Aggregated Shapley effects: nearest-neighbor estimation procedure and confidence intervals. Application to snow avalanche modeling.

María Belén Heredia, Clémentine Prieur*, Nicolas Eckert+

* Grenoble Alpes University, Inria, AIRSEA
+ Grenoble Alpes University, INRAE, ZTNA

Grenoble, 18 September 2020.
Natural phenomena are complex.
Natural phenomena are complex.

Given some initial conditions:

\[
\frac{\partial h}{\partial t} + \frac{\partial h v}{\partial x} = 0
\]

\[
\frac{\partial h v}{\partial t} + \frac{\partial}{\partial x} \left( h v^2 + \frac{h^2}{2} \right) = h (g \sin \phi - \Gamma)
\]

where \( v = \| \vec{v} \| \) is the flow velocity, \( h \) is the flow depth, \( \phi \) is the local angle, \( t \) is the time, \( g \) is the gravity constant and \( \Gamma = \| \vec{F} \| \) is the Voellmy frictional force,

\[
F = \mu g \cos \phi + \frac{g}{\xi} v^2,
\]

where \( \mu \) and \( \xi \) are the friction parameters (see more detail in [Naaim et al., 2004]).
Natural phenomena are complex.

Black box model
Aggregated Shapley effects

Motivation

\[ f : \mathcal{D} = \text{inputs} \rightarrow \mathbb{R} \]

\[ Y = \left\{ \right. \]

\[ x_{\text{min}}, x_{\text{max}}, \in \mathbb{R} \rightarrow \mathbb{R} \]

\[ h : \mathcal{D} = \text{outputs} \rightarrow \mathbb{R} \]

\[ x_{\text{start}}, x_{\text{end}}, \in \mathbb{R} \rightarrow \mathbb{R} \]

\[ x_{\text{amount}} \in \mathbb{R} \]
To get meaningful samples, we apply acceptance-rejection (AR) sampling:
The ingredients for our global sensitivity analysis (GSA) problem are:

- input parameters leading to significant snow avalanches are dependent,
- the sample is given from the AR sampling and not drawn based on a specific estimation strategy (pick-freeze, replicated designs,...),
- two of the three outputs are functional.
GSA framework

We aim at determining which input parameters contribute the most to a given quantity of interest (defined from the output of the model).
We aim at determining which input parameters contribute the most to a given quantity of interest (defined from the output of the model).
GSA framework

We aim at determining which input parameters contribute the most to a given quantity of interest (defined from the output of the model).

\[
\begin{align*}
\text{Inputs} & \quad X = (X_1, \ldots, X_d) \\
& \quad X_i : \Omega \to \mathbb{R} \quad \forall i \in \{1, \ldots, d\} \\
\Rightarrow & \quad \text{Model} \\
& \quad f(X) = Y \\
& \quad Y : \Omega \to \mathbb{R}^p
\end{align*}
\]

Depending on the quantities of interest:

- variance-based (Sobol’ indices [Sobol’, 1993], Shapley effects [Owen, 2014]),
- density-based indices or moment-free measures [Borgonovo, 2007, Da Veiga, 2015],
- derivative-based measures [Sobol’ and Kucherenko, 2009, Lamboni et al., 2013].
GSA framework

We aim at determining which input parameters contribute the most to a given quantity of interest (defined from the output of the model).

Inputs
\[ X = (X_1, \ldots, X_d) \]
\[ X_i : \Omega \rightarrow \mathbb{R} \text{ } \forall i \in \{1, \ldots, d\} \]

Model
\[ f(X) = Y \]
\[ Y : \Omega \rightarrow \mathbb{R}^p \]

Shapley effects are the ideal framework to our problem!

- they are meaningful even for dependent inputs [Owen and Prieur, 2017, Loos and Prieur, 2019],
- there exists a given data estimation method [Broto et al., 2020].

Moreover,
- we can extend them to multivariate and functional outputs adapting the propositions in [Campbell et al., 2006, Lamboni et al., 2009, Gamboa et al., 2013, Alexanderian et al., 2020],
- we propose a bootstrap strategy to build confidence intervals.
Shapley effects

If $Y$ is scalar.

Shapley effect [Owen, 2014] (coopetative game theory [Shapley, 1953]) of $i$:

$$Sh_i = \frac{1}{d \text{Var}(Y)} \sum_{u \subseteq -\{i\}} (\frac{d-1}{|u|})^{-1} \left( \text{Var}(\mathbb{E}(Y|X_u \cup i)) - \text{Var}(\mathbb{E}(Y|X_u)) \right).$$
If $Y$ is scalar,

Shapley effect \cite{Owen_2014} (coopetative game theory \cite{Shapley_1953}) of $i$:

$$Sh_i = \frac{1}{d \cdot \text{Var}(Y)} \sum_{u \subseteq \{i\}} \binom{d-1}{|u|}^{-1} \left( \text{Var}(E(Y|X_{u U i})) - \text{Var}(E(Y|X_u)) \right).$$

**Inputs (coop. game players)**

- $X_1$
- $X_2$
- $X_d$
- $X_3$
- $X_i$

**Price**

- $\text{Var}(Y)$

**Shapley value**

- $\text{Var}(Y)$
Shapley effects

If $Y$ is scalar.

Shapley effect [Owen, 2014] (coopetative game theory [Shapley, 1953]) of $i$:

$$Sh_i = \frac{1}{d \cdot \text{Var}(Y)} \sum_{u \subseteq \{1, \ldots, d\} \setminus \{i\}} \left( \frac{d-1}{|u|} \right)^{d-1} (\text{Var}(E(Y|X_{u \cup i})) - \text{Var}(E(Y|X_u))).$$

Inputs (coop. game players)

$X_1, X_2, X_3, X_d, X_i$

Price

$\text{Var}(Y)$

Shapley value

$\text{Var}(Y)$

[Shapley, 1953] proved that this is the fairest way to divide a price among players (efficiency, symmetry, dummy, additive).
Shapley effects

If $Y$ is scalar.

The Shapley effect \cite{Owen_2014} (cooperative game theory \cite{Shapley_1953}) of $i$:

$$Sh_i = \frac{1}{d \Var(Y)} \sum_{u \subseteq \{-i\}} \left(\frac{d-1}{|u|}\right)^{-1} \left(\Var(\mathbb{E}(Y|X_{u \cup \{i\}})) - \Var(\mathbb{E}(Y|X_u))\right).$$

Inputs (coop. game players)

- $X_1$
- $X_2$
- $X_i$
- $X_d$
- $X_3$

Price

- $\Var(Y)$

Shapley value

- $\Var(Y)$

\cite{Shapley_1953} proved that this is the fairest way to divide a price among players (efficiency, symmetry, dummy, additive).

Shapley effect properties:

- $0 \leq Sh_i \leq 1$ for all $i \in \{1, \ldots, d\}$,
- $\sum_{i=1}^{d} Sh_i = 1$. 
Aggregated Shapley effects

If output is multivariate or the discretization of a functional output $Y = (Y_1, \ldots, Y_p)$.

Aggregated Shapley effects of input $X_i$:

$$GSh_i = \frac{\sum_{j=1}^{p} \text{Var}(Y_j) Sh_{i,j}}{\sum_{j=1}^{p} \text{Var}(Y_j)},$$

Aggregated Shapley effects accomplish the natural requirements for a sensitivity measure [Heredia et al., 2020]:

- $0 \leq GSh_i \leq 1$,
- $GSh_i(\lambda f(X))) = GSh_i(f(X))$ for all $\lambda \in \mathbb{R}$,
- $GSh_i(Of(X)) = GSh_i(f(X)))$ for all $O \in \mathbb{R}^{p \times p}$ and $O^t O = I$.

If the output dimension $p >> 1$, dimension reduction techniques such as pca, fpca [Yao et al., 2005] [Ramsay and Silverman, 2005] should be performed.
Estimation using nearest neighbors

For all $1 \leq i \leq d$ and all $1 \leq j \leq p$ to estimate $S_{j|i}$ and $G_{Sh_i}$, we need to estimate

$$\text{Var}(\mathbb{E}(Y_j|X_u)) \quad \text{or} \quad \mathbb{E}(\text{Var}(Y_j|X_{-u}))$$

for all $u \subseteq \{1, \ldots, d\}$, with $-u = \{1, \ldots, d\} \setminus u$.

In our context, we have to estimate from the given data $(X, Y)$ obtained from the AR sampling.
For all $1 \leq i \leq d$ and all $1 \leq j \leq p$ to estimate $Sh^i_j$ and $GSh^i$, we need to estimate

$$\text{Var}(\mathbb{E}(Y_j|X_u)) \quad \text{or} \quad \mathbb{E}(\text{Var}(Y_j|X_{-u}))$$

for all $u \subseteq \{1, \ldots, d\}$, with $-u = \{1, \ldots, d\} \setminus u$.

In our context, we have to estimate from the given data $(X, Y)$ obtained from the AR sampling.

[Broto et al., 2020] proposed to estimate $E_u = \mathbb{E}(\text{Var}(Y_j|X_{-u}))$ using nearest-neighbors. The estimator $\hat{E}_u$ converges in probability to $E_u$ under mild assumptions (theorem 6.6 of [Broto et al., 2020]).

Combining what they call the subset $W$-aggregation procedure with the estimates $\hat{E}_u$, [Broto et al., 2020, proposition 6.12] propose a consistent estimator for each Shapley effect.
Adaptation to the estimation of both Shapley and aggregated Shapley effects, with the construction of bootstrap confidence intervals:

Inputs: (i) a \( n \) sample \((x, y)\), (ii) \( N_{\text{tot}} \) the estimation cost, (iii) \( 1 \leq N_u \leq n \), the cost for estimation of \( F_u \) \( (N_u \) depends on \( N_{\text{tot}} \) and can be chosen in order to minimize the variance of the estimation), (iv) a \( N_u \) random sample \((s_\ell)_{1 \leq \ell \leq N_u}\) from \([1, \ldots, n]\), (v) \( N_l \) number of neighbors.
Adaptation to the estimation of both Shapley and aggregated Shapley effects, with the construction of bootstrap confidence intervals:

Inputs: (i) a \( n \) sample \((x, y)\), (ii) \( N_{\text{tot}} \) the estimation cost, (iii) \( 1 \leq N_u \leq n \), the cost for estimation of \( E_{u} \) (\( N_u \) depends on \( N_{\text{tot}} \) and can be chosen in order to minimize the variance of the estimation), (iv) a \( N_u \) random sample \((s_\ell)_{1 \leq \ell \leq N_u} \) from \( \{1, \ldots, n\} \), (v) \( N_I \) number of neighbors.

1. For all \( u \subset \{1, \ldots, d\} \) and for all \( 1 \leq \ell \leq N_u \), compute:

\[
\hat{E}_{u,s_\ell}^j = \frac{1}{N_I - 1} \sum_{v,v' \in \mathcal{B}_{-u,\ell}} \left( y_j^v - \frac{1}{N_I} \bar{y}_{s_\ell} \right)^2 \quad \text{with} \quad \bar{y}_{s_\ell} = \frac{1}{N_I} \sum_{u,v,v' \in \mathcal{B}_{u,\ell}} y_j^v
\]

with \( \mathcal{B}_{-u,\ell} \) the set of \( N_I \) closest neighbors of \( x_{s_\ell}^u \) where
\( x_{s_\ell}^u = (x_{w_1}^{s_\ell}, \ldots, x_{w_k}^{s_\ell}) \) with \( -u = \{w_1, \ldots, w_k\} \) and \( k = |{-u}| \).
Adaptation to the estimation of both Shapley and aggregated Shapley effects, with the construction of bootstrap confidence intervals:

Inputs: (i) a sample \((x, y)\), (ii) \(N_{\text{tot}}\) the estimation cost, (iii) \(1 \leq N_u \leq n\), the cost for estimation of \(E_u\) (\(N_u\) depends on \(N_{\text{tot}}\) and can be chosen in order to minimize the variance of the estimation), (iv) a random sample \((s_\ell)_{1 \leq \ell \leq N_u}\) from \([1, \ldots, n]\), (v) \(N_I\) number of neighbors.

1. For all \(u \in \{1, \ldots, d\}\) and for all \(1 \leq \ell \leq N_u\), compute:

\[
\hat{E}_{u,s_\ell}^j = \frac{1}{N_I - 1} \sum_{v,x_\ell \in \mathcal{B}_{-u,\ell}} \left( y_j^\prime - \frac{1}{N_I} \bar{y}_{s_\ell} \right)^2 \quad \text{with} \quad \bar{y}_{s_\ell} = \frac{1}{N_I} \sum_{v,x_\ell \in \mathcal{B}_{u,\ell}} y_j^\prime
\]

with \(\mathcal{B}_{-u,\ell}\) the set of \(N_I\) closest neighbors of \(x_{-u}^{s_\ell}\) where \(x_{-u}^{s_\ell} = (x_{w_1}^{s_\ell}, \ldots, x_{w_k}^{s_\ell})\) with \(-u = \{w_1, \ldots, w_k\}\) and \(k = | -u |\).
2.1 Compute for all $u \subset \{1, \ldots, d\}$.

$$\hat{E}_u^j = \frac{1}{N_u} \sum_{\ell=1}^{N_u} \hat{E}_{u,s_\ell}^j. \tag{1}$$

2.2 Compute $B$ bootstrap samples (the idea of block-bootstrap is adapted from [Benoumechiara and Elie-Dit-Cosaque, 2019]) from (1):

2.2.1 Create $N_u$ bootstrap samples from $\hat{E}_{u,s_\ell}^j$ by sampling with replacement from $\left(\hat{E}_{u,s_\ell}^j\right)_{1 \leq \ell \leq N_u}$.

2.2.2 Compute for all $b \in \{1, \ldots, B\}$:

$$\hat{E}_u^{j,(b)} = \frac{1}{N_u} \sum_{\ell=1}^{N_u} \hat{E}_{u,s_\ell}^{j,(b)}. \tag{2}$$
3.1. Compute $\tilde{Sh}_i^j$ for all $j \in \{1, \ldots, p\}$ according to:

$$
\tilde{Sh}_i^j = \frac{1}{d\hat{\sigma}_j^2} \sum_{u \subseteq -i} \left( \frac{d-1}{|u|} \right)^{-1} \left( \hat{E}_{u \cup \{i\}}^j - \hat{E}_u^j \right),
$$

(3)

where $\hat{\sigma}_j^2$ is the empirical variance of $y_j$.

3.2 Compute $B$ bootstrap samples of $\tilde{Sh}_i^j$ using (2) in (3):

$$
\tilde{Sh}_i^{j,(b)} = \frac{1}{d\hat{\sigma}_{j}^{2(b)}} \sum_{u \subseteq -i} \left( \frac{d-1}{|u|} \right)^{-1} \left( \hat{E}_{u \cup \{i\}}^{j,(b)} - \hat{E}_u^{j,(b)} \right),
$$

where $\hat{\sigma}_j^{2(b)}$ is the empirical variance of a bootstrap sample of $y_j$. 
4.1 Compute $\hat{GSh}_i$ for all $i \in \{1, \ldots, d\}$ according to:

$$\hat{GSh}_i = \frac{1}{d \sum_{j=1}^{p} \hat{\sigma}_j^2} \sum_{j=1}^{p} \sum_{u \subset -i} \left( d - 1 \right)^{-1} \left( \hat{E}_{j, u \cup \{i\}} - \hat{E}_{u} \right),$$

4.2 compute $B$ bootstrap samples of $\hat{Gh}_i$:

$$\hat{GSh}_i^{(b)} = \frac{1}{d \sum_{j=1}^{p} \hat{\sigma}_j^{2,(b)}} \sum_{j=1}^{p} \sum_{u \subset -i} \left( d - 1 \right)^{-1} \left( \hat{E}_{j, u \cup \{i\}}^{(b)} - \hat{E}_{u}^{(b)} \right).$$
4.1 Compute $\widehat{GSh}_i$ for all $i \in \{1, \ldots, d\}$ according to:

$$
\widehat{GSh}_i = \frac{1}{d \sum_{j=1}^{p} \hat{\sigma}_{j}^{2}} \sum_{j=1}^{p} \sum_{u \subsetneq i} \left( d - 1 \right)^{-1} \left( \hat{E}_{u \cup \{i\}} - \hat{E}_{u} \right),
$$

4.2 Compute $B$ bootstrap samples of $\widehat{Gh}_i$:

$$
\widehat{GSh}_{i}^{(b)} = \frac{1}{d \sum_{j=1}^{p} \hat{\sigma}_{j}^{2(b)}} \sum_{j=1}^{p} \sum_{u \subsetneq i} \left( d - 1 \right)^{-1} \left( \hat{E}_{u \cup \{i\}}^{(b)} - \hat{E}_{u}^{(b)} \right).
$$

5 Compute simultaneous bootstrap confidence intervals (correction of Bonferroni) with bias correction (see e.g., [Efron, 1981]).
Linear Gaussian model with two inputs

Model from [Iooss and Prieur, 2019].

\[ Y = \beta_0 + \beta^t X \]

with \( X_i \sim \mathcal{N}(0, 1) \), \( \beta_1 = 1 \), \( \beta_2 = 0 \), \( X_1 \) and \( X_2 \) correlated \( \rho = 0.4 \).

Figure: Mean absolute error of the estimation of scalar Shapley effects in \( N=300 \) i.i.d. samples in function of \( N_{tot} \). \( N_1 = 3 \). 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 (Bonferroni correction) is displayed with dotted lines. The theoretical probability of coverage 0.9 is shown with a plain gray line. The bootstrap sample size is fixed to \( B = 500 \).
Multivariate Linear Gaussian model with two inputs

\[ Y = (Y_1, Y_2, Y_3) = \beta_0 + \beta^t X \]

with \( X_i \sim \mathcal{N}(0, 1) \), \( X_1 \) and \( X_2 \) correlated \( \rho = 0.4 \), and \( \beta \in \mathbb{R}^{2 \times 3} \):

\[ \beta = \begin{bmatrix} 1 & 4 & 0.1 \\ 1 & 3 & 0.9 \end{bmatrix}. \]

**Figure:** Mean absolute error of the estimation of aggregated Shapley effects in \( N=300 \) i.i.d. samples in function of \( N_{tot} \). \( N_I \). 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 (Bonferroni correction) is displayed with dotted lines. The theoretical probability of coverage 0.9 is shown with a plain gray line.
**Input** | **Description** | **Distribution**
--- | --- | ---
$\mu$ | Static friction coefficient | $\mathcal{U}[0.05, 0.65]$ |
$\xi$ | Turbulent friction [m/s²] | $\mathcal{U}[400, 10000]$ |
$l_{start}$ | Length of the release zone [m] | $\mathcal{U}[5,300]$ |
$h_{start}$ | Mean snow depth in the release zone [m] | $\mathcal{U}[0.05, 3]$ |
$x_{start}$ | Release abscissa [m] | $\mathcal{U}[0, 1600]$ |

We consider $v_{ol, start} = l_{start} \times h_{start} \times 72.3 / \cos(35^\circ)$ instead of $h_{start}$ and $l_{start}$.

**AR rules:**

- avalanche simulation is flowing in $[1600m, 2412m]$,
- $vol > 7000m^3$,
- runout distance $< 2500m$ (end of the path).

From $n_0 = 100000$, AR sample size $n_1 = 6152$. 
**AR rules:**
- Avalanche simulation is flowing in $[1600 \text{m}, 2412 \text{m}]$,
- $vol > 7000 \text{m}^3$,
- Runout distance $< 2500 \text{m}$ (end of the path).

From $n_0 = 100000$, AR sample size $n_1 = 6152$. 
Ubiquitous Shapley effects

Figure: Shapley effects are estimated with a sample of size 6152 and Ntot=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 bootstrap sample size is fixed to B = 500.
Figure: **Aggregated Shapley effects** are estimated with a sample of size 6152 and Ntot=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 10000 years. The bootstrap sample size is fixed to B = 500.
Conclusions

- We extended Shapley effects to models with multivariate or functional outputs.
- We proposed an algorithm to construct bootstrap confidence intervals for estimation.
- The bootstrap confidence intervals have accurate coverage probability.
- Aggregated Shapley effects are more stable and easier to interpret (observed by [Alexanderian et al., 2020] for Sobol’ indices).

Perspectives

- In order to estimate with samples of higher size, build a surrogate model of our avalanche model.
- To perform a GSA in several corridors in order to see if there exist correlations between the local slope and the ubiquitous effects.
- To study theoretically the asymptotic properties of our estimator.
Thanks! Questions?
References I

Variance-based sensitivity analysis for time-dependent processes.

Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms.

A new uncertainty importance measure.

Variance Reduction for Estimation of Shapley Effects and Adaptation to Unknown Input Distribution.

Campbell, K., McKay, M. D., and Williams, B. J. (2006).
Sensitivity analysis when model outputs are functions.
The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004).

An accelerated-time model for response curves.

Global sensitivity analysis with dependence measures.
Nonparametric standard errors and confidence intervals. 

*Local polynomial modelling and its applications.* 

Sensitivity indices for multivariate outputs. 

Aggregated Shapley effects: nearest-neighbor estimation procedure and confidence intervals. Application to snow avalanche modeling. working paper or preprint.

Shapley effects for sensitivity analysis with correlated inputs: Comparisons with Sobol’ indices, numerical estimation and applications. 

Derivative-based global sensitivity measures: General links with Sobolâ€™ indices and numerical tests. 

Multivariate global sensitivity analysis for dynamic crop models. 
References

Dense snow avalanche modeling: flow, erosion, deposition and obstacle effects.
Snow And Avalanches: Papers Presented At The European Geophysical Union Conference, Nice, April 2003. Dedicated To The Avalanche Dynamics Pioneer Dr. B. Salm.

Sobol’ Indices and Shapley Value.

On Shapley value for measuring importance of dependent inputs.
*SIAM/ASA Journal on Uncertainty Quantification*, 5(1).

*Functional Data Analysis.*

Estimating the mean and covariance structure nonparametrically when the data are curves.

Shapley, L. (1953).
A Value for n-Person Games.

Derivative based global sensitivity measures and their link with global sensitivity indices.
Sensitivity analysis for non-linear mathematical models.  

Functional data analysis for sparse longitudinal data.  
Appendix
Shapley value [Shapley, 1953]

Given a set of \( d \) players in a coalitional game and a characteristic function \( \text{val} : 2^d \rightarrow \mathbb{R}, \text{val}(\emptyset) = 0 \), 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 \( \text{val}(u \cup \{i\}) = \text{val}(u \cup \{\ell\}) \) for all \( u \subseteq \{1, \ldots, d\} - \{i, j\} \), then \( \phi_i = \phi_\ell. \)

3. **(Dummy)** If \( \text{val}(u \cup \{i\}) = \text{val}(u) \) for all \( u \subseteq \{1, \ldots, d\} \), then \( \phi_i = 0. \)

4. **(Additivity)** If \( \text{val} \) and \( \text{val}' \) have Shapley values \( \phi \) and \( \phi' \) respectively, then the game with characteristic function \( \text{val} + \text{val}' \) has Shapley value \( \phi_i + \phi'_i \) for \( i \in \{1, \ldots, d\} \).

It is proved in [Shapley, 1953] that according to the Shapley value, the amount that player \( i \) gets given a coalitional game \( (\text{val}, d) \) is:

\[
\phi_i = \frac{1}{d} \sum_{u \subseteq \{1, \ldots, d\} - \{i\}} \left( \frac{d-1}{|u|} \right)^{-1} (\text{val}(u \cup \{i\}) - \text{val}(u)) \quad \forall \ i \in \{1, \ldots, d\}.
\]
Functional principal component analysis [Yao et al., 2005]

We have a collection of $n$ independent trajectories of a smooth random function $f(.,X)$ with unknown mean $\mu(s) = \mathbb{E}(f(s,X)), s \in \tau$, where $\tau$ is a bounded and closed interval in $\mathbb{R}$, and covariance function:

$$G(s_1, s_2) = \text{Cov}(f(s_1, X), f(s_2, X)), s_1, s_2 \in \tau.$$ 

We assume that $G$ has a $L^2$ orthogonal expansion in terms of eigenfunction $\xi_k$ and non increasing eigenvalues $\lambda_k$ such that:

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

The Karhunen-Loève orthogonal expansion of $f(s,X)$ is:

$$f(s, X) = \mu(s) + \sum_{k \geq 1} \alpha_k(X) \xi_k(s) \approx \mu(s) + \sum_{k=1}^{q} \alpha_k(X) \xi_k(s), s \in \tau, \quad (4)$$

where $\alpha_k(X) = \int_{\tau} f(s, X) \xi_k(s) \, ds$ is the $k$-th functional principal component (fPC) and $q$ is a truncation level.

For fPCs estimation, the authors in [Yao et al., 2005] propose first to estimate $\hat{\mu}(s)$ using local linear smoothers and to estimate $\hat{G}(s_1, s_2)$ using local linear surface smoothers ([Fan and Gijbels, 1996]).
The estimates of eigenfunctions and eigenvalues correspond then to the solutions of the following integral equations:

\[ \int_\tau \hat{G}(s_1, s) \hat{\xi}_k(s_1) ds_1 = \lambda_k \hat{\xi}_k(s), s \in \tau, \]

with \( \int_\tau \hat{\xi}(s) ds = 1 \) and \( \int_\tau \hat{\xi}_k(s) \hat{\xi}_m(s) = 0 \) for all \( m \neq k \leq q \). The problem is solved by using a discretization of the smoothed covariance (see further details in [Rice and Silverman, 1991] and [Capra and Müller, 1997]). Finally, fPCs \( \hat{\alpha}_k(X) = \int_\tau f(s, X) \hat{\xi}_k(s) ds \) are solved by numerical integration.

Aggregated Shapley effects are computed with only the \( q \) first fPCs:

\[
\tilde{G}Sh_i = \frac{1}{d \sum_{k=1}^q \lambda_k} \sum_{k=1}^q \sum_{u \subseteq -i} \left( d - 1 \right)^{-1} \left( \mathbb{E}(\text{Var}(\alpha_k(X)|X_{u \cup \{i\}})) - \mathbb{E}(\text{Var}(\alpha_k(X)|X_u)) \right).
\]

(5)
Theorem (Theorem 6.6 [Broto et al., 2020])

If \( f \) is bounded, the \( \hat{E}_u \) converges to \( E_u \) in probability when \( n \) and \( N_u \) if:

- For all \( i \in \{1, \ldots, d\} \), \( (\mathcal{X}_i, d_i) \) is a Polish space with metric \( d_i \) with \( \mathcal{X}_i \) the domain of \( X_i \), and \( X = (X_1, \ldots, X_d) \) has a density \( f_X \) with respect to a finite measure \( \mu = \bigotimes_{i=1}^{d} \mu_i \) which is bounded and \( \mathbb{P}_X \) almost everywhere continuous.

- The closest neighbors in \( B_{-u, \ell} \) are two by two distinct.
The bias-corrected percentile method [Efron, 1981]
Given bootstrap samples $B$ of $\hat{G}_{Sh_i}$, $\mathcal{R}_i = \{\hat{G}_{Sh_i}^{(1)}, \ldots, \hat{G}_{Sh_i}^{(B)}\}$.
We compute a bias correction constant $z_0$:

$$
\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{G}_{Sh_i}^{(b)} \in \mathcal{R}_i \text{ s. t. } \hat{G}_{Sh_i}^{(b)} \leq \hat{G}_{Sh_i}\}}{B}\right)
$$

where $\Phi$ the standard normal cumulative distribution function.
The corrected quantile estimate $\hat{q}(\beta)$:

$$
\hat{q}_i(\beta) = \Phi(2\hat{z}_0 + z_{\beta}),
$$

where $z_{\beta}$ satisfies $\Phi(z_{\beta}) = \beta$.
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 $g(\hat{G}_{Sh_i}) \sim \mathcal{N}(G_{Sh_i} - \tau z_0, \tau^2)$ and $g(\hat{G}_{Sh_i}^*) \sim \mathcal{N}(\hat{G}_{Sh_i} - \tau z_0, \tau^2)$ where $\hat{G}_{Sh_i}^*$ is the bootstrapped $\hat{G}_{Sh_i}$ for fixed sample (see [Efron, 1981]).
Probability of coverage with Bonferroni correction
The probability of coverage with Bonferroni correction is the probability that \([\hat{q}_i(\alpha/(2d)), \hat{q}_i(1 - \alpha/(2d))]\) contains \(GSh_i\) for all \(i \in \{1, \ldots, d\}\) simultaneously.

The POC is estimated as

\[
\hat{\text{POC}} = \frac{1}{N} \sum_{k=1}^{N} w^k,
\]

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.