using a discretization of the smoothed covariance (see further details in [53] and [10]). Finally, fPCs $\hat{\alpha}_k(\mathbf{X}) = \int_\tau f(s,\mathbf{X})\widehat{\xi}_k(s)ds$ are solved by numerical integration.

Aggregated Shapley effects are approximated using the low rank KL model representation with truncation level q, in other words, they are computed with only the q first fPCs:

$$\widetilde{GSh}_i = \frac{1}{d\sum_{k=1}^{q}\lambda_k} \sum_{k=1}^{q} \sum_{\mathfrak{u}\subseteq -i} \binom{d-1}{|\mathfrak{u}|}^{-1} \left(\mathbb{E}(\mathrm{Var}(\alpha_k(\mathbf{X})|\mathbf{X}_{\mathfrak{u}\cup\{i\}})) - \mathbb{E}(\mathrm{Var}(\alpha_k(\mathbf{X})|\mathbf{X}_{\mathfrak{u}}))\right). \tag{3.6}$$

Remark 3.1. (3.6) can be estimated as (3.4).

In unreported numerical test cases, we noticed that using the same sample to perform fPCA and to estimate the Shapley effects provides better results than splitting the sample in two parts.

4. Bootstrap confidence intervals with percentile bias correction.

Confidence intervals are a valuable tool to quantify uncertainty in estimation. We consider non parametric bootstrap confidence intervals with bias percentile correction (see, e.g., [19, 20]). More precisely, we propose to construct confidence intervals, with a block bootstrap procedure, following ideas in [5]. Indeed, bootstrap by blocks is necessary to preserve the nearest-neighbor structure in Equation (3.2) and to avoid potential equalities in distance (see Assumption 6.3 in [9]). We describe in Algorithm 4.1 how to create B bootstrap samples for scalar Shapley effect estimators $\widehat{Sh}_i^j$ and aggregated Shapley effect estimators $\widehat{GSh}_i$, and then we describe the percentile bias correction method.

If model output is scalar, only Steps 1 to 3 of Algorithm 4.1 should be used. The block bootstrap procedure is described by Steps 3.1 to 3.3. Also, the same sample $(\mathbf{x}, \mathbf{y})$ is used to estimate the variance of the outputs Y_j, $1 \leq j \leq p$, and the Shapley effects. In unreported numerical experiments, we noticed once more that using one sample gives better results than

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

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 $\emptyset \subsetneq \mathfrak{u} \subsetneq \{1,\dots,d\}$, a $N_\mathfrak{u}$ random sample $(s_\ell)_{1\le\ell\le N_\mathfrak{u}}$ from $\{1,\dots,n\}$.

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

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 $1 \le j \le p$, $\widehat{\sigma}_j^{2,(b)}$ the empirical variance of $\mathbf{y}_j^{(b)}$.
3. For each $j \in \{1,\dots,p\}$:
   3.1. For all $\mathfrak{u}$ and for all $(s_\ell)_{1\le\ell\le N_\mathfrak{u}}$ compute $\widehat{E}^j_{\mathfrak{u},s_\ell}$ using (3.2).
   3.2. For all $\mathfrak{u}$, create a $N_\mathfrak{u}$ bootstrap sample $\widehat{E}^{j,(b)}_{\mathfrak{u},s_\ell}$ by sampling with replacement from $\left(\widehat{E}^j_{\mathfrak{u},s_\ell}\right)_{1\le\ell\le N_\mathfrak{u}}$ computed in Step 3.1.
   3.3. Compute $\widehat{c_j}^{(b)}(\mathfrak{u}) = \frac{1}{N_\mathfrak{u}} \sum_{\ell=1}^{N_\mathfrak{u}} \widehat{E}^{j,(b)}_{\mathfrak{u},s_\ell}$ for all $\mathfrak{u}$ using (3.1).
   3.4. Compute the b bootstrap sample of $\widehat{Sh}_i^j$ according to (3.3):
$$\widehat{Sh}_i^{j,(b)} = \frac{1}{d\,\widehat{\sigma}_j^{2,(b)}} \sum_{\mathfrak{u}\subseteq -i} \binom{d-1}{|\mathfrak{u}|}^{-1} \left(\widehat{c_j}^{(b)}(\mathfrak{u}\cup\{i\}) - \widehat{c_j}^{(b)}(\mathfrak{u})\right).$$
4. Compute the b bootstrap sample of $\widehat{GSh}_i$ using (3.4):
$$\widehat{GSh}_i^{(b)} = \frac{1}{d\sum_{j=1}^p \widehat{\sigma}_j^{2,(b)}} \sum_{j=1}^p \sum_{\mathfrak{u}\subseteq -i} \binom{d-1}{|\mathfrak{u}|}^{-1} \left(\widehat{c_j}^{(b)}(\mathfrak{u}\cup\{i\}) - \widehat{c_j}^{(b)}(\mathfrak{u})\right).$$

end for

splitting the sample in two parts: one for estimating the variance of the outputs, and the other to estimate the Shapley effects.

For $1 \le i \le d$, $1 \le j \le p$, let $\mathcal{R}_i = \{\widehat{GSh}_i^{(1)}, \dots, \widehat{GSh}_i^{(B)}\}$ and $\mathcal{R}_i^j = \{\widehat{Sh}_i^{j,(1)}, \dots, \widehat{Sh}_i^{j,(B)}\}$, the bias-corrected percentile method presented in [20] is applied. Let us denote by Φ the standard normal cumulative distribution function and by Φ^{-1} its inverse. A bias correction constant z_0, estimated as $\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\widehat{GSh}_i^{(b)} \in \mathcal{R}_i \text{ s. t. } \widehat{GSh}_i^{(b)} \le \widehat{GSh}_i\}}{B}\right)$ is computed (similar for $\widehat{Sh}_i^j$). Then, the corrected quantile estimate $\hat{q}(\beta)$ for $\beta \in]0,1[$ is defined as $\hat{q}_i(\beta) = \Phi(2\hat{z}_0 + z_\beta)$, where z_β satisfies $\Phi(z_\beta) = \beta$. Corrected bootstrap confidence interval of level $1-\alpha$ is estimated by the interval whose endpoints are $\hat{q}_i(\alpha/2)$ and $\hat{q}_i(1-\alpha/2)$.

To guarantee the validity of the previous BC corrected confidence interval $[\hat{q}_i(\alpha/2), \hat{q}_i(1-\alpha/2)]$, there must exist an increasing transformation g, $z_0 \in \mathbb{R}$ and $\tau > 0$ such that $g(\widehat{GSh}_i) \sim \mathcal{N}(GSh_i - \tau z_0, \tau^2)$ and $g(\widehat{GSh}_i^*) \sim \mathcal{N}(\widehat{GSh}_i - \tau z_0, \tau^2)$ where $\widehat{GSh}_i^*$ is the bootstrapped $\widehat{GSh}_i$ for fixed sample (see [19]). Normality hypothesis can be tested using traditional normality tests as Shapiro test or using graphical methods as empirical normal quantile-quantile plots.

In our application and test cases, we observed that g can be chosen as the identity. To prove empirically the performance of the procedure described in Algorithm 4.1, we compute the empirical probability of coverage (POC) of simultaneous intervals using Bonferroni correction. The POC with Bonferroni correction is the probability that the interval $[\hat{q}_i(\alpha/(2d)), \hat{q}_i(1-\alpha/(2d))]$ contains GSh_i for all $i \in \{1, \ldots, d\}$ simultaneously. To be more precise, if the confidence intervals are computed in N independent samples of size n of $(\mathbf{X}, \mathbf{Y})$. The POC is estimated as $\widehat{POC} = \sum_{k=1}^{N} \frac{w^k}{N}$, where w^k is equal to 1 if $\hat{q}_i(\alpha/(2d)) \leq GSh_i \leq \hat{q}_i(1-\alpha/(2d))$ for all i, and 0 otherwise.

5. Test cases.

In this section, we numerically study the performance of the estimation procedure and the probability coverage of the boostrap confidence intervals we introduced in the previous section. We consider two test cases: a multivariate linear Gaussian model and the functional mass spring model proposed in the work of [24]. To estimate the scalar Shapley effects, we use the function `shapleySubsetMc` of the `R` package `sensitivity` corresponding to the estimation procedure defined by (3.1), (3.2) and (3.3). Functional PCA is performed using the `R` package `FPCA` [12].

5.1. Multivariate linear Gaussian model.

We consider a multivariate linear model with two Gaussian inputs based on the example from [31]. To this toy function, there is an analytical expression of the scalar and aggregated Shapley effects (see [31]).

The model f is defined as $\mathbf{Y} = f(\mathbf{X}) = B^T\mathbf{X}$ 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 of $\mathbf{X}$ is given by:

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

and the coefficients of $B = (\beta_{ij}) \in \mathbb{R}^{2\times 3}$ are chosen as:

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

The variance of the output Y_j with $j \in \{1, 2, 3\}$ is $\sigma_{Y_j}^2 = \beta_{1j}^2\sigma_1^2 + 2\rho\beta_{1j}\beta_{2j}\sigma_1\sigma_2 + \beta_{2j}^2\sigma_2^2$.

The scalar Shapley effects are:

$$\sigma_{Y_j}^2\phi_1^j = \beta_{1j}^2\sigma_1^2\left(1 - \frac{\rho^2}{2}\right) + \rho\beta_{1j}\beta_{2j}\sigma_1\sigma_2 + \beta_2^2\sigma_2^2\frac{\rho^2}{2},$$

$$\sigma_{Y_j}^2\phi_2^j = \beta_{2j}^2\sigma_2^2\left(1 - \frac{\rho^2}{2}\right) + \rho\beta_{1j}\beta_{2j}\sigma_1\sigma_2 + \beta_1^2\sigma_1^2\frac{\rho^2}{2}.$$

Then, the aggregated Shapley effects for $i \in \{1, 2\}$ are calculated according to (3.4).

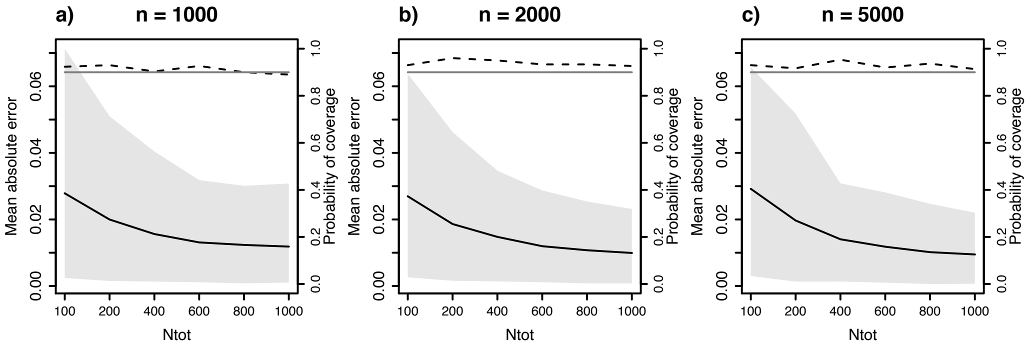

Figure 1. *Linear Gaussian model: mean absolute error of the estimation of scalar Shapley effects of the output Y_1 in N=300 i.i.d. samples in function of N_{tot} using different sample sizes a) $n = 1000$, b) $n = 2000$ and c) $n = 5000$. The 0.05 and 0.95 pointwise quantiles of the absolute error are drawn with gray polygons. The probability of coverage of the 90% bootstrap simultaneous intervals is displayed with dotted lines. The theoretical probability of coverage 0.9 is also shown with a plain gray line. The bootstrap sample size is fixed to $B = 500$.*

First, we focus on scalar Shapley effect estimation and the associated confidence intervals, for example scalar Shapley effects for Y_1 output. For Y_1 output, the most important input is X_2 with a Shapley effect of 0.66. In Figure 1, we analyze estimation accuracy and POC evolution in function of n and N_{tot}. n and N_{tot} values are fixed according to our computation budget. For each combination of n and N_{tot}, $N = 300$ independent random samples are used. To estimate the bootstrap confidence intervals, we use $B = 500$ bootstrap samples. The 95% quantile of the absolute error are displayed. Scalar Shapley effects estimation depends on n and N_{tot}. As expected, bias decreases when n and N_{tot} increase. If n is fixed, bias decreases when N_{tot} increases. In particular, bias is the smallest with $n = 5000$ and $N_{tot} = 1000$. Regardless sample sizes, POCs estimated vary around 0.9 as expected.

The estimation of the bias for aggregated Shapley effects and the POC evolution by varying n and N_{tot} are displayed in Figure 2. Similarly as for scalar effects, POC is close to 0.9, regardless the sample size and, bias reduces when n and N_{tot} increase.

We estimate Shapley effects and aggregated Shapley effects if inputs correlation is higher ($\rho = 0.9$). POC and bias results are also satisfactory (not shown). In fact, POC values vary also around 0.9 and bias decreases and goes to 0 when n and N_{tot} increases. For this simple test case, we have shown that confidence intervals using Algorithm 4.1 reach accurate coverage probability and that bias reduces when n and N_{tot} increase. Nevertheless in this test case, estimation is effortless because $d = 2$.

5.2. Mass-spring model. The method is illustrated on a test case with discretized functional output: the functional mass-spring model proposed by [24], where the displacement of a mass connected to a spring is considered:

$$m\ell''(t) + c\ell'(t) + k\ell(t) = 0, \tag{5.1}$$

with initial conditions $\ell(0) = l$, $\ell'(0) = 0$, and $t \in [1, 40]$. There exists an analytical

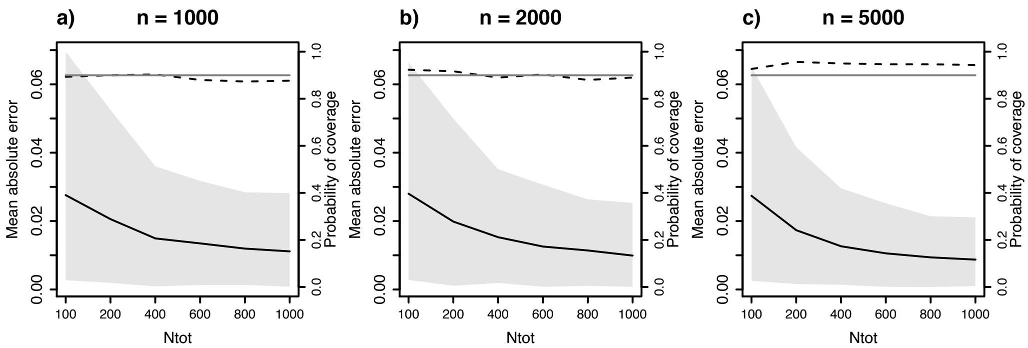

Figure 2. *Linear Gaussian model: mean absolute error of the estimation of aggregated Shapley effects in N=300 i.i.d. samples in function of N_{tot} using different sample sizes a) $n = 1000$, b) $n = 2000$ and c) $n = 5000$. The 0.05 and 0.95 pointwise quantiles of the absolute error are drawn with gray polygons. The probability of coverage of the 90% bootstrap simultaneous intervals is displayed with dotted lines. The theoretical probability of coverage 0.9 is also shown with a gray plain line. The bootstrap sample size is fixed to $B = 500$.*

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

Table 1

Mass spring model: Inputs description and uncertainty intervals. $\mathcal{U}$ denotes the uniform distribution.

solution to Equation (5.1). This model has four inputs (see more details in Table 1). The model output is the vector $\mathbf{Y} = f(\mathbf{X}) = (\ell(t_1), \ldots, \ell(t_{800})), \quad t_i = 0.05i$ with $i \in \{1, \ldots, 800\}$.

Inputs are considered independent. The true aggregated Shapley effects are unknown but they are approximated using a high sample size $n = 25\,000$ and $N_{tot} = 10\,000$. 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 [24]).

The discretized output is high-dimensional ($p = 800$). We perform fPCA (see Subsection 3.3) to estimate the effects using the first $q \ll p$ fPCs. Figure 3 shows the POC and bias evolution if different values for n and N_{tot} are used for the aggregated effects estimation. We use the first 6 fPCs which explain 95% of the output variance (see Figure 3 a). For each n and N_{tot} combination, the aggregated Shapley effects are estimated for $N = 100$ independent samples and confidence intervals are estimated with $B = 500$ bootstrap samples. Bias is large if sample size is small $n = 1000$ (see Figure 3 b). However, it reduces drastically when sample sizes increases as expected. In particular, if $n = 5000$ and $N_{tot} = 2002$ bias is the smallest (see Figure 3 d). If n and N_{tot} are too small, POC estimated values are lower than 0.9. This might be a consequence of bias in the estimation (see Figure 3 b). But when N_{tot} increases, POC is close to 0.9. In general in our experiments, confidence intervals are correct because POC values are around 0.9 when N_{tot} increases.

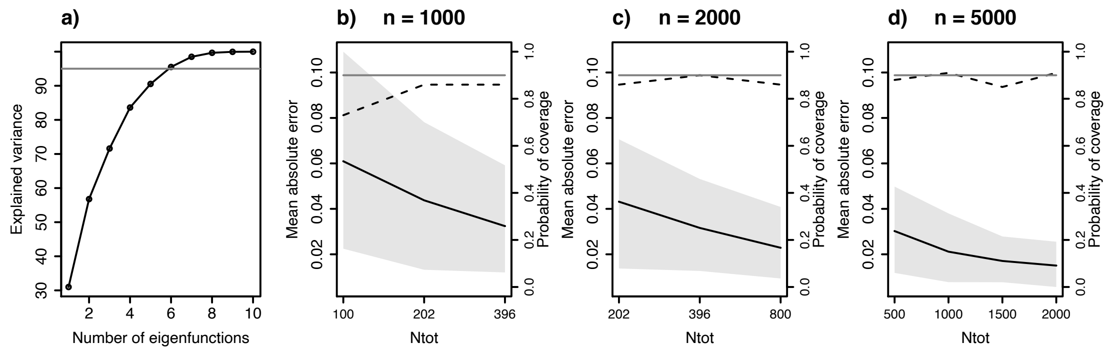

Figure 3. *Mass spring model: a) Explained variance as a function of the decomposition basis size. The gray line is displayed at 95% of the variance explained which corresponds to 6 eigenfunctions. The mean absolute error of the estimation of aggregated Shapley effects using the first 6 eigenfunctions in $N = 100$ i.i.d. samples in function of N_{tot} using sample of size b) $n = 1000$, c) $n = 2000$ and d) $n = 5000$. The 0.05 and 0.95 pointwise quantiles of the absolute error are drawn with gray polygons. The probability of coverage of the 90% bootstrap simultaneous intervals is displayed with a dotted line. The 0.9 value is also highlighted with a plain gray line. The bootstrap sample size is fixed to $B = 500$.*

6. Avalanche long term forecasting.

Our GSA method is applied to the avalanche model proposed by [45] in a general framework for a better understanding of the numerical model and in a context of risk management focusing on a well documented avalanche corridor. The objective is to determine which are the most influential input parameters on specific outputs of interest.

6.1. Model.

The avalanche model is based on depth-averaged Saint-Venant equations and considers the avalanche as a fluid in motion. In more detail, the Saint-Venant model 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 runout length distance measured from the avalanche starting abscissa. Under these assumptions, shallow-water approximations of the mass and momentum equations can be used:

$$\frac{\partial h}{\partial t} + \frac{\partial hv}{\partial x} = 0$$

$$\frac{\partial hv}{\partial t} + \frac{\partial}{\partial x}\left(hv^2 + \frac{h^2}{2}\right) = h\left(g\sin\phi - \mathrm{F}\right)$$

where $v = \|\vec{\mathbf{v}}\|$ is the flow velocity, h is the flow depth, ϕ is the local angle, t is the time, g is the gravity constant and $\mathrm{F} = \|\vec{\mathbf{F}}\|$ is a frictional force. The model uses the Voellmy frictional force $\mathrm{F} = \mu g cos\phi + \frac{g}{\xi h}v^2$, where μ and ξ are friction parameters. The equations are solved with a finite volumes scheme [43].

The numerical model depends on six inputs: the friction parameters μ and ξ, the length $\mathrm{l}_{\mathrm{start}}$ of the avalanche release zone, the snow depth $\mathrm{h}_{\mathrm{start}}$ within the release zone, the begining of the release zone denoted by $\mathrm{x}_{\mathrm{start}}$ and the discretized topography of the flow path, denoted

Input	Description	Distribution
μ	Static friction coefficient	$\mathcal{U}[0.05, 0.65]$
ξ	Turbulent friction [m/s^2]	$\mathcal{U}[400, 10000]$
$\mathrm{l_{start}}$	Length of the release zone [m]	$\mathcal{U}[5, 300]$
$\mathrm{h_{start}}$	Flow depth at the release zone [m]	$\mathcal{U}[0.05, 3]$
$\mathrm{x_{start}}$	Release abscissa [m]	$\mathcal{U}[0, 1600]$

Table 2

Avalanche model, scenario 1: Input description and uncertainty intervals. In the computation of the GSA measures, we consider $vol_{start} = l_{start} \times h_{start} \times 72.3/\cos(35°)$.

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 runout length from the starting point of the avalanche release zone 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. We use for D the topography of a path located in Bessans, France. We chose this particular path because it has been well studied in other works for example, in [16, 15, 18]. The model outputs are the flow velocity, flow depth trajectories in the path D and runout distance of an avalanche, the last one corresponds to the avalanche's distance traveled. Note that the model has two functional and one scalar outputs and these three outputs are the objects of the GSA study.

We develop our GSA in two contexts or scenarios by considering different input distributions. In the first one, input distributions are uniforms, thus GSA is applied in a general context. In the second one, input distributions are more precise and based on the results of a propagation model, then GSA is developed in the context of local avalanche risk assessment. For hazard zoning, return periods derived from rounout distances are usually considered [15]. Rougly speaking, a return period is the mean time in which a given runout distance is reached or exceeded at a given path's position [54]. In our GSAs, we put a particular emphasis on locations where avalanche events are significant with return periods varying from 10 to 10 000 years, according to the preliminary study in [15].

6.2. Scenario 1. We wish here to determine the most influential input parameters in a general context with few knowledge on input parameter distribution. We expect from GSA a better understanding of the numerical model.

6.2.1. Description. Uniform distributions are used for all the inputs. Inputs μ, ξ vary in their physical value ranges. Inputs $\mathrm{l_{start}}$ and $\mathrm{h_{start}}$ vary in their spectrum of reasonable values given by the avalanche path characteristics. The $\mathrm{x_{start}}$ input distribution is determined by calculating the abscissa interval where the release zone average slope is superior to 30°. Indeed, the slope remains above 30° during the first 1600m of the path. A good approximation of avalanche release zones is commonly obtained this way. In the following we consider that inputs $\mathrm{l_{start}}$ and $\mathrm{h_{start}}$ are related by the equation: $\mathrm{vol_{start}} = l_{start} \times h_{start} \times 72.3/\cos(35°)$, where $\mathrm{vol_{start}}$ is an approximation of the avalanche volume at the release zone, with the mean width and slope of the release zone equal to 72.3m and 35°, respectively. We then replace inputs $\mathrm{l_{start}}$ and $\mathrm{h_{start}}$ in the analysis by a single input $\mathrm{vol_{start}}$. These input scenario and their uncertainty intervals are described in Table 2. The input correlations are close to 0 since we assume they are a priori independent.

For a given avalanche simulation, its functional velocity and flow depth outputs have a high number of zeros because they are null before the release zone and after the runout zone. Also, there might be some avalanche simulations that are meaningless in physical or risk terms. Therefore to perform GSA, we select simulations that accomplish the following acceptance-rejection (AR) rules: (i) avalanche simulation is flowing in the interval $[1600m, 2412m]$, (ii) its volume is superior to 7000 m^3 and, (iii) avalanche runout zone is inferior to 2500m which corresponds to the end of the path. Indeed physically and in terms of risk assessment, only this set of avalanches is interesting for the GSA study because first, the return periods in the interval $[1600m, 2412m]$ vary from 1 to 10 000 years. Second, we focus on medium, large and very large avalanches which have a high potential damage and third, our GSA is focus on topography D, thus runout zones outside the path are not useful for our study purpose. From the initial simulations, we only keep the ones satisfying (i) to (iii), which is the AR sample used to carry out the GSA.

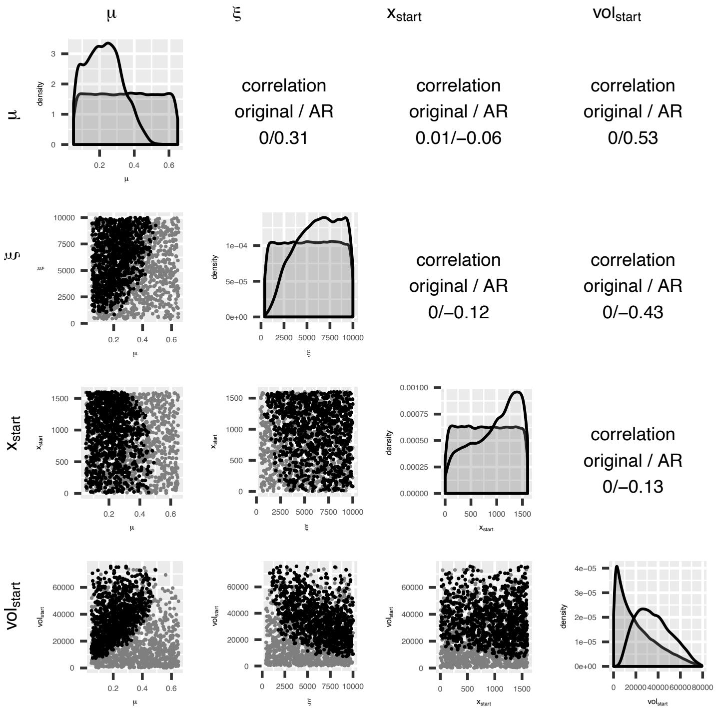

Figure 4. *Avalanche model, scenario 1: scatter-plots of initial (black points) and acceptance rejection (gray points) samples. In the figure's diagonal, the density function of the initial (gray color) and AR (transparent) samples are displayed. Input correlations of the original and AR samples are shown. 1000 subsamples of original and AR samples are used for illustration purpose.*

6.2.2. 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 2. Then, by applying (i) to (iii) our AR sample size was reduced to $n_1 = 6152$. The main characteristics of the AR sampling can be observed on Figure 4, on which we have drawn the initial sample with black points and the AR sample with gray points. Even if the initial sample size is high $n_0 = 100\,000$ and if the corresponding input parameter sample does not present any significant correlation structure, the AR sample size is low and we can observe a correlation structure. For example, inputs μ and ξ were independent for the initial sample but the correlation computed after the AR algorithm is 0.31. Note that the input parameter correlations induced by the AR algorithm were the main motivation to compute Shapley effects and not Sobol' indices in this first scenario.

On Figure 5 are plotted highest density region (HDR) boxplots for the velocity and the snow depth curves in the GSA studied interval, obtained by using the R package `rainbow` developed by [29]. The HDR boxplot is a vizualization tool for functional data based on the density estimation of the first two components of the PCA decomposition of the observed functions (see [28] for further details). In the interval, the avalanche velocity ranges from $0.1\mathrm{ms}^{-1}$ to $71.56\mathrm{ms}^{-1}$ and avalanches are in deceleration phase (see Figure 5 a). Flow depths vary from 0.03m to 7.52m. The flow depth curves exhibit high fluctuations in $[2100m, 2300m]$ (see Figure 5 b) which corresponds to a path's convexity region. Runout distances vary from 815.2m to 2478.2m (see Figure 5 c). Long runout distances characterize very large avalanches.

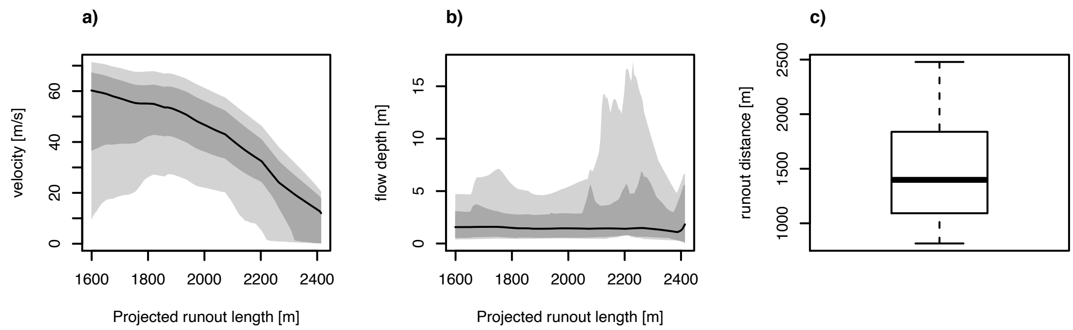

Figure 5. *Avalanche model, scenario 1: a) and b) functional HDR boxplots of velocity and flow depth curves, resp. It is shown 50% HDR (light gray), 100% HDR(dark gray) and modal curve (black line). c) runout distance boxplot. The AR sample size is* $n_1 = 6152$.

On Figure 6 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 track but its importance decreases along the path and conversely, the importance of the other inputs increases. For snow depth output, the most important input is $\mathrm{vol}_{\mathrm{start}}$ 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 6 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. For runout distance, the most relevant input is $\mathrm{x}_{\mathrm{start}}$.

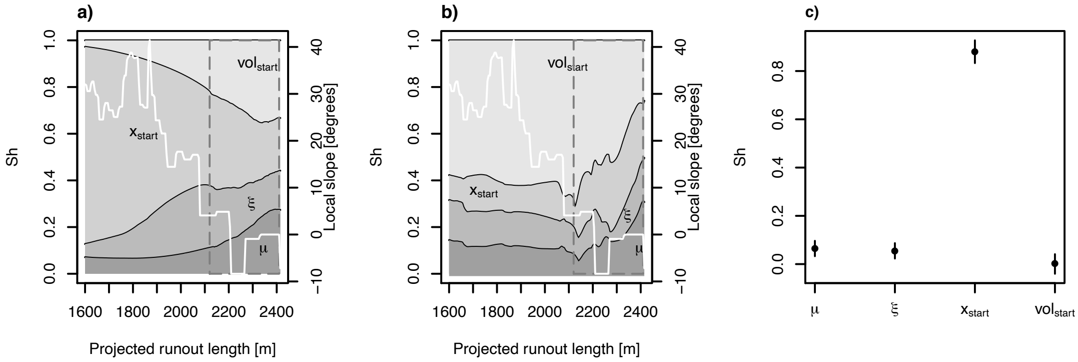

Figure 6. *Avalanche model, scenario 1: a) and b) ubiquitous Shapley effects of velocity and flow depth curves, resp. and, c) runout distance Shapley effects. Shapley effects are estimated with a sample of size 6152 and Ntot=2000. The local slope is displayed with a white line. A gray dotted rectangle box is displayed at interval [2017, 2412] where return periods vary from 10 to 10 000 years. The bootstrap sample size is fixed to* $B = 500$.

Figure 7 shows aggregated Shapley effects and 90% confidence intervals computed over space intervals $[x, 2412]$ where $x \in \{1600, 1700, \ldots, 2412\}$. 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 local slope high variation regions (see Figure 7 compared to Figure 6). For explaining more than 95% of the velocity output variance, 2 fPCs are required, while, or explaining more than 95% of the flow depth output variance, at most 4 fPCs are required, depending on x. Note that on Figure 7, the Shapley effects that are computed are integrated on the interval $[x, 2412m]$. For the velocity output, the most important input is $\mathrm{x_{start}}$ in the interval $[1600m, 2100m]$ but its importance decreases along the path. In the interval $[2017m, 2412m]$ where return periods are non trivial, $\mathrm{x_{start}}$ and $\mathrm{vol_{start}}$ are the most important followed by μ and ξ. For the flow depth output, $\mathrm{vol_{start}}$ 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 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.

6.3. Scenario 2. The aim is now to determine the most influential inputs in a local avalanche risk context with a strong knowledge of input distribution.

6.3.1. Description. In [15], the authors considered a Bayesian framework in a long-term avalanche hazard assessment to estimate input distribution in the path under study. Input ξ is fixed to 1300. In avalanche literature, it is assumed that ξ depends on the path topography and given that D is fixed it seems reasonable to use a constant ξ value. Input parameters in this scenario are dependent. The dependence between $\mathrm{h_{start}}$ and $\mathrm{l_{start}}$ is modeled with a linear function $\mathrm{l_{start}} = 31.25 + 87.5\mathrm{h_{start}}$, and similarly as in scenario 1, we consider $\mathrm{vol_{start}}$ as input

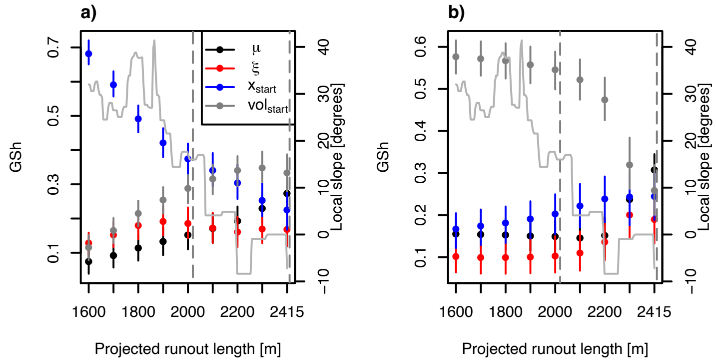

Figure 7. *Avalanche model, scenario 1: a) and b) aggregated Shapley effects of velocity and flow depth curves calculated over space intervals* $[x, 2412m]$ *where* $x \in \{1600m, 1700m, \ldots, 2412m\}$*. Shapley effects are estimated with samples of size 6152 and Ntot=2000. Effects are estimated using 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 return periods vary from 10 to 10 000 years. The bootstrap sample size is fixed to* $B = 500$.

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

Table 3

Avalanche model: Scenario 2. Input description and uncertainty intervals. x_{nstart} *is a normalization of* x_{start}*. There is a well known linear relationship between* h_{start} *and* l_{start} *in the avalanche path. In the computation of the GSA measures, we consider* $vol_{start} = l_{start} \times h_{start} \times 72.3/\cos(35°)$.

instead of $\mathrm{h_{start}}$ and $\mathrm{l_{start}}$. The complete input distribution resulting from the study in [15] is described in Table 3. Input correlations have been computed. As an example, the correlation between μ and $\mathrm{vol_{start}}$ is 0.8.

To perform GSA in this scenario, our AR rules are: (i) avalanche is flowing in the interval [1600m, 2204m] where return periods vary from 10 to 300 years, (ii) avalanche volume is superior to $7000\mathrm{m}^3$ and, (iii) μ coefficient is inferior to 0.39 as we focus on dry snow avalanches. Under these conditions, we recover a set of potential threat avalanches which could cause strong material or human damages.

6.3.2. Global sensitivity analysis results. We first ran $n_0 = 100\,000$ avalanches from an i.i.d. sample of input distribution following Table 3. After applying the AR algorithm, the sample size was reduced to $n_2 = 1284$ and the input distribution suffers some changes. For example, μ and $\mathrm{vol_{start}}$ correlation changes from 0.8 to 0.2 which is still non negligible. Ubiquitous Shapley effects are displayed on Figure 8 panels a and b. For the velocity, the three inputs have a similar importance till 1900m, then $\mathrm{vol_{start}}$ importance decreases and μ and $\mathrm{x_{start}}$ importance increases (see Figure 8 a). Similarly as in scenario 1, the effects show fluctuations which correspond to changes in local slope. In particular for the flow depth output, input effects suffer radical changes when the local slope decreases from 20° to 10° (see Figure 8 b). For runout distance, the most relevant input is x_{start} (see Figure 8 c).

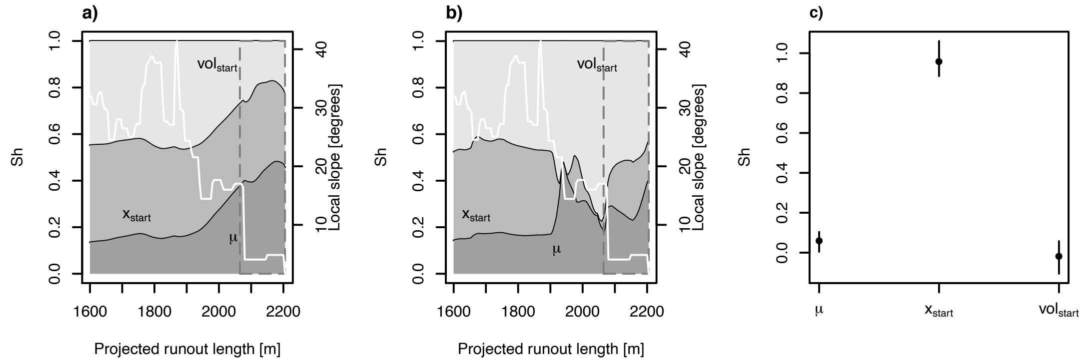

Figure 8. *Avalanche model, scenario 2: a) and b) ubiquitous Shapley effects of velocity and flow depth curves, c) runout distance Shapley effects. Shapley effects are estimated with samples of size 1284 and Ntot=800. The local slope is displayed with a white line. A gray dotted rectangle shows the interval [2064, 2204] where return periods vary from 10 to 300 years. The bootstrap sample size is fixed to* $B = 500$.

Aggregated effects (see Figure 9) present less fluctuations and are easier to interpret (see Figure 8). 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 9). To outperform the estimation accuracy at the end of the path, it would be interesting to generate a larger initial sample of avalanches. Then the costs would be prohibitive, thus it would be necessary to first learn a surrogate model and then to use it for running simulations.

7. Conclusions and perspectives. In this work, we extended Shapley effects to models with multivariate or functional outputs. We proved that aggregated Shapley effects accomplish the natural requirements for a GSA measure. For the estimation, we proposed to extend the subset aggregation procedure with double Monte Carlo given data estimator of [9]. Also, we proposed an algorithm to construct bootstrap confidence intervals for scalar and aggregated Shapley effects based on the ideas of [5]. In test cases, the convergence of our estimator was empirically studied. Also, we proved empirically that the bootstrap confidence intervals we proposed have accurate coverage probability. Estimation and bootstrap confidence interval algorithms well behave. Nevertheless, high sample sizes ($n = 5000$ and $N_{tot} = 2000$) are required to guarantee accurate results. Remark that it is well known that Shapley effects estimation is costly. It would be interesting to study theoretically the asymptotic properties of our estimator, but this study is out of the scope of this paper. Recently, in the R package `sensitivity` the function `sobolshap_knn` to estimate Shapley effects with n and N_{tot} from a given data sample has been implemented. This function uses a tree based technique to approximate nearest-neighbor search which reduces drastically computation times. The function is partic-

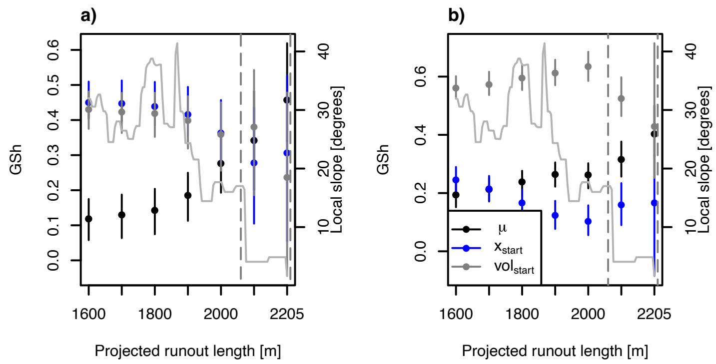

Figure 9. *Avalanche model, scenario 2: a) and b) aggregated Shapley effects of velocity and flow depth curves calculated over space intervals $[x, 2204]$ where $x \in \{1600, 1700, \dots, 2204\}$ and using the first fPCs which have 95% of output variance. Shapley effects are estimated with samples of size 1284 and Ntot=800. The local slope is displayed with a gray line. A gray dotted rectangle is displayed at [2017m, 2204m] where return periods vary from 10 to 300 years. The bootstrap sample size is fixed to $B = 500$.*

ularly attractive if n and N_{tot} are high, we could even use $N_{tot} = (2^d - 2) \times n$. We did not use in the present work this function as we were not able to obtain confidence intervals with accurate coverage probability for the estimation it computes. We rather used the `shapleySubsetMc` function which corresponds to the estimator introduced in [9] on which our estimator for aggregated Shapley effects is based. Our GSA method was applied to an avalanche model whose outputs are velocity, flow depth trajectories and runout distance. Model samples for this application were obtained from an acceptance-rejection (AR) algorithm. Moreover, input parameters in this application were not necessarily confined in a rectangular region. For these reasons, it was not possible to consider independence of input parameters. The key advantages of the procedure we proposed in this paper are that it does not require independence of input parameters and that it handles functional outputs such as space and/or time dependent processes. We considered two different settings, a general one where we have little knowledge of input distributions, and a local one which focuses on a well documented avalanche corridor. In the application, we observed that the estimation of aggregated Shapley effects was more stable and easier to interpret than ubiquitous effects. The same observation was done by [1] in the case of aggregated Sobol' indices. Thus depending on the GSA study objectives, users might rather use aggregated Shapley effects than ubiquitous effects. Application is challenging because AR samples are generally of moderate size, for example, from the 100 000 initial sample, the AR sampling produced a 6000 to 1200 sample, depending on the scenario. In a future work, it would be useful to construct a surrogate of the avalanche model to generate larger AR samples. Indeed within larger samples, we could improve the accuracy of aggregated Shapley effect estimation and thus reduce confidence intervals width.

Acknowledgments. We thank Sébastien Da Veiga for fruitful discussions on nearest-neighbor estimation of Shapley effects. M.B. Heredia holds a Ph.D. grant from OSUG@2020 labex. Within the CDP-Trajectories framework, this work is supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-

15-IDEX-02). Part of the computations were performed using the Froggy platform of the CIMENT infrastructure (https://ciment.ujf-grenoble.fr), which is supported by the Rhône-Alpes region (GRANT CPER07 13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the program "Investissements d'avenir" supported by the Agence Nationale pour la Recherche.

REFERENCES

[1] A. Alexanderian, P. A. Gremaud, and R. C. Smith, *Variance-based sensitivity analysis for time-dependent processes*, Reliability Engineering & System Safety, 196 (2020), p. 106722, https://doi.org/https://doi.org/10.1016/j.ress.2019.106722, http://www.sciencedirect.com/science/article/pii/S0951832019303837.

[2] A. Antoniadis, C. Helbert, C. Prieur, and L. Viry, *Spatio-temporal metamodeling for West African monsoon*, Environmetrics, 23 (2012), pp. 24–36, https://doi.org/10.1002/env.1134, https://onlinelibrary.wiley.com/doi/abs/10.1002/env.1134.

[3] I. Antoniano-Villalobos, E. Borgonovo, and X. Lu, *Nonparametric estimation of probabilistic sensitivity measures*, Statistics and Computing, (2019), https://doi.org/10.1007/s11222-019-09887-9, https://doi.org/10.1007/s11222-019-09887-9.

[4] B. Auder and I. B., *Global sensitivity analysis based on entropy*, in In Safety, Reliability and Risk Analysis - Proceedings of the ESREL 2008 Conference, CRC Press, 2008, pp. 2107–2115.

[5] Benoumechiara, Nazih and Elie-Dit-Cosaque, Kevin, *Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms*, ESAIM: ProcS, 65 (2019), pp. 266–293, https://doi.org/10.1051/proc/201965266, https://doi.org/10.1051/proc/201965266.

[6] E. Borgonovo, *A new uncertainty importance measure*, Reliability Engineering & System Safety, 92 (2007), pp. 771 – 784, https://doi.org/https://doi.org/10.1016/j.ress.2006.04.015, http://www.sciencedirect.com/science/article/pii/S0951832006000883.

[7] E. Borgonovo, G. B. Hazen, and E. Plischke, *A Common Rationale for Global Sensitivity Measures and Their Estimation*, Risk Analysis, 36 (2016), pp. 1871–1895, https://doi.org/10.1111/risa.12555, https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.12555.

[8] E. Borgonovo and E. Plischke, *Sensitivity analysis: A review of recent advances*, European Journal of Operational Research, 248 (2016), pp. 869–887, https://doi.org/10.1016/j.ejor.2015.06.032, https://doi.org/10.1016/j.ejor.2015.06.032.

[9] B. Broto, F. Bachoc, and M. Depecker, *Variance Reduction for Estimation of Shapley Effects and Adaptation to Unknown Input Distribution*, SIAM/ASA Journal on Uncertainty Quantification, 8 (2020), pp. 693–716, https://doi.org/10.1137/18M1234631, https://doi.org/10.1137/18M1234631.

[10] W. B. Capra and H.-G. Müller, *An accelerated-time model for response curves*, Journal of the American Statistical Association, 92 (1997), pp. 72–83, https://doi.org/10.1080/01621459.1997.10473604, https://doi.org/10.1080/01621459.1997.10473604.

[11] G. Chastaing, F. Gamboa, and C. Prieur, *Generalized Hoeffding-Sobol decomposition for dependent variables - application to sensitivity analysis*, Electron. J. Statist., 6 (2012), pp. 2420–2448, https://doi.org/10.1214/12-EJS749, https://doi.org/10.1214/12-EJS749.

[12] Y. Chen, C. Carroll, X. Dai, J. Fan, P. Z. Hadjipantelis, K. Han, H. Ji, H.-G. Mueller, and J.-L. Wang, *fdapace: Functional Data Analysis and Empirical Dynamics*, 2020, https://CRAN.R-project.org/package=fdapace. R package version 0.5.1.

[13] S. Da Veiga and F. Gamboa, *Efficient estimation of sensitivity indices*, Journal of Nonparametric Statistics, 25 (2013), pp. 573–595, https://doi.org/10.1080/10485252.2013.784.

[14] G. Delipei, *Développement d'une méthodologie de Quantification d'Incertitudes pour une analyse Multi-Physique Best Estimate et application sur un Accident d'Éjection de Grappe dans un Réacteur à Eau Pressurisée*, PhD thesis, 2019, http://www.theses.fr/2019SACLX078. Thèse de doctorat dirigée par Garnier, Josselin Mathématiques appliquées Université Paris-Saclay (ComUE) 2019.

[15] N. Eckert, M. Naaim, and E. Parent, *Long-term avalanche hazard assessment with a Bayesian depth averaged propagation model*, Journal of Glaciology, 56 (2010), pp. 563–586, https://doi.org/10.3189/

002214310793146331.

[16] N. Eckert, E. Parent, T. Faug, and 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 (2008), pp. 1067–1081, https://doi.org/10.5194/nhess-8-1067-2008, https://www.nat-hazards-earth-syst-sci.net/8/1067/2008/.

[17] N. Eckert, E. Parent, M. Naaim, and D. Richard, *Bayesian stochastic modelling for avalanche predetermination: from a general system framework to return period computations*, Stochastic Environmental Research and Risk Assessment, (2008), pp. 185–206, https://doi.org/10.1007/s00477-007-0107-4, https://doi.org/10.1007/s00477-007-0107-4.

[18] Eckert, Nicolas, Naaim, Mohamed, Giacona, Florie, Favier, Philomène, Lavigne, Aurore, Richard, Didier, Bourrier, Franck, and Parent, Eric, *Repenser les fondements du zonage règlementaire des risques en montagne « récurrents »*, La Houille Blanche, (2018), pp. 38–67, https://doi.org/10.1051/lhb/2018019, https://doi.org/10.1051/lhb/2018019.

[19] B. Efron, *Nonparametric standard errors and confidence intervals*, Canadian Journal of Statistics, 9 (1981), pp. 139–158, https://doi.org/10.2307/3314608, https://onlinelibrary.wiley.com/doi/abs/10.2307/3314608.

[20] B. Efron and R. Tibshirani, *Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy*, Statist. Sci., 1 (1986), pp. 54–75, https://doi.org/10.1214/ss/1177013815, https://doi.org/10.1214/ss/1177013815.

[21] J. Fan and I. Gijbels, *Local polynomial modelling and its applications*, no. 66 in Monographs on statistics and applied probability series, Chapman & Hall, London, 1996.

[22] P. Favier, D. Bertrand, N. Eckert, and M. Naaim, *A reliability assessment of physical vulnerability of reinforced concrete walls loaded by snow avalanches*, Natural Hazards and Earth System Sciences, 14 (2014), pp. 689–704, https://doi.org/10.5194/nhess-14-689-2014.

[23] F. Gamboa, P. Gremaud, T. Klein, and A. Lagnoux, *Global sensitivity analysis: a new generation of mighty estimators based on rank statistics*, arXiv preprint arXiv:2003.01772, (2020).

[24] F. Gamboa, A. Janon, T. Klein, and A. Lagnoux, *Sensitivity analysis for multidimensional and functional outputs*, ArXiv e-prints, (2013), https://arxiv.org/abs/1311.1797.

[25] F. Gamboa, A. Janon, T. Klein, and A. Lagnoux, *Sensitivity indices for multivariate outputs*, Comptes Rendus Mathematique, 351 (2013), pp. 307 – 310, https://doi.org/https://doi.org/10.1016/j.crma.2013.04.016, http://www.sciencedirect.com/science/article/pii/S1631073X13000824.

[26] Q. Ge and M. Menendez, *Extending Morris method for qualitative global sensitivity analysis of models with dependent inputs*, Reliability Engineering & System Safety, 162 (2017), pp. 28 – 39, https://doi.org/https://doi.org/10.1016/j.ress.2017.01.010, http://www.sciencedirect.com/science/article/pii/S0951832017300625.

[27] J. Hart and P. A. Gremaud, *An approximation theoretic perspective of sobol' indices with dependent variables*, International Journal for Uncertainty Quantification, 8 (2018), pp. 483–493, https://doi.org/10.1615/Int.J.UncertaintyQuantification.2018026498.

[28] R. J. Hyndman, *Computing and graphing highest density regions*, The American Statistician, 50 (1996), pp. 120–126, https://doi.org/10.2307/2684423.

[29] R. J. Hyndman and H. L. Shang, *Rainbow plots, bagplots, and boxplots for functional data*, Journal of Computational and Graphical Statistics, 19 (2010), pp. 29–45, https://doi.org/10.1198/jcgs.2009.08158, https://doi.org/10.1198/jcgs.2009.08158.

[30] B. Iooss and P. Lemaître, *A Review on Global Sensitivity Analysis Methods*, Springer US, Boston, MA, 2015, pp. 101–122, https://doi.org/10.1007/978-1-4899-7547-8_5, https://doi.org/10.1007/978-1-4899-7547-8_5.

[31] B. Iooss and C. Prieur, *Shapley effects for sensitivity analysis with correlated inputs: Comparisons with Sobol' indices, numerical estimation and applications*, International Journal for Uncertainty Quantification, 9 (2019), pp. 493–514, https://doi.org/10.1615/Int.J.UncertaintyQuantification.2019028372.

[32] M. Irwin and Z. Wang, *Dynamic Systems Modeling*, American Cancer Society, 2017, pp. 1–12, https://doi.org/10.1002/9781118901731.iecrm0074, https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118901731.iecrm0074.

[33] J. Jacques, C. Lavergne, and N. Devictor, *Sensitivity analysis in presence of model uncertainty and correlated inputs*, Reliability Engineering & System Safety, 91 (2006), pp. 1126 – 1134, https:

//doi.org/https://doi.org/10.1016/j.ress.2005.11.047, http://www.sciencedirect.com/science/article/pii/S0951832005002231. The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004).

[34] A. Janon, M. Nodet, and C. Prieur, *Uncertainties assessment in global sensitivity indices estimation from metamodels*, International Journal for Uncertainty Quantification, 4 (2014), pp. 21–36, https://doi.org/10.1615/Int.J.UncertaintyQuantification.2012004291.

[35] S. Kucherenko, O. Klymenko, and N. Shah, *Sobol' indices for problems defined in non-rectangular domains*, Reliability Engineering & System Safety, 167 (2017), pp. 218 – 231, https://doi.org/https://doi.org/10.1016/j.ress.2017.06.001, http://www.sciencedirect.com/science/article/pii/S0951832016301363. Special Section: Applications of Probabilistic Graphical Models in Dependability, Diagnosis and Prognosis.

[36] S. Kucherenko, S. Tarantola, and P. Annoni, *Estimation of global sensitivity indices for models with dependent variables*, Computer Physics Communications, 183 (2012), pp. 937 – 946, https://doi.org/https://doi.org/10.1016/j.cpc.2011.12.020, http://www.sciencedirect.com/science/article/pii/S0010465511004085.

[37] M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, and H. Monod, *Multivariate global sensitivity analysis for dynamic crop models*, Field Crops Research, 113 (2009), pp. 312 – 320, https://doi.org/https://doi.org/10.1016/j.fcr.2009.06.007, http://www.sciencedirect.com/science/article/pii/S0378429009001531.

[38] M. Lamboni, H. Monod, and D. Makowski, *Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models*, Reliability Engineering & System Safety, 96 (2011), pp. 450 – 459, https://doi.org/https://doi.org/10.1016/j.ress.2010.12.002, http://www.sciencedirect.com/science/article/pii/S0951832010002504.

[39] G. Li, H. Rabitz, P. E. Yelvington, O. O. Oluwole, F. Bacon, C. E. Kolb, and J. Schoendorf, *Global sensitivity analysis for systems with independent and/or correlated inputs*, The Journal of Physical Chemistry A, 114 (2010), pp. 6022–6032, https://doi.org/10.1021/jp9096919, https://doi.org/10.1021/jp9096919. PMID: 20420436.

[40] A. López-Benito and R. Bolado-Lavín, *A case study on global sensitivity analysis with dependent inputs: The natural gas transmission model*, Reliability Engineering & System Safety, 165 (2017), pp. 11 – 21, https://doi.org/https://doi.org/10.1016/j.ress.2017.03.019, http://www.sciencedirect.com/science/article/pii/S0951832017303447.

[41] T. A. Mara and S. Tarantola, *Variance-based sensitivity indices for models with dependent inputs*, Reliability Engineering & System Safety, 107 (2012), pp. 115 – 121, https://doi.org/https://doi.org/10.1016/j.ress.2011.08.008, http://www.sciencedirect.com/science/article/pii/S0951832011001724. SAMO 2010.

[42] T. A. Mara, S. Tarantola, and P. Annoni, *Non-parametric methods for global sensitivity analysis of model output with dependent inputs*, Environmental Modelling & Software, 72 (2015), pp. 173 – 183, https://doi.org/https://doi.org/10.1016/j.envsoft.2015.07.010, http://www.sciencedirect.com/science/article/pii/S1364815215300153.

[43] M. Naaim, *Dense avalanche numerical modeling: interaction between avalanche and structures*, in 25 years of snow avalanche research, Voss, NOR, 12-16 May 1998, Norway, 1998, pp. 187–191, https://hal.inrae.fr/hal-02582417.

[44] M. Naaim, T. Faug, F. Naaim, and N. Eckert, *Return period calculation and passive structure design at the Taconnaz avalanche path, France*, Annals of Glaciology, 51 (2010), pp. 89–97, https://doi.org/10.3189/172756410791386517.

[45] M. Naaim, F. Naaim-Bouvet, T. Faug, and A. Bouchet, *Dense snow avalanche modeling: flow, erosion, deposition and obstacle effects*, Cold Regions Science and Technology, 39 (2004), pp. 193 – 204, https://doi.org/https://doi.org/10.1016/j.coldregions.2004.07.001, http://www.sciencedirect.com/science/article/pii/S0165232X04000643. Snow And Avalanches: Papers Presented At The European Geophysical Union Conference, Nice, April 2003. Dedicated To The Avalanche Dynamics Pioneer Dr. B. Salm.

[46] A. Owen, *Sobol' Indices and Shapley Value*, SIAM/ASA Journal on Uncertainty Quantification, 2 (2014), pp. 245–251, https://doi.org/10.1137/130936233, https://doi.org/10.1137/130936233.

[47] A. B. Owen and C. Prieur, *On Shapley value for measuring importance of dependent inputs*,

SIAM/ASA Journal on Uncertainty Quantification, 5 (2017), https://doi.org/10.1137/16M1097717.
[48] E. Plischke, *An effective algorithm for computing global sensitivity indices (EASI)*, Reliability Engineering & System Safety, 95 (2010), pp. 354 – 360, https://doi.org/https://doi.org/10.1016/j.ress.2009.11.005, http://www.sciencedirect.com/science/article/pii/S0951832009002579.
[49] E. Plischke, E. Borgonovo, and C. L. Smith, *Global sensitivity measures from given data*, European Journal of Operational Research, 226 (2013), pp. 536–550, https://doi.org/10.1016/j.ejor.2012.11.06.
[50] E. Plischke, G. Rabitti, and E. Borgonovo, *Computing Shapley Effects for Sensitivity Analysis*, arXiv e-prints, (2020), arXiv:2002.12024, p. arXiv:2002.12024, https://arxiv.org/abs/2002.12024.
[51] G. Rabitti and E. Borgonovo, *A Shapley–Owen index for interaction quantification*, SIAM/ASA Journal on Uncertainty Quantification, 7 (2019), pp. 1060–1075, https://doi.org/10.1137/18M1221801, https://doi.org/10.1137/18M1221801.
[52] M. I. Radaideh, S. Surani, D. OGrady, and T. Kozlowski, *Shapley effect application for variance-based sensitivity analysis of the few-group cross-sections*, Annals of Nuclear Energy, 129 (2019), pp. 264 – 279, https://doi.org/https://doi.org/10.1016/j.anucene.2019.02.002, http://www.sciencedirect.com/science/article/pii/S0306454919300714.
[53] J. A. Rice and B. W. Silverman, *Estimating the mean and covariance structure nonparametrically when the data are curves*, Journal of the Royal Statistical Society: Series B (Methodological), 53 (1991), pp. 233–243, https://doi.org/10.1111/j.2517-6161.1991.tb01821.x, https://doi.org/10.1111%2Fj.2517-6161.1991.tb01821.x.
[54] R. Schläppy, N. Eckert, V. Jomelli, M. Stoffel, D. Grancher, D. Brunstein, M. Naaim, and 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), pp. 12 – 26, https://doi.org/https://doi.org/10.1016/j.coldregions.2013.12.001, http://www.sciencedirect.com/science/article/pii/S0165232X13001900.
[55] L. Shapley, *A Value for n-Person Games*, Contributions to the Theory of Games (AM-28), Princeton: Princeton University Press., 2 (1953).
[56] I. M. Sobol', *Sensitivity analysis for non-linear mathematical models*, Mathematical Modelling and Computational Experiment, 1 (1993), pp. 407–414.
[57] M. Solís, *Non-parametric estimation of the first-order Sobol indices with bootstrap bandwidth*, Communications in Statistics - Simulation and Computation, 0 (2019), pp. 1–16, https://doi.org/10.1080/03610918.2019.1655575, https://doi.org/10.1080/03610918.2019.1655575.
[58] E. Song, B. Nelson, and J. Staum, *Shapley effects for global sensitivity analysis: Theory and computation*, SIAM/ASA Journal on Uncertainty Quantification, 4 (2016), pp. 1060–1083, https://doi.org/10.1137/15M1048070, https://doi.org/10.1137/15M1048070.
[59] Y. Sun, D. W. Apley, and J. Staum, *Efficient nested simulation for estimating the variance of a conditional expectation*, Operations Research, 59 (2011), pp. 998–1007, https://doi.org/10.1287/opre.1110.0932, https://doi.org/10.1287/opre.1110.0932.
[60] S. D. Veiga, *Global sensitivity analysis with dependence measures*, Journal of Statistical Computation and Simulation, 85 (2015), pp. 1283–1305, https://doi.org/10.1080/00949655.2014.945932, https://doi.org/10.1080/00949655.2014.945932.
[61] C. Xu, *Decoupling correlated and uncorrelated parametric uncertainty contributions for nonlinear models*, Applied Mathematical Modelling, 37 (2013), pp. 9950 – 9969, https://doi.org/https://doi.org/10.1016/j.apm.2013.05.036, http://www.sciencedirect.com/science/article/pii/S0307904X13003570.
[62] C. Xu and G. Z. Gertner, *Uncertainty and sensitivity analysis for models with correlated parameters*, Reliability Engineering & System Safety, 93 (2008), pp. 1563 – 1573, https://doi.org/https://doi.org/10.1016/j.ress.2007.06.003, http://www.sciencedirect.com/science/article/pii/S0951832007001652.
[63] F. Yao, H.-G. Müller, and J.-L. Wang, *Functional data analysis for sparse longitudinal data*, Journal of the American Statistical Association, 100 (2005), pp. 577–590, https://doi.org/10.1198/016214504000001745, https://doi.org/10.1198/016214504000001745.
[64] K. Zhang, Z. Lu, L. Cheng, and F. Xu, *A new framework of variance based global sensitivity analysis for models with correlated inputs*, Structural Safety, 55 (2015), pp. 1 – 9, https://doi.org/https://doi.org/10.1016/j.strusafe.2014.12.005, http://www.sciencedirect.com/science/article/pii/S0167473015000181.