



Analyse de sensibilité avec entrées dépendantes : cas linéaire, effets de Shapley et extensions récentes

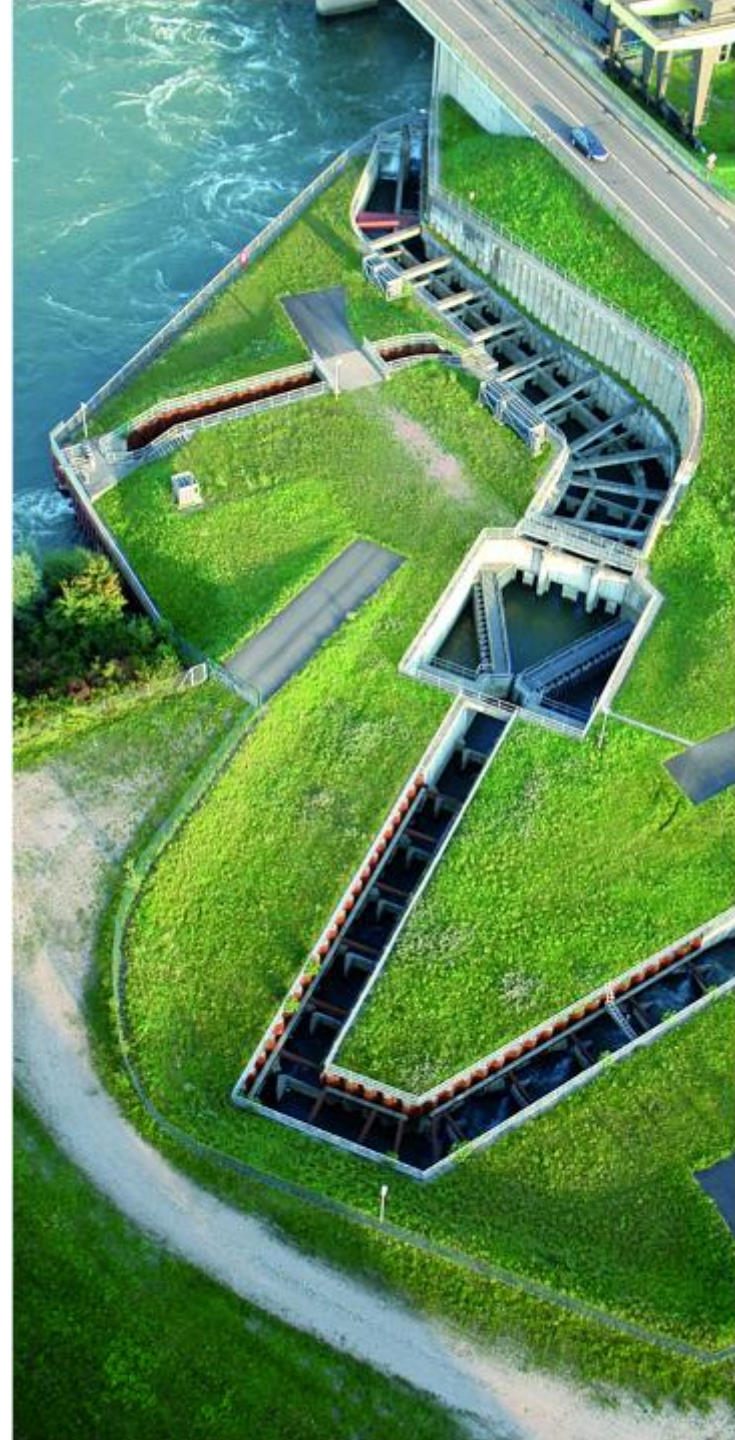
Bertrand IOOSS, *EDF R&D, Chatou, France*

SINCLAIR AI Lab Saclay, France

avec Vincent Chabridon, Laura Clouvel, Margot Héryn,
Marouane Il Idrissi (EDF R&D)

et Clémentine Prieur (Université Grenoble-Alpes)

Rencontres MEXICO 2021, INRAE Toulouse, 29/11/2021

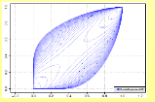



Étape C : Propagation des sources d'incertitude



Étape B: Quantification des sources d'incertitudes

Modélisation par des distributions

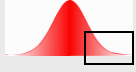


Étape A : Spécification du problème

Variables d'entrée
Incertaines : X
(de dimension d)

Modèle numérique
(ou processus de mesure)
 $G(X)$

Variables d'intérêt
 $Y = G(X)$

Quantité d'intérêt
Ex: variance, probabilité ..


Étape C' : Analyse de sensibilité, Analyse de robustesse



Révision des modèles

Critère de décision
Ex: Probabilité $< 10^{-b}$



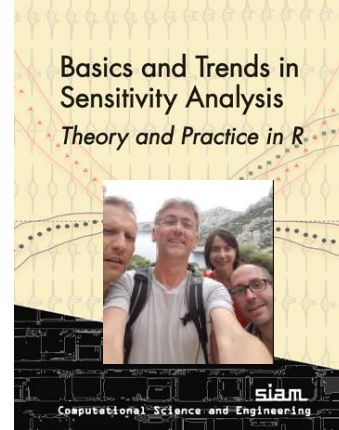


INTRODUCTION – RAPPELS

ANALYSE DE SENSIBILITÉ AVEC ENTRÉES INDÉPENDANTES

Settings of global sensitivity analysis (SA)

SÉBASTIEN DA VEIGA • FABRICE GAMBOA
BERTRAND IOOSS • CLÉMENTINE PRIEUR



1. **Understand the behaviour of the model output wrt inputs** (direction of variation, interaction, discontinuity, ...)
=> Visualization techniques, Metamodeling, ...

2. **Simplify the study (dimension reduction)** **Screening**

Determine the non-influential inputs (that can be fixed)

=> Design of experiments, Morris & derivative-based methods

=> Given-data method from machine learning (e.g., HSIC)



3. **Prioritize the uncertainty sources to reduce the output uncertainty**

Inputs whose uncertainty must be reduced to mostly decrease the **output uncert.**

=> Best routes:

- Sobol' indices: $S_i = \frac{\text{Var}[E(Y|X_i)]}{\text{Var}(Y)}$, $S_{ij} = \frac{\text{Var}[E(Y|X_i X_j)]}{\text{Var}(Y)} - S_i - S_j, \dots$

Quantitative partitioning

- Easy interpretation of total Sobol' indices: $S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + \dots$

- Shapley effects in case of dependent inputs



4. **Analyze the robustness of the quantity of interest wrt the uncertainty on the input uncertainty models**

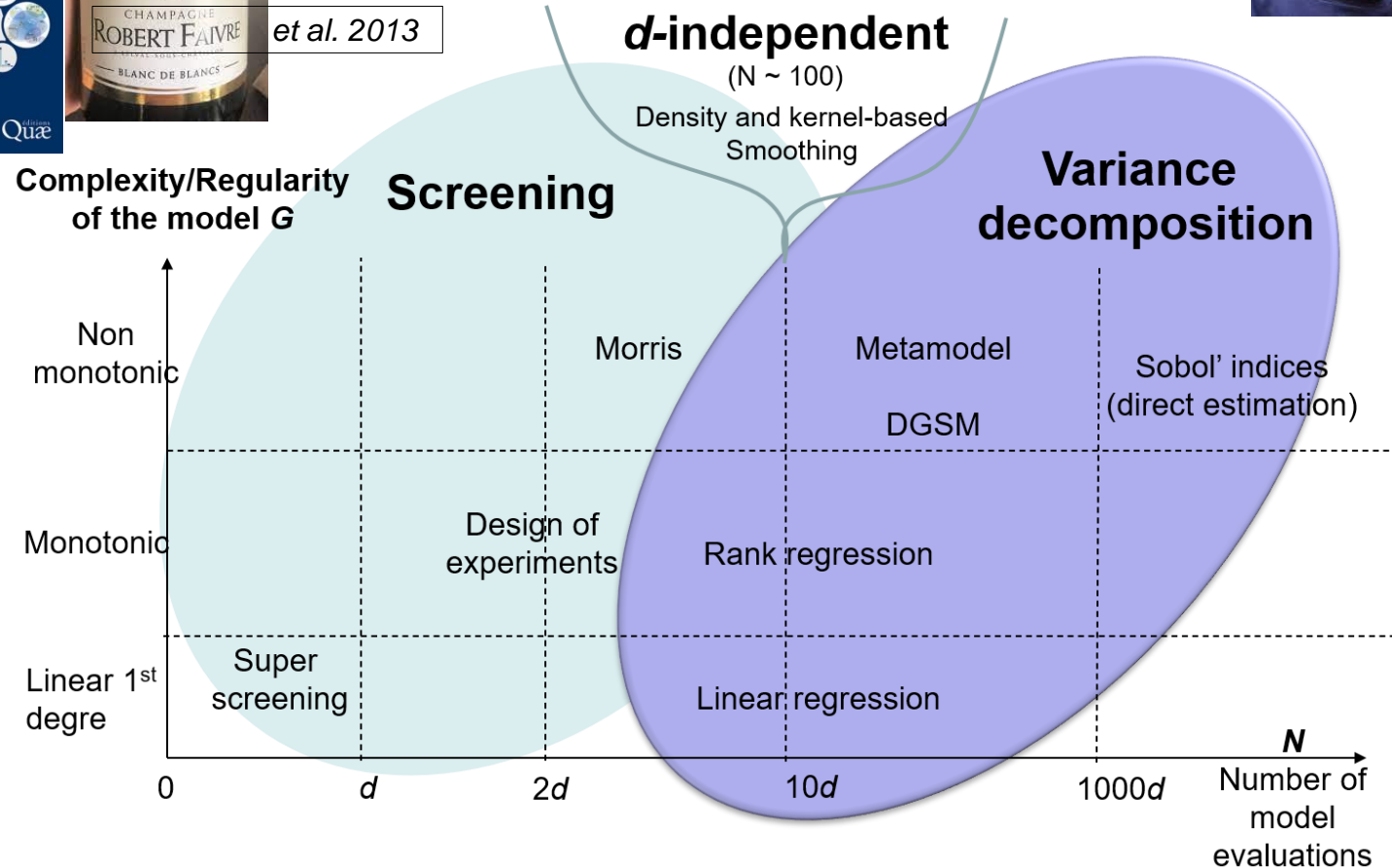
Robustness analysis

The MEXICO-user point of view!

Global methodology, inspired from the papers of JRC Ispra (Andrea Saltelli, ...), Tilburg Univ. (Jack Kleijnen, ...) and Sandia Nat. Lab. (Jon Helton, ...)



et al. 2013



This kind of classification grids greatly help the end-users to navigate inside the jungle of different methods, each better than the others (wrt authors' opinion), **then to choose the methods most adapted to their specific problem**

Functional ANOVA:

(strong hypothesis: independence of the X_j)

$$\text{Var}(Y) = \sum_{i=1}^d V_i(Y) + \sum_{i<j}^d V_{ij}(Y) + \dots + V_{12\dots d}(Y)$$

$$V_i = \text{Var}_{X_i}[E(Y|X_i)]$$

$$S_i = \frac{V_i}{\text{Var}(Y)}; \quad S_{ij} \quad ; \quad \dots \quad ; \quad T_i = S_i + \sum_j S_{ij} + \sum_{j,k} S_{ijk} + \dots$$

1st order

2nd order

Total index

$$T_i = E[\text{Var}(Y|X_{-i})]/\text{Var}(Y)$$

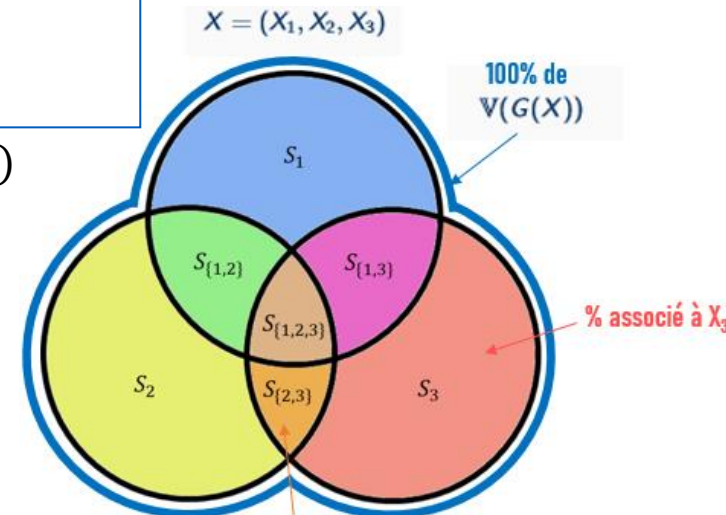
Important properties:

$$0 \leq S_i \leq T_i \leq 1$$

$$\sum_{u \subseteq \{1, \dots, d\}} S_u = 1$$

« Good » estimation methods:

- For S_i and S_{ij} : Replicated OA-LHS - Cost: $N = 2n$ (in practice: $n \sim 1000$)
- For S_i and S_{T_i} : Pick-freeze Monte Carlo - Cost: $N = n(d+2)$
- Quasi Monte Carlo, FAST and extensions: Potential reduction of n
- Approximation of the computer code by a metamodel (building cost $N \sim 10d$)
Then: Monte Carlo algo. applied on the metamodel (e.g. kriging)





ANALYSE DE SENSIBILITÉ AVEC ENTRÉES DÉPENDANTES

CAS LINÉAIRE

Rappels - Cumul quadratique – Développement d'ordre 1

Calcul de la moyenne de Y :

$$E(Y) = G(\mu) \quad \text{où } \mu \text{ est la moyenne de } \mathbf{X}$$

La moyenne de la réponse est égale, au premier ordre, à la réponse calculée aux valeurs moyennes des entrées

Calcul de la variance de Y :

$$\begin{aligned} \text{Var}(Y) &= E(Y - E(Y))^2 \\ &= E \left(G(\mu) + \sum_{i=1}^d \frac{\partial G}{\partial X_i} \Big|_{X=\mu} (X_i - \mu_i) - G(\mu) \right)^2 \\ &= \sum_{i=1}^d \sum_{j=1}^d \frac{\partial G}{\partial X_i} \Big|_{X=\mu} \frac{\partial G}{\partial X_j} \Big|_{X=\mu} E[(X_i - \mu_i)(X_j - \mu_j)] \end{aligned}$$

où μ_i est la moyenne de X_i

$$\text{Var}(Y) = \sum_{i=1}^d \sum_{j=1}^d \frac{\partial G}{\partial X_i} \Big|_{X=\mu} \frac{\partial G}{\partial X_j} \Big|_{X=\mu} \rho_{ij} \sigma_i \sigma_j$$

où σ_i est l'écart type de X_i
et ρ_{ij} est le coef. de corrélation entre X_i et X_j

Remarques :

++ Ne nécessite que moyenne et covariance de \mathbf{X}

-- A ne pas utiliser pour les modèles $G(\cdot)$ fortement non linéaires

Rappels - Cumul quadratique – Variables indépendantes

Calcul de la variance si les X_i sont indépendantes :

$$\text{Var}(Y) = \sum_{i=1}^d \underbrace{\left(\frac{\partial G}{\partial X_i} \Big|_{X=\mu} \right)^2}_{\text{Contribution de chaque variable d'entrée à l'incertitude sur la variable de sortie}} \sigma_i^2 \quad \text{où } \sigma_i^2 \text{ est la variance de } X_i$$

Contribution de chaque variable d'entrée à l'incertitude sur la variable de sortie



- Termes « déterministes » → composants du gradient de $G(\bullet)$
- Termes liés à l'incertitude de la variable X_i (variance)

Indices de sensibilité qui en découlent :

$$\eta_i^2 = \frac{1}{\text{Var}(Y)} \left(\frac{\partial G}{\partial X_i} \Big|_{X=\mu} \right)^2 \sigma_i^2$$

où $\text{Var}(Y)$ est estimée par la formule du cumul quadratique

Cumul quadratique – Variables dépendantes

À partir de la décomposition, un indice (erronné), nommé CVD pour « covariance decomposition », est possible :

$$\eta_i^2 = \frac{1}{\text{Var}(Y)} \left[\left(\frac{\partial G}{\partial X_i} \Big|_{X=\mu} \right)^2 \sigma_i^2 + \sum_{\substack{j=1 \\ j \neq i}}^d \frac{\partial G}{\partial X_i} \Big|_{X=\mu} \frac{\partial G}{\partial X_j} \Big|_{X=\mu} \rho_{ij} \sigma_i \sigma_j \right]$$

Remarques :

++ *Faciles à calculer*

++ *Les η_i^2 somment à la variance de Y*

-- **Peuvent être négatifs**, donc ne peuvent être interprétés comme un pourcentage de contribution à la variance (il n'y a pas de part négative dans un camembert)

Exemple : $Y = X_1 + 2 X_2$ avec $X_1 \sim \mathcal{N}(0;1)$, $X_2 \sim \mathcal{N}(0;1)$ et $\rho(X_1, X_2) = -0,9$

On a : $V = \text{Var}(Y) = 1,4$; $\eta_1^2 = -0,8 / V$; $\eta_2^2 = 2,2 / V$

Mesure d'importance pour modèle linéaire

Échantillon d'apprentissage (par exemple Monte Carlo) : N réalisations de (\mathbf{X}, Y)

Modèle linéaire pour la sortie : $Y = \beta_0 + \sum_{i=1}^d \beta_i X_i$; prédicteur : $\hat{Y}(\mathbf{x}) = \hat{\beta}_0 + \sum_{i=1}^d \hat{\beta}_i x_i$

Valider le modèle linéaire : $R_{Y(\mathbf{X})}^2 = 1 - \frac{\sum_{k=1}^N (\hat{Y}_k - Y_k)^2}{\sum_{k=1}^N (Y_k - \bar{Y})^2}$; diagnostics (résidus); Q^2 (R^2 sur base de test)

Importance Measure [Johnson & LeBreton 2004]: "The proportionate contribution each variable makes to R^2 , considering both its direct effect and its effect when combined with other variables"

2 propriétés « phares » à respecter pour les indices : 1) somment au R^2 , 2) non négatifs

[Grömping 2007]

Cas indépendant : « Standard Regression Coef. » $\text{SRC}(X_i) = \beta_i \sqrt{\frac{\text{Var}(X_i)}{\text{Var}(Y)}}$ ($\text{SRC}(X_i)^2 = \eta_i^2$)

Somment au R^2 et si le modèle linéaire est exact : $\text{SRC}(X_i) = \rho(X_i, Y)$

Cas dépendant :

- Préconisation classique : « Partial Correlation Coef. » $\text{PCC}(X_i) = \rho(X_i - \hat{X}_{-i}, Y - \hat{Y}_{-i})$
-- *PCC n'est qu'un indice de linéarité entre Y et $X_i \Rightarrow$ **Cet indice ne doit pas être utilisé***
- Semi-PCC $\text{SPCC}(X_i) = \rho(X_i - \hat{X}_{-i}, Y) = r_{Y, (X_i | \mathbf{X}_{-i})}$
On a : $r_{Y, (X_i | \mathbf{X}_{-i})}^2 = R_{Y(\mathbf{X})}^2 - R_{Y(\mathbf{X}_{-i})}^2$ (pouvoir explicatif additionnel de X_i)
-- SPCC ne somment pas au R^2 et écrasent l'effet d'une variable très corrélée à d'autres

Modèle linéaire et entrées dépendantes

Multicollinéarité, mesurée par Variance Inflation Factor : $VIF(X_j) = 1/[1 - R_{X_j(x_{-j})}^2]$

⇒ indices LMG:

[Lindeman–Merenda–Gold 1980]

$$LMG_j = \frac{1}{d!} \sum_{\pi \in \text{permutations of } \{1, \dots, d\}} r_{Y, (X_j | X_\pi)}^2$$

$$r_{Y, (X_j | X_\pi)}^2 = R_{Y(X_{v \cup \{j\}})}^2 - R_{Y(X_v)}^2$$

où v sont les indices entrant avant j dans π

Exemple précédent : $Z = X_1 + 2 X_2$ avec $X_1 \sim \mathcal{N}(0;1)$, $X_2 \sim \mathcal{N}(0;1)$ et $\rho(X_1, X_2) = -0,9$

On a : $V = \text{Var}(Y) = 1,4$; $LMG_1 = 0,415 / V$; $LMG_2 = 0,985 / V$

$LMG_1 \sim 30\%$; $LMG_2 \sim 70\%$

⇒ PMVD (Proportional Marginal Variance Decomposition):

$$PMVD_j = \frac{\sum_{\pi \in \text{permutations of } \{1, \dots, d\}} L(\pi) r_{Y, (X_j | X_\pi)}^2}{\sum_{\pi} L(\pi)}$$

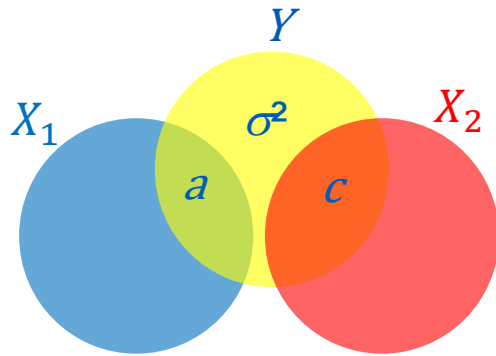
où $L(\pi)$ prend de grandes valeurs pour les permutations qui ont le bon ordre d'influence décroissante

$L(\pi)$ est construite en utilisant un axiome qui dit que si $SPCC_j=0$ alors $PMVD_j=0$

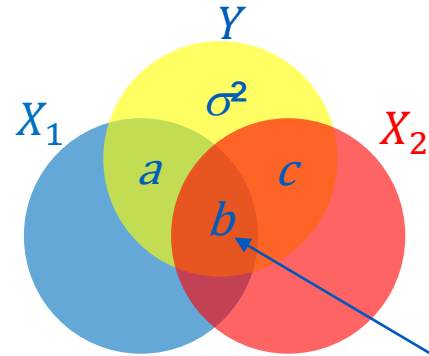
[Feldman 2000;
Grömping 2007;
Il Idrissi et al. 2021]

- LMG et PMVD sont basés sur une idée de **partage** des parts de variance dues à la corrélation entre variables (**équitable** pour LMG / **proportionnel** pour PMVD)
- **Leur somme vaut R^2 et ils sont non négatifs**
- PMVD respecte **la propriété d'exclusion** : attribue **un poids nul** à une variable qui a un coef. de régr. nul

Exemple analytique : 2 variables dépendantes



Cas indépendant



Cas dépendant

$$R_{Y(X_1, X_2)}^2 = \frac{a + b + c}{a + b + c + \sigma^2}$$

part due à la corrélation

Indices de sensibilité

Coefficient de corrélation au carré : $r_{Y, X_1}^2 = \frac{a + b}{a + b + c + \sigma^2}$ and $r_{Y, X_2}^2 = \frac{c + b}{a + b + c + \sigma^2}$.

PCC : $r_{(Y, X_1) | X_2}^2 = \frac{a}{a + \sigma^2}$ and $r_{(Y, X_2) | X_1}^2 = \frac{c}{c + \sigma^2}$.

SPCC : $r_{Y, (X_1 | X_2)}^2 = \frac{a}{a + b + c + \sigma^2}$ and $r_{Y, (X_2 | X_1)}^2 = \frac{c}{a + b + c + \sigma^2}$.

LMG : $LMG_1 = (a + b/2) / (a + b + c + \sigma^2)$
 $LMG_2 = (c + b/2) / (a + b + c + \sigma^2)$

PMVD : $PMVD_1 = a + b \times a / (a + c) / (a + b + c + \sigma^2)$
 $PMVD_2 = c + b \times c / (a + c) / (a + b + c + \sigma^2)$

Calcul des LMG et PMVD

- Les indices LMG et PMVD peuvent être calculés grâce à une combinatoire relativement complexe

cf. implémentation optimisée par Marouane El Idrissi dans package R « sensitivity » [I. et al. 2021]

- Problème majeur : indices trop coûteux à évaluer quand $d > 10$**

=> Alternatives à l'étude pour la grande dimension :

- 1) Si \mathbf{X} est (multi-)gaussien, on obtient une formule analytique (lourde) pour les LMG :

$$\text{LMG}_i = \frac{1}{d \text{Var}(Y)} \sum_{u \subseteq \{1, \dots, d\} \setminus \{i\}} \binom{d-1}{|u|} \left[E(\text{Var}(Y | \mathbf{X}_u)) - E(\text{Var}(Y | \mathbf{X}_{u \cup \{i\}})) \right]$$

Les termes $E(\text{Var}(Y | \mathbf{X}_u))$ font intervenir des blocs de la matrice de covariance de \mathbf{X}

[Broto et al. 2019]

- 2) Les indices de Johnson ou RWA (relative weight allocations), basés sur une simple SVD (et donc surmontant le fléau de la dimension), **approchent bien les LMG**

[Clouvel 2019]

Illustrations sur modèles jouets (1/3)

$$g(X) = X_1^2 + X_2 + X_3$$

$\mathcal{U}[0.5, 1.5]$ $\mathcal{U}[1.5, 4.5]$ $\mathcal{U}[4.5, 13.5]$

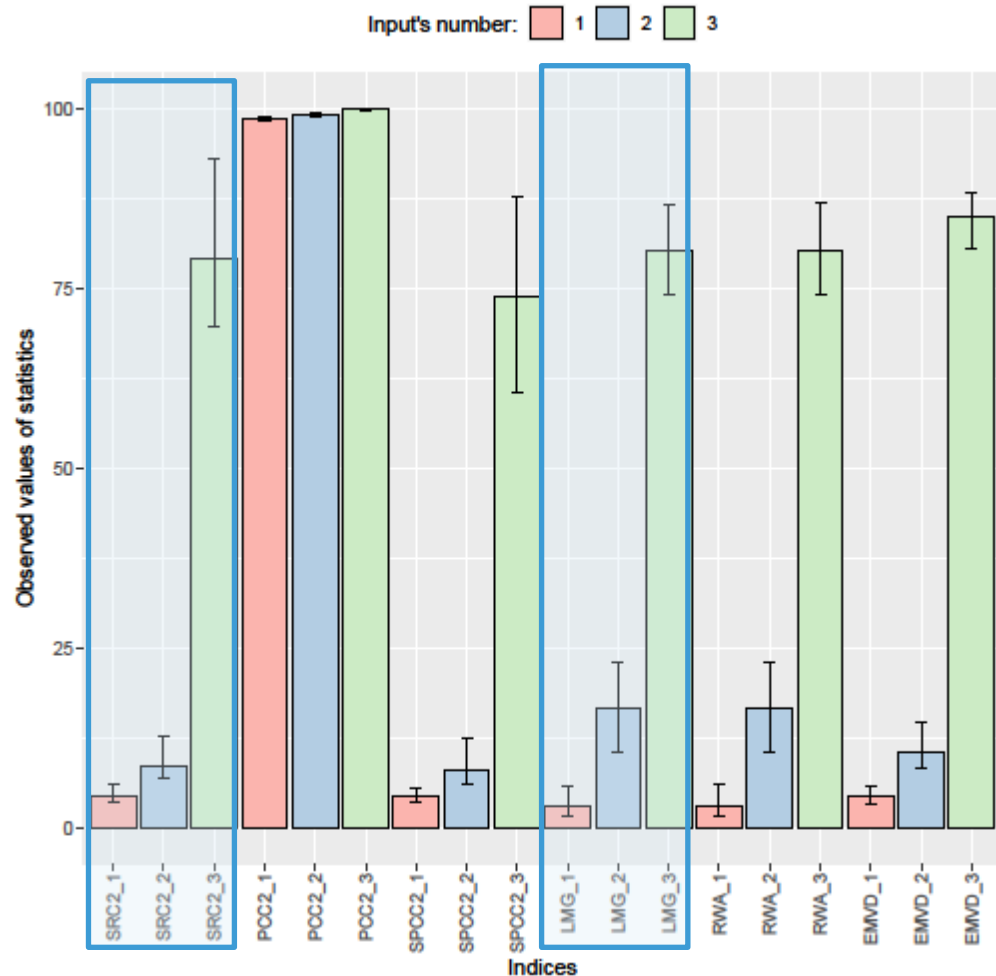
3 variables indépendantes

$$Y = g(X)$$

$$n = 100$$

Régression linéaire

$$\Rightarrow R^2 = Q^2 = 1$$



Illustrations sur modèles jouets (2/3)

$$g(X) = X_1^2 + X_2 + X_3$$

$\mathcal{U}[0.5, 1.5]$ $\mathcal{U}[1.5, 4.5]$ $\mathcal{U}[4.5, 13.5]$

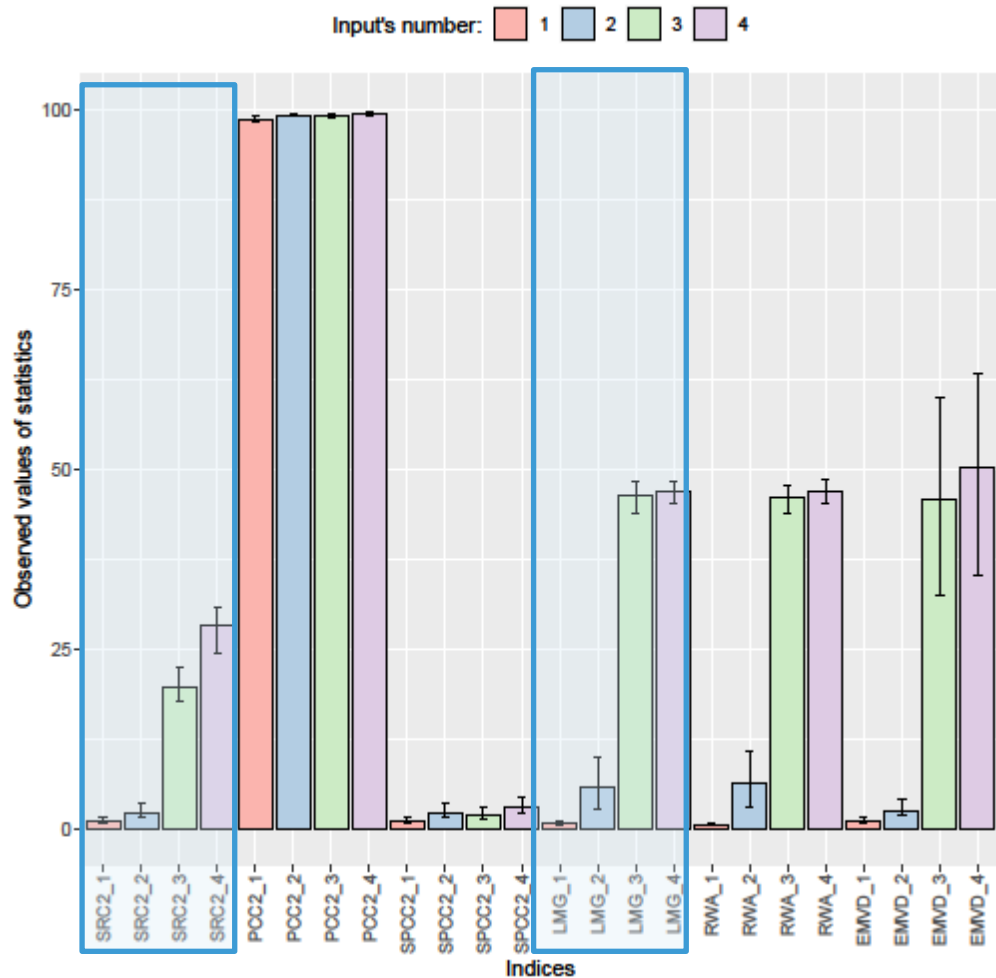
4 variables dont 2 corrélées

$$Y = g(X) + X_4$$

avec $X_4 = X_3 + \eta$
 $\eta \sim \mathcal{N}(0, 1)$

$n = 100$

Régression linéaire
 $\Rightarrow R^2 = Q^2 = 1$



Illustrations sur modèles jouets (3/3)

$$g(X) = X_1 + \eta$$

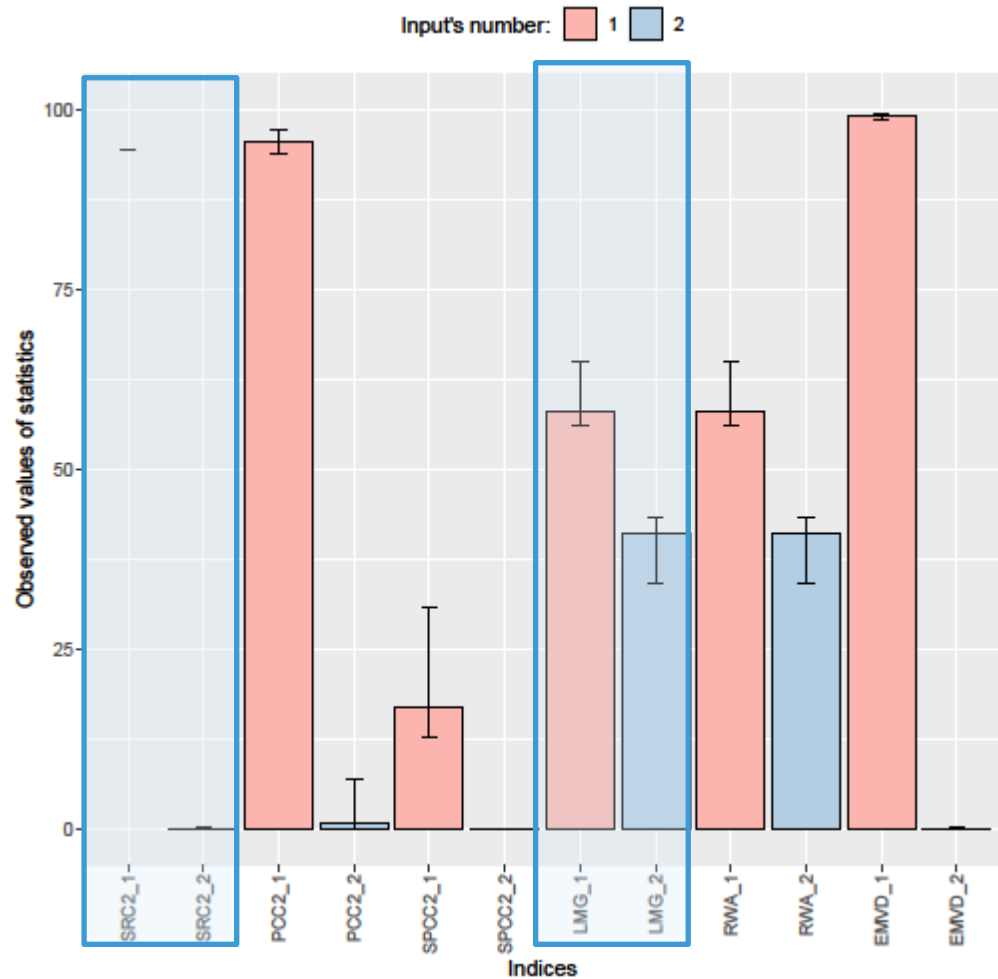
$\eta \sim \mathcal{N}(0, 0.01)$

2 variables très corrélées
(dont une « fausse » variable)

$$Y = g(X)$$

$$n = 100$$

Régression linéaire
 $\Rightarrow R^2 = Q^2 = 0,99$



Exemple sur jeu de données public (2/2)

$$R^2 = 0.74 \quad Q^2 = 0.72$$

VIF = Mesure de multicollinéarité

PMVD

E-MVD

	Input	n°	VIF	SRC ²	PCC ²	SPCC ²	LMG	RWA	E-MVD
	crim	1	1.79	1.09	2.28	0.61	2.79	3.29	0.72
	zn	2	2.30	1.51	2.46	0.66	2.50	2.81	0.67
	indus	3	3.95	0.10	0.10	0.03	3.74	3.66	0.06
	nox	4	4.39	4.79	4.02	1.09	3.31	3.68	1.54
nombre de chambres	rm	5	1.93	8.59	14.57	4.45	19.01	20.59	22.71
	age	6	3.09	0.01	0.09	0.00	2.20	2.70	0.00
	dis	7	3.95	12.02	10.44	3.04	3.17	1.86	2.18
accessibilité autoroutes	rad	8	7.40	9.56	4.72	1.29	2.46	2.10	0.83
taux des taxes	tax	9	8.88	6.73	2.82	0.76	3.87	3.64	1.07
nb élèves / nb enseignants	ptratio	10	1.78	5.15	9.97	2.89	7.93	8.70	6.48
	b	11	1.34	0.92	2.56	0.69	2.37	2.97	1.12
Indice de pauvreté de la population	lstat	12	2.93	17.64	18.76	6.02	20.59	17.92	36.56
	Sum		43.74	68.10	72.70	21.51	73.93	73.93	73.93

Autres entrées : pollution, criminalité, distance aux centres d'emplois, ...

À interpréter bien sûr au regard de la causalité potentielle entre variables

Remarque : - effet de seuillage brutal de la mesure PMVD

CAS NON LINÉAIRE - SHAPLEY EFFECTS

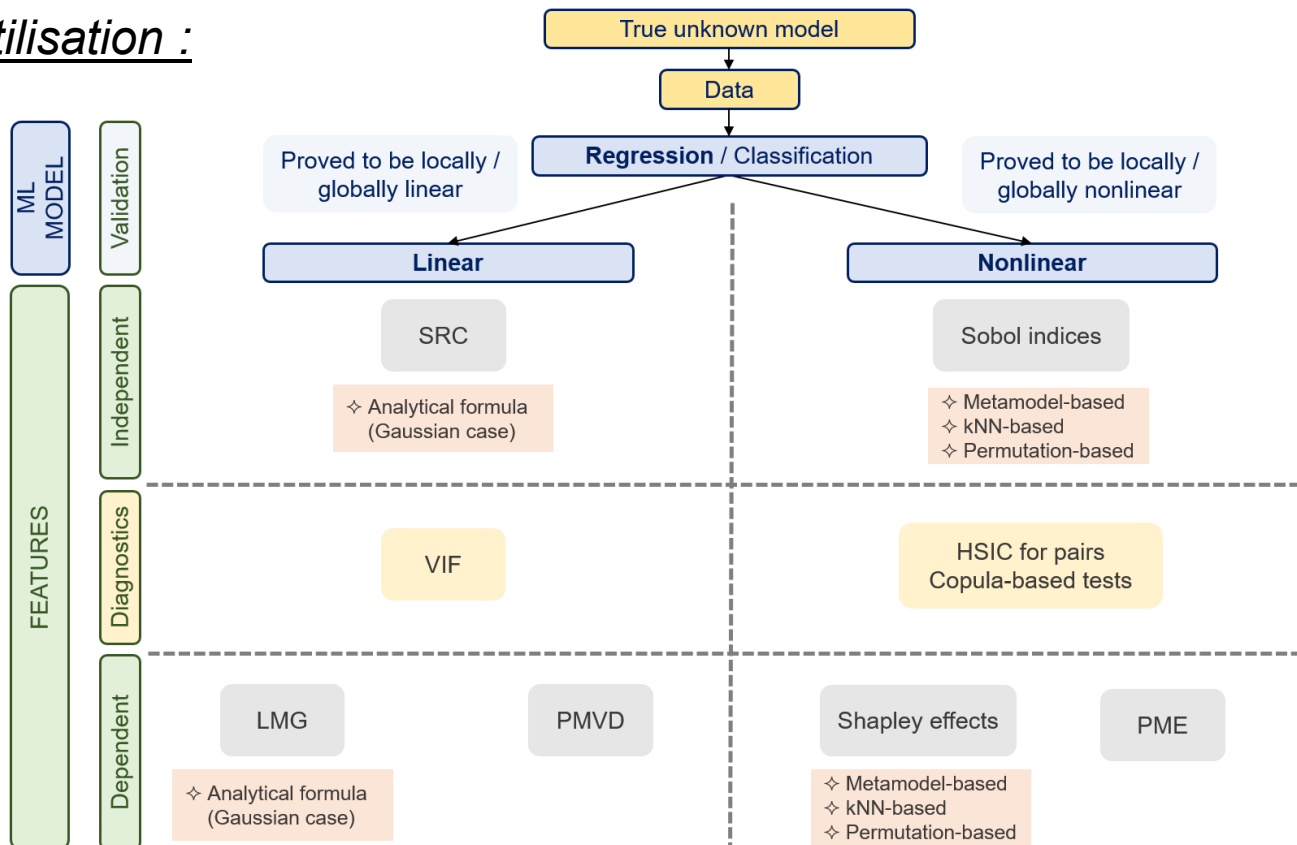
Indices de sensibilité dans le cas non linéaire

- Cas indépendant** : indices de Sobol' (extension SRC²)

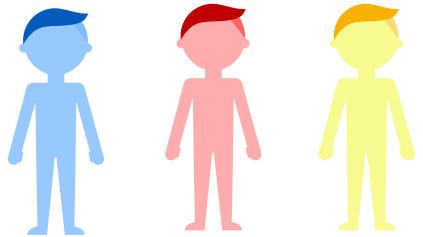
$$S_i = \text{Var}[E(Y|X_i)]/\text{Var}(Y)$$

$$T_i = E[\text{Var}(Y|X_{-i})]/\text{Var}(Y)$$
- Cas dépendant** : ANCOVA : déconseillée (indices négatifs)
 - => **indices de Shapley** (extension LMG : indices de Sobol remplacent les R^2)
 - => *NEW (en cours de consolidation) : Proportional Marginal Effects (PME)*

Guide d'utilisation :



Les valeurs de Shapley par l'exemple



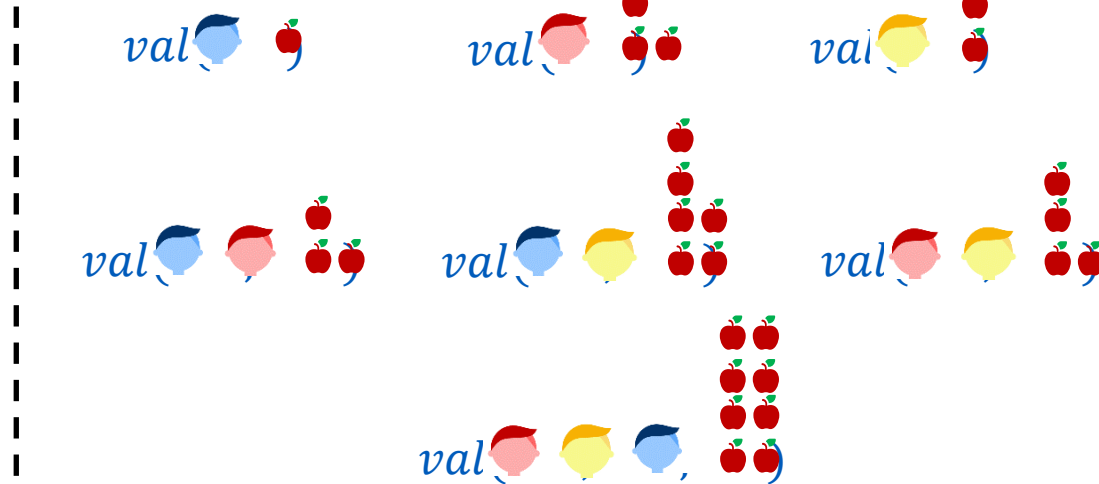
Timide Leader Diplomate



Comment redistribuer leur production de manière coopérative ?

4 Axiomes :

1. **(Efficacité)** La somme des valeurs réparties doit être égale à la production totale
2. **(Symétrie)** Si, pour toute coalition, l'apport marginal de deux joueurs est égale, leur part doit être égale.
3. **(Joueur nul)** Si, pour toute coalition, l'apport marginal d'un joueur est nul, sa part est nulle.
4. **(Additivité)** Si un jeu peut être décomposé en deux sous-jeux distincts, les valeurs de Shapley de ce jeu sont la somme de ceux des deux sous-jeux.



L'unique solution est :

$$\phi_j = \frac{1}{d} \sum_{A \subset -j} \binom{d-1}{|A|}^{-1} (val(A \cup \{j\}) - val(A))$$

Shapley effects for sensitivity analysis

In a cooperative game, Shapley values allocate the earnings to each actor of the game

Let $\text{val}(\cdot)$ be the cost function and $\mathcal{D} = \{1, \dots, d\}$

$$Sh_i = \sum_{u \subseteq \mathcal{D} \setminus \{i\}} \frac{(d - |u| - 1)! |u|!}{d!} |\text{val}(u \cup \{i\}) - \text{val}(u)|$$

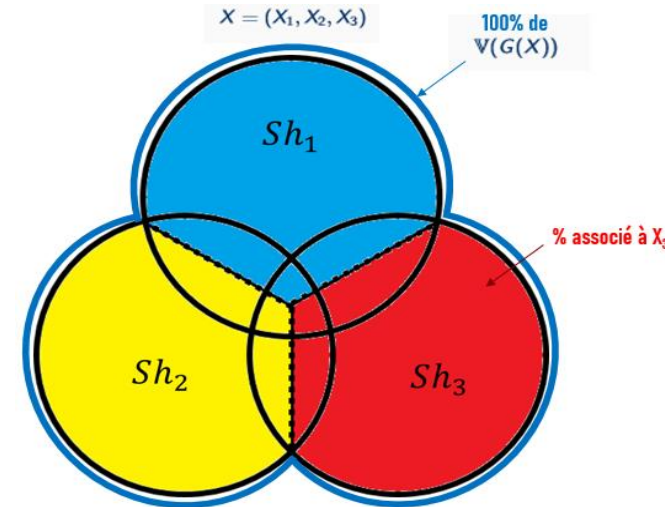
[Shapley 53]

Variance-based SA => Shapley effects as sensitivity indices

by taking $\text{val}(u) = S_u = \text{Var}[E(Y|X_u)] / \text{Var}(Y)$ [Owen 14]

Important properties:

- 1) $Sh_i \geq 0 \forall i = 1, \dots, d$ and $\sum_{i=1}^d Sh_i = 1$
- 2) In case of independent inputs: $Sh_i = \sum_{u \ni i} \frac{S_u}{|u|}$
then $S_i \leq Sh_i \leq T_i$ (not respected in general case)



- Ne requiert pas d'hypothèse d'indépendance
- Permet une décomposition interprétable de la variance de la variable d'intérêt

- Les parts dues aux interactions, et celles dues à la dépendance ne sont pas explicites

Simple linear models: Shapley = LMG

[I. & Prieur 2019]

$\mathbf{X} \sim \mathcal{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $Y = \beta_0 + \boldsymbol{\beta}^\top \mathbf{X}$, with

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \rho \in [-1, 1], \sigma_i > 0.$$

We have $\sigma^2 = V[Y] = \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2$. Then

$$\begin{cases} \sigma^2 Sh_1 = \beta_1^2\sigma_1^2\left(1 - \frac{\rho^2}{2}\right) + \rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2\frac{\rho^2}{2}, \\ \sigma^2 S_1 = \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \rho^2\beta_2^2\sigma_2^2 \text{ and } \sigma^2 S_1^{\text{tot}} = \beta_1^2\sigma_1^2(1 - \rho^2). \end{cases}$$

Simple linear models: examples

[I. & Prieur 2019]

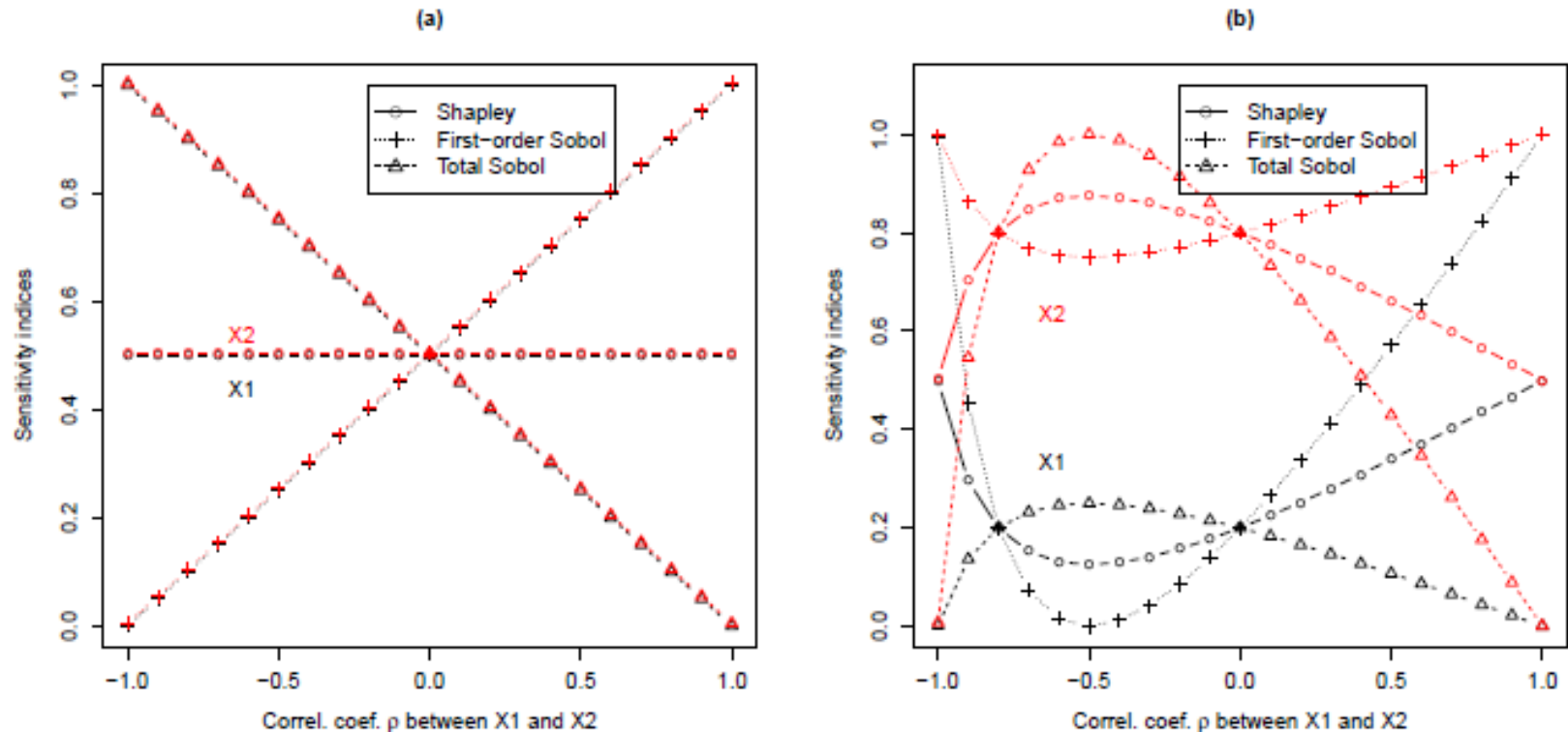


Figure: Sensitivity indices on the linear model ($\beta_1 = 1, \beta_2 = 1$) with two Gaussian inputs. (a): $(\sigma_1, \sigma_2) = (1, 1)$. (b): $(\sigma_1, \sigma_2) = (1, 2)$.

Another 2 simple examples

[I. & Prieur 2019]

$$G(X) = X_1 + X_2 X_3,$$
$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \rho \\ 0 & 1 & 0 \\ \rho & 0 & 1 \end{pmatrix} \right).$$

The three inputs have Shapley effects:

$$Sh_1 = 0.5 - 2\rho^2/12$$

$$Sh_2 = 0.25 + \rho^2/12$$

$$Sh_3 = 0.25 + \rho^2/12$$

$$\mathbf{x} \sim N \left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

$$Y = f(x_1, x_2) = x_1$$

$$V = \text{Var}(Y) = 1$$

$$Sh_1 = 1 - \rho^2/2; Sh_2 = \rho^2/2$$

The so-called Shapley's joke

The Shapley effects can be understood as an averaging index over the interaction and dependence effects

Numerical estimation algos in the general case

$$Sh_i = \frac{1}{d!} \sum_{\pi \in \Pi(\mathcal{D})} |\text{val}(P_i(\pi) \cup \{i\}) - \text{val}(P_i(\pi))|$$

$\Pi(\mathcal{D}) = \text{set of all permutations, } P_i(\pi) \text{ variables that precede } X_i \text{ in } \pi$

Instead of $\text{val}(u) = S_u$, take $\text{val}(u) = T_u = E[\text{Var}(Y|X_{-u})]/\text{Var}(Y)$

[Song et al. 2016]

Tuning parameters:

N_v : sample size for $\text{Var}(Z)$ estimation ; m : number of permutations

N_i : inner loop size (conditional variance) ; N_o : outer loop size (expectation)

$$\Rightarrow \text{Cost} = N_i N_o m (d-1) + N_v$$

1) Exact permutation algorithm: $m = d!$

Optimal with $N_i=3$ \Rightarrow Cost: $N = 3 d! N_o (d-1) + N_v$

2) Random permutation algorithm: Monte Carlo (MC) on permutation loop

Optimal with $N_i=3$ and $N_o=1$ \Rightarrow Cost: $N = 3 m (d-1) + N_v$

Important remarks:

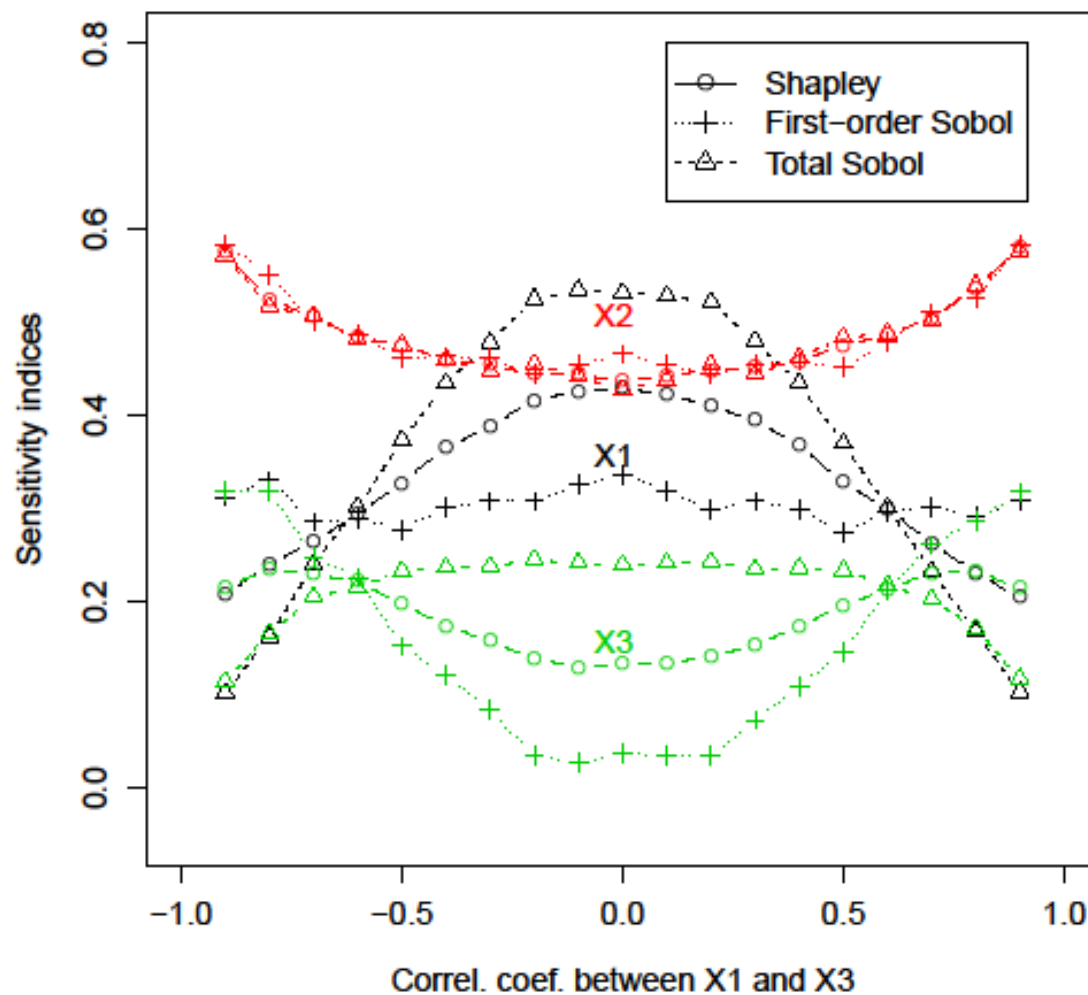
1) *Bootstrap confidence intervals can be computed [Benoumechiara & Elie-Dit-Cosaque 19]*

2) *A costly model can be replaced by a metamodel (ex: Gaussian process)*

3) *Nearest-neighbors estimation can be used instead of MC (\Rightarrow given-data algo which does not require to know & simulate conditional input distributions)*

Numerical test on Ishigami function

$$\begin{cases} f(X_1, X_2, X_3) = \sin X_1 + 7(\sin X_2)^2 + 0.1X_3^4 \sin X_1 \\ X_i \sim U_{[-\pi; \pi]} \text{ for } i = 1, \dots, 3 \end{cases}$$



Random permut. algo

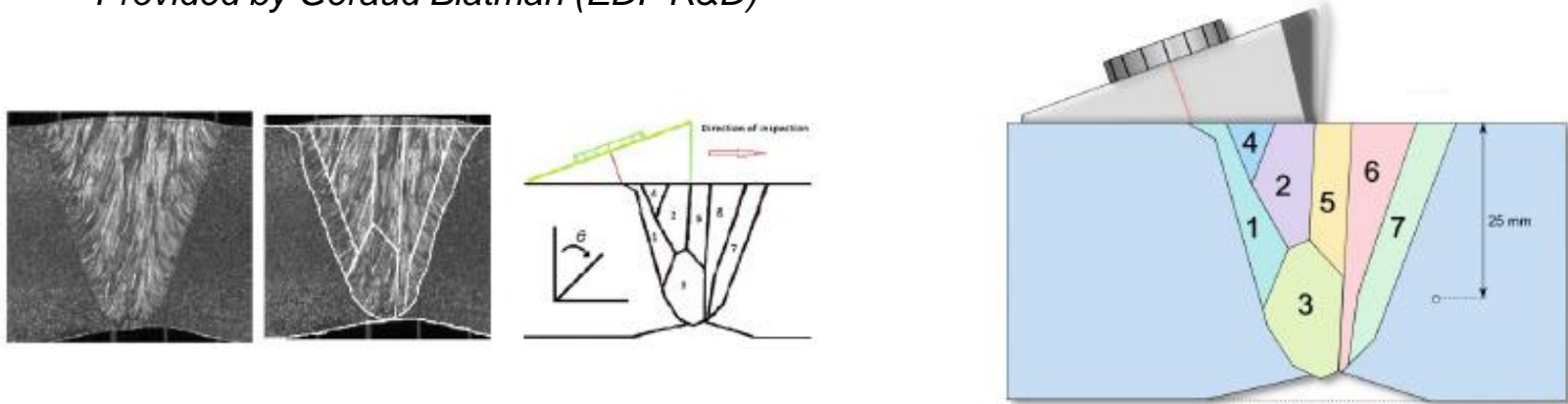
$N_v = 1e4$
 $N_i = 3$
 $N_o = 1$
 $m = 5e4$

Industrial application

Non Destructive Control for a weld inspection with ultrasonic waves

[Rupin et al. 14]

Provided by Géraud Blatman (EDF R&D)



Numerical simulations by Finite Element model (ATHENA 2D)

Output: signal amplitude

Inputs (11):

- 4 non-correlated elastic constants (Gaussian)
- 7 correlated orientations (Gaussian)

500 numerical experiments (Sobol' sequence)

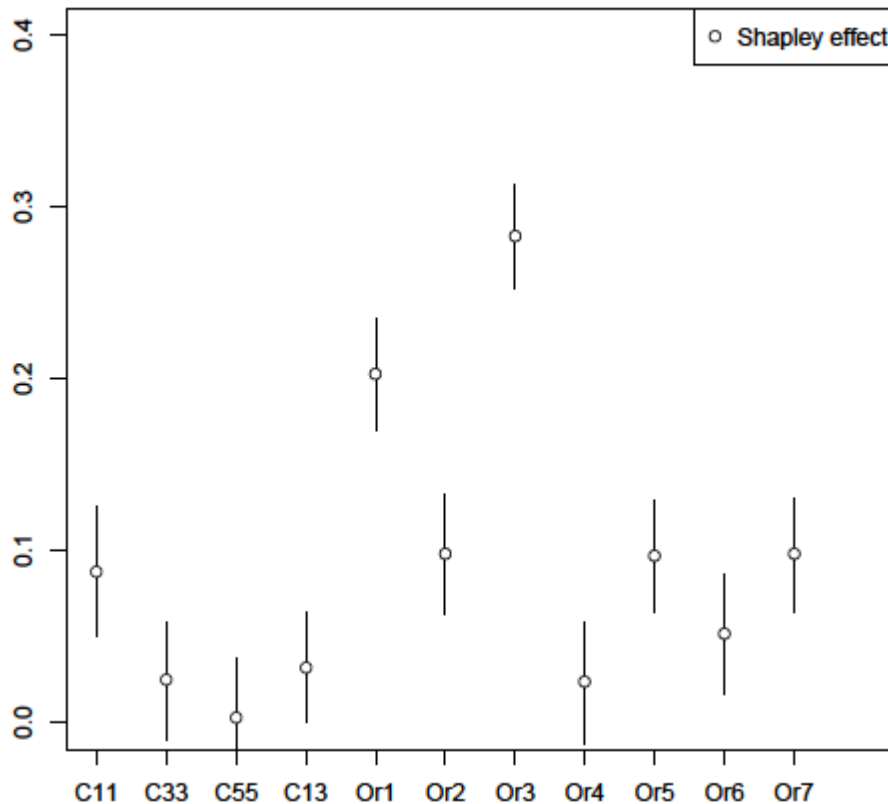
3 h of cpu-time on a HPC cluster

$$\begin{pmatrix}
 1 & 0.80 & 0.74 & 0.69 & 0.31 & 0.23 & 0.20 \\
 & 1 & 0.64 & 0.53 & 0.59 & 0.51 & 0.46 \\
 & & 1 & 0.25 & 0.60 & 0.57 & 0.54 \\
 & & & 1 & -0.25 & -0.34 & -0.33 \\
 & & & & 1 & 0.96 & 0.84 \\
 & & & & & 1 & 0.95 \\
 & & & & & & 1
 \end{pmatrix}$$

Results: Kriging-predictor based Shapley effects

Linear model: $R^2 = 25\%$

Kriging metamodel: $Q^2=86.5\%$



$N_v = 1e4$

$N_i = 3$

$N_o = 1$

$m = 1e4$

(R memory allocation problem
with larger m values)

Travaux récents sur les valeurs proportionnelles : Analogie avec LMG/PMVD => Shapley/PME

- La valeur de Shapley est la seule allocation **efficace** vérifiant la propriété « **contributions équilibrées** » : la présence simultanée de deux joueurs i et j **ajoute une même valeur** aux valeurs qu'ils auraient obtenues en l'absence l'un de l'autre
Autrement dit, le gain apporté par l'interaction de i et j est additionnel.
- La valeur proportionnelle est l'unique allocation **efficace** vérifiant la propriété « **gains proportionnels égaux** » : la présence simultanée de deux joueurs i et j **multiplie par une même valeur** les valeurs qu'ils auraient obtenues en l'absence l'un de l'autre
Autrement dit, le gain apporté par l'interaction de i et j est multiplicatif

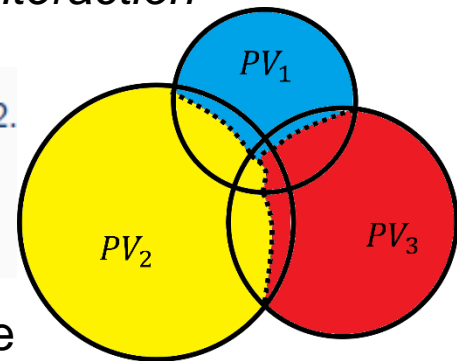
=> C'est un partage différent du surplus de valeur créée par leur *interaction*

$$PV\left(\left(\{1, 2\}, v\right)\right)_i = v(\{i\}) + \frac{v(\{i\})}{v(\{1\}) + v(\{2\})} \left(v(\{1, 2\}) - v(\{1\}) - v(\{2\})\right), \quad i = 1, 2.$$

$$Shap\left(\left(\{1, 2\}, v\right)\right)_i = v(\{i\}) + \frac{1}{2} \left(v(\{1, 2\}) - v(\{1\}) - v(\{2\})\right), \quad i = 1, 2.$$

PV => Proportional Marginal Effects (PME), qui résolvent la blague de Shapley

Les effets d'interaction restent à approfondir (cf. notion de synergie)



[Hérin et al. 2021]

Merci - Références

- N. Benoumechiara and K. Elie-Dit-Cosaque. Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms, *ESAIM: Proceedings and Surveys*, 2019.
- B. Broto, F. Bachoc, M. Depecker and J-M. Martinez. Sensitivity indices for independent groups of variables, *MCS*, 2019
- B. Broto, F. Bachoc, and M. Depecker. Variance reduction for estimation of Shapley effects and adaptation to unknown input distribution, *SIAM/ASA JUQ*, 2020
- L. Clouvel, *Uncertainty quantification of the fast flux calculation for a PWR vessel*, Thèse de l'Université Paris-Saclay, 2019
- S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. Basics and trends in sensitivity analysis. Theory and practice in R. *SIAM*, 2021
- U. Grömping, Relative importance for linear regression in R: the Package relaimpo. *JSS*, 2006
- M. Hérin, M. Il Idrissi, V. Chabridon and B. Iooss, Proportional marginal effects for sensitivity analysis with correlated inputs, *Submitted abstract to SAMO2022*, 2021
- M. Il Idrissi, V. Chabridon and B. Iooss, Developments and applications of Shapley effects to reliability-oriented sensitivity analysis with correlated inputs, *EMS*, 2021
- M. Il Idrissi, B. Iooss and V. Chabridon, Mesures d'importance relative par décompositions de la performance de modèles de régression, *52èmes JdS*, pp 497-502, Nice, France, Juin 2021
- B. Iooss, S. Da Veiga, A. Janon and G. Pujol, R sensitivity package v1.27.0, 2021
- B. Iooss and C. Prieur. Shapley effects for sensitivity analysis with dependent inputs: comparisons with Sobol' indices, numerical estimation and applications, *IJUQ*, 2019
- Johnson, J. and LeBreton, J. History and use of relative importance indices in organizational research, *ORM*, 2004
- R.H. Lindeman, P.F. Merenda and R.Z. Gold, *Introduction to bivariate and multivariate analysis*, Scott Foresman / Company, 1980
- A.B. Owen. Sobol' indices and Shapley value. *SIAM/ASA JUQ*, 2014
- A.B. Owen and C. Prieur. On Shapley value for measuring importance of dependent inputs, *SIAM/ASA JUQ*, 2017
- L.S. Shapley. A value for n-persons game. In H.W. Kuhn and A.W. Tucker, editors, Contributions to the theory of games II, Annals of mathematic studies. Princeton University Press, Princeton, NJ, 1953
- I.M. Sobol, On sensitivity estimation for nonlinear mathematical models, *MMCE*, 1993
- E. Song, B.L. Nelson, and J. Staum. Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA JUQ*, 2016