Importance ranking of parameter uncertainties in geo-hazard assessments

Jeremy Rohmer

UNIVERSITE DE LORRAINE
Ecole Nationale Supérieure des Mines de Nancy
Laboratoire GeoRessources
Ecole Doctorale RP2E

THESE

Présentée en vue du grade de

DOCTEUR DE L'UNIVERSITE DE LORRAINE
en Génie Civil-Hydrosystèmes-Géotechnique

Par

Jérémy ROHMER

Importance ranking of parameter uncertainties in geo-hazard assessments

Analyse de sensibilité des incertitudes paramétriques dans les évaluations d'aléas géotechniques

le 16 Novembre 2015

Devant le jury composé de

Michael OBERGUGGENBERGER– University of Innsbruck – Austria	Rapporteur
Michael BEER – University of Liverpool – UK	Rapporteur
Anne GEGOUT-PETIT – Université de Lorraine – France	Examinateur
Gilles GRANDJEAN – BRGM – France	Examinateur
Thierry VERDEL – Université de Lorraine – France	Directeur de thèse
Jack-Pierre PIGUET – Université de Lorraine – France	Co-directeur de thèse

Acknowledgements

I thank T. Verdel and J.-P. Piguet (Uni. Lorraine) for accepting to supervise this PhD thesis. I am very grateful to my co-authors, C. Baudrit (INRA), E. Foerster (CEA), A. Nachbaur (BRGM) and T. Verdel (Uni. Lorraine), for the constructive work and fruitful discussions which led to the publications supporting the present manuscript.
I also thank the BRGM directorate of "Risks and Prevention" division (J.L. Foucher, H. Fabriol and G. Grandjean) for supporting my "personal and professional" project.
I am also grateful to D. Guyonnet (BRGM/ENAG) and G. Grandjean (BRGM) for introducing me to Fuzzy sets and new uncertainty theories and O. Sedan (BRGM) for his very valuable advice on practical / operational aspects of natural risk assessments.
Last but not least, I am very grateful to my family (Julia, my parents, my parents in law and my sister in law) for supporting me. Special thanks to the French train company SNCF for its repetitive delays, which let me enough time for writing this manuscript.

Orleans, 16 November 2015 J. Rohmer.

Technical Summary

The central topic of the present thesis is the treatment of epistemic uncertainty in geo-hazard assessments (like landslide, earthquake, etc.). Contrary to aleatory uncertainty (*aka* randomness, variability), epistemic uncertainty can be reduced through additional measurements (lab tests, in site experiments, etc.) or modelling (e.g., through numerical simulations) or extra research efforts. Among the different types of epistemic uncertainties, we focused here on the parametric one: this corresponds to the incomplete knowledge of the correct setting of the input parameters (like values of soil properties) of the model supporting the geo-hazard assessment. A possible option to manage this type of uncertainty is through sensitivity analysis: 1. identify the contribution of the different input parameters in the uncertainty on the final hazard outcome; 2. rank them in terms of importance; 3. decide accordingly the allocation of additional characterisation studies.

For this purpose, variance-based global sensitivity analysis (VBSA) is a powerful procedure, which allows: i. incorporating the effect of the range of the input variation and of the nature of the probability distribution (normal, uniform, etc.); ii. exploring the sensitivity over the whole range of variation (i.e. in a global manner) of the input random variables; iii. fully accounting for possible interactions among them; and iv. providing a quantitative measure of sensitivity without introducing a priori assumptions on the model's mathematical structure (i.e. model-free). The most important sources of parameter uncertainty can then be identified (using the main effects) as well as the parameters of negligible influence (using the total effects). Besides, some key attributes of the model behaviour can be identified (using the sum of the main effects). Yet, to the author's best knowledge, this kind of analysis has rarely been conducted in the domain of geo-hazard assessments. This can be explained by the specificities of the domain of geo-hazard assessments, which impose considering several constraints, which are at the core of the present work.

Most numerical models supporting geo-hazard assessments have moderate-to-high computation time (typically several minutes, even hours), either because they are large-scale (e.g., landslide susceptibility assessment at the spatial scale of a valley), or because the underlying processes are difficult to be numerically solved (e.g., complex elastoplastic rheology law like the Hujeux model describing the complex coupled hydromechanical behaviour of a slip sur-

face). Despite the extensive research work on the optimization of the computation algorithms, VBSA remains computationally intensive, as it imposes to run a large number of simulations (> 1,000). In this context, VBSA can be made possible via the combination with meta-modelling techniques. This technique consists in replacing the long-running numerical model by a mathematical approximation referred to as "meta-model" (also named "response surface", or "surrogate model"), which corresponds to a function constructed using a few computer experiments (typically 50-100, i.e. a limited number of time consuming simulations), and aims at reproducing the behaviour of the "true" model in the domain of model input parameters and at predicting the model responses with a negligible computation time cost.

The applicability of the combination VBSA and meta-models was demonstrated using the model developed by Laloui and co-authors at EPFL (Lausanne) for studying the Swiss La Frasse landslide. We focused on the sensitivity of the surface displacements to the seven parameters of the Hujeux law assigned to the slip surface. In this case, a single simulation took 4 days of calculation. On the other hand, evaluating the main effects (first order sensitivity indices) should require >1,000 different simulations, which is here hardly feasible using the numerical simulator. This computation burden was alleviated using a kriging-type meta-model constructed using 30 different simulations. Furthermore, the impact of the meta-model error (i.e. the additional uncertainty introduced because the true simulator was replaced by an approximation) was discussed by treating the problem under the Bayesian formalism. This allowed assigning confidence intervals to the derived sensitivity measures: the importance ranking could then be done accounting for the limited knowledge on the "true" simulator (i.e. through only 30 different long-running simulations), hence increasing the confidence in the analysis. To the author's best knowledge, the application of such kinds of technique is original in the domain of landslide risk assessment.

The second limitation of VBSA is related to the nature of the parameters (input or output): they are scalar. Yet, in the domain of geo-hazard, parameters are often functional, i.e. they are complex functions of time or space (or both). This means that parameters can be vectors with possible high dimension (typically 100-1,000). For instance, the outputs of the La Frasse model correspond to temporal curves of the displacements (discredited in 300 steps) at any nodes of the mesh, i.e. the outputs are vectors of size 300 at any spatial location. Another example is the spatial distribution of hydraulic conductivities of a soil formation. Focusing first on the functional output case, a methodology to carry out dynamic (global) sensitivity analysis of landslide models was described combining: 1. basis set expansion to reduce the dimensionality of the functional model output; 2. extraction of the dominant modes of variation in the overall structure of the temporal evolution; 3. meta-modelling techniques to achieve the computation, using a limited number of simulations, of the sensitivity indices associated to each of the modes of variation. These were interpreted by adopting the perspective of the risk

practitioner in the following fashion: “identifying the properties, which influence the most the possible occurrence of a destabilization phase (acceleration) over the whole time duration or on a particular time interval”. However, a limitation was underlined, namely the physical interpretation of the dominant modes of variation, especially compared to the traditional time-varying VBSA (more easily interpretable, but also intractable for very long time series). Based on the study on the functional output, the applicability of the proposed methodology was also investigated for the case of functional inputs using as an example, a random field assigned to the heterogeneous porosity of a soil formation.

Finally, a third limitation of VBSA is related to the representation of uncertainty. By construction, VBSA relies on tools/procedures of the probabilistic setting. Yet, in the domain of geo-hazard assessments, the validity of this approach can be debatable, because data are often scarce, incomplete or imprecise. In this context, the major criticisms available in the literature against the systematic use of probability in such situations were reviewed. On this basis, the use of a flexible uncertainty representation tool was investigated, namely Fuzzy Sets to handle different situations of epistemic uncertainty. For each situation, examples of real cases in the context of geo-hazard assessments were used:

- Vagueness due to the use of qualitative statements. The application of Fuzzy Sets was illustrated in the context of susceptibility assessment of abandoned underground structures. In particular, it is shown how the so-called “threshold effect” can be handled when the expert defines classes of hazard / susceptibility;
- Reasoning with vague concepts. This is handled using Fuzzy Logic. This is illustrated with the treatment of imprecision associated to the inventory of assets at risk in the context of seismic risk analysis;
- Imprecision. This is handled by defining possibility distributions, which have a strong link with Fuzzy Sets. This is illustrated with the representation of uncertainty on the amplification factor of lithological site effects in the context of seismic risk analysis;
- Imprecision on the parameters of a probabilistic model. This is handled in the setting of fuzzy random variables. This is illustrated using a probabilistic damage assessment model in the context of seismic risk analysis, namely the RISK-UE, level 1 model.

On this basis, the issue of sensitivity analysis considering a mixture of randomness and imprecision was addressed. Based on a literature review, a major limitation was outlined, namely the computation time cost: new tools for uncertainty representation basically rely on interval-valued tools, the uncertainty propagation then involves optimisation procedure, which can be highly computationally intensive. In this context, an adaptation of the contribution to failure probability plot of [Li and Lu, 2013], to both handle probabilistic and possibilistic information was proposed. The analysis can be conducted in a post-processing step, i.e. using only the samples of random intervals and of random numbers necessary for the propagation

phase, hence at no extra computational cost. Besides, it allows placing on the same level random and imprecise parameters, i.e. it allows the comparison of their contribution in the probability of failure so that concrete actions from a risk management perspective can be decided accordingly. The applicability of this easy-to-use tool was demonstrated using real cases, where it is questionable to use probabilities to treat uncertainty. The first application case corresponds to stability analysis of steep slopes. The main imprecise parameters in this case are the tension crack's height located in the upper part of the cliff and the toe height. The second one corresponds to the stability analysis of an abandoned underground quarry, where the extraction ratio was imprecise because it could only be estimated with great difficulties (due to the particular geometry of the quarry). Finally, a third example was used, namely the stability analysis of a mine pillar presenting thin layers of clay, whose properties are difficult to evaluate in practice. This last example imposed to rely on meta-modelling techniques to ease the joint uncertainty propagation phase using the long-running mechanical numerical code. In summary, the present work should be seen as an effort to handle epistemic parameter uncertainties in geo-hazard assessments. First, the achievement is of methodological nature (methodology for conducting VBSA using long running simulators, methodology for conducting VBSA adapted to functional outputs, methodology for conducting sensitivity analysis when both imprecision and randomness are present). This methodological work takes advantages of the recent advances in the statistical community (VBSA, basis set expansion, Fuzzy Sets, Fuzzy random variables, hybrid propagation, etc.) to answer practical questions (what drives the uncertainty on the results of the hazard assessment? How to conduct multiple simulations when the simulation code takes one hour to be run? How should the uncertainty be treated when the only pieces of information available restrict to vague statements and a few quantitative estimates?). A great attention has been paid to investigate the applicability of each proposed technique / procedure i.e. by highlighting the pros and cons through the confrontation to real cases. This constitutes the second achievement of the present work.

Résumé étendu

Le présent travail de thèse se concentre sur le traitement des incertitudes de type "épistémique" dans les modèles d'évaluation des aléas dits géotechniques (e.g., séismes, glissements de terrain, subsidences liées aux mines, etc.). Contrairement aux incertitudes aléatoires (aussi dénommées "variabilité naturelle"), les incertitudes épistémiques peuvent être réduites en menant des études supplémentaires/complémentaires (mesures en laboratoire, in situ, ou via des modélisations numériques, redéfinition du protocole expérimental, etc.). Parmi les différentes formes d'incertitudes épistémiques, nous nous concentrons ici sur celle dite "paramétrique", qui est liée aux difficultés de quantifier les valeurs des paramètres d'entrée du modèle utilisé pour l'analyse de ces aléas. Une possible stratégie pour gérer ce type d'incertitude repose sur l'analyse de sensibilité consistant à : 1. identifier la contribution de chacun des paramètres dans l'incertitude totale de l'évaluation de l'aléa ; 2. de les ordonner selon cette contribution, i.e. selon leur influence / importance ; 4. d'identifier les paramètres d'influence négligeable permettant ainsi de simplifier le modèle d'évaluation ; 4. de décider alors l'allocation des études, analyses et efforts futurs en matière de caractérisation.

Dans un premier chapitre, nous nous concentrons sur les méthodes identifiées dans la littérature comme étant les plus avancées pour traiter ce problème, à savoir l'analyse globale de sensibilité reposant sur la décomposition de la variance fonctionnelle VBSA. Cette approche définit des indices de sensibilité quantitatifs (entre 0 et 1) correspondant aux indices de Sobol' : ils permettent non seulement d'évaluer la part de la variance résultant de la variation d'un seul paramètre d'entrée (i.e., pourcentage de la variance de la réponse expliquée), mais également de la contribution résultant de la variation de plusieurs paramètres (interactions). L'avantage de cette analyse est d'être globale (toutes les valeurs paramètres sont modifiées en même temps) et de prendre en compte l'information probabiliste sur la représentation des incertitudes d'entrée. Nous avons appliqué ce type d'analyse à un modèle analytique de "pente infinie" pour évaluer la susceptibilité de glissement de terrain via un facteur de stabilité. La figure 1 donne un exemple du résultat pour les six paramètres du modèle de pente (C : cohésion du sol ; ϕ : angle de frottement ; γ : masse volumique ; θ : angle de la pente ; m : hauteur de la partie saturée du sol ; z : hauteur du glissement). Nous évaluons la contribution de chaque paramètre dans l'incertitude globale du facteur de stabilité de la

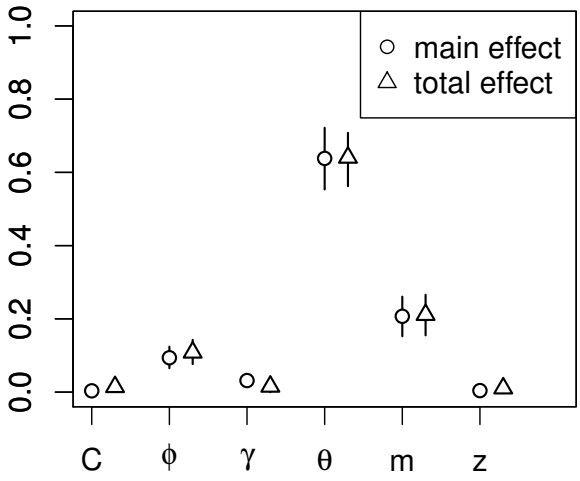

FIGURE 1 – Exemple d'un résultat dérivant dune analyse globale de sensibilité sur le modèle de pente infinie à partir des indices de Sobol' de 1er ordre et indices totaux.

pente en utilisant les indices de Sobol' de 1er ordre (symbole rond sur la figure 1). Ici, l'angle de la pente a la plus grande influence sur la variabilité du facteur de stabilité. Nous utilisons également les indices totaux (symbole triangle sur la figure 1), qui correspondent à la seule contribution du paramètre considéré et de ses interactions avec tous les autres paramètres : il est alors possible d'identifier les paramètres dont la contribution peut être considérée comme négligeable (à savoir ceux dont la valeur de l'indice total est proche de zéro), ainsi que d'avoir des informations sur la structure mathématique de la relation entre facteur de stabilité et paramètres d'entrée. Ici, les paramètres pouvant être négligés sont la cohésion C et la hauteur du glissement z.

Cependant, malgré la richesse de l'information que l'on est capable d'obtenir, ce type d'analyse n'est pas systématiquement utilisé dans le domaine des risques naturels. Cela peut s'expliquer par trois particularités de ce domaine, qui imposent de considérer plusieurs contraintes / limitations, qui sont au coeur du présent travail.

0.1 Limitation n°1 : gérer le temps de calcul

L'implémentation de VBSA exige un nombre important de simulations. Dans l'exemple de la figure 1, un algorithme basé sur l'échantillonnage Monte-Carlo a été utilisé : 80 000 simulations différentes ont été ici nécessaires pour le calcul des indices de sensibilité. Ce coût calculatoire

reste abordable dans le cas de modèles (semi-)analytiques à l'instar de celui utilisé pour la "pente infinie". Cependant, les estimations de l'aléa peuvent se baser sur des modèles numériques dont le temps de calcul n'est pas négligeable (plusieurs minutes voire heures). Par exemple, ce temps de calcul peut s'expliquer par l'obligation de traiter l'évaluation à grande échelle et donc d'utiliser des maillages avec un grand nombre de mailles/cellules (e.g., évaluation de la susceptibilité de glissement de terrain à l'échelle d'un bassin versant). Un autre exemple est un modèle numérique d'un glissement de terrain prenant en compte le couplage entre processus mécaniques et hydrauliques, dont la résolution numérique peut être ardue. Afin de surmonter ce problème, nous avons proposé de combiner VBSA avec la technique de méta-modélisation, qui consiste à capturer la relation mathématique entre les paramètres d'entrée et la variable de sortie du modèle numérique via une approximation mathématique construite avec un nombre restreint de simulations intensives (typiquement 50-100).

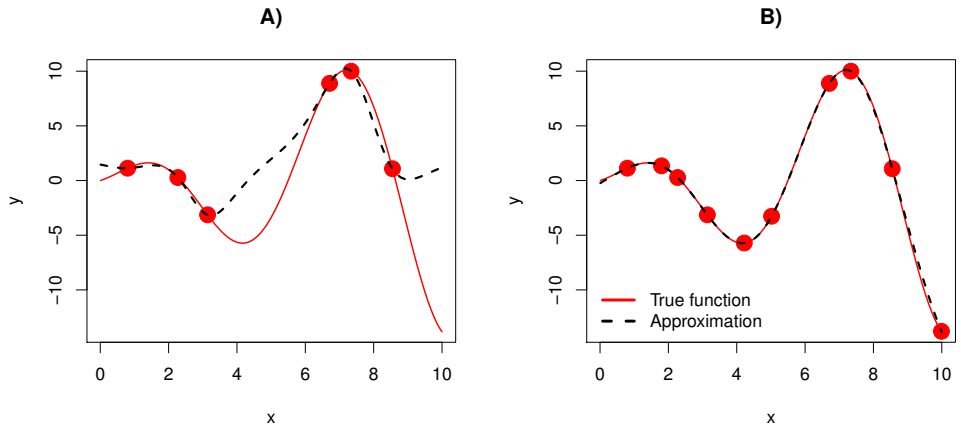

FIGURE 2 – Approximation d'un modèle 1d (rouge) par un méta-modèle de type krigeage (en noir) construit à partir des configurations indiquées par des points rouges : A) 6 simulations différentes ; B) 10 simulations différentes.

Plusieurs types de méta-modèles existent et nous nous sommes concentrés sur le krigeage numérique. Cette technique repose sur les outils d'interpolation spatiale de la géostatistique. Dans notre cas, les valeurs interpolées ne sont pas des coordonnées géographiques, mais sont les paramètres d'entrée du modèle numérique. A titre illustratif, la figure 2 donne l'exemple d'une fonction simple avec un paramètre d'entrée x : $y = x(\cos(x) + \sin(x))$ (en rouge) qui est approximée (en noir) en utilisant soit 6 (figure 2A)) ou 10 (figure 2B)) différentes configurations (valeurs) du paramètre x (points rouges). Le krigeage est associé à une tendance constante et une fonction de corrélation de type Matérn sans effet pépite. La figure 2 montre que seulement 10 valeurs de x sont nécessaires pour approximer de manière satisfaisante la vraie fonction.

Cette stratégie a été appliquée au cas réel du glissement de terrain de La Frasse (Suisse) dont le modèle numérique a un temps de calcul de plusieurs jours, car le comportement rhéologique du matériau au niveau de la surface de glissement suit une loi complexe (loi Hujeux). A partir d'un nombre limité de simulations (ici une trentaine), nous avons pu approximer les déplacements horizontaux en surface en fonction des valeurs des propriétés du matériau constituant la surface de glissement. En vérifiant la qualité d'approximation par une méthode par validation croisée, nous avons remplacé le code numérique couteux en temps de calcul par le méta-modèle et avons ainsi pu dériver les effets principaux (indices de Sobol' de 1er ordre) pour étudier la sensibilité sur les déplacements. Cependant, un prix à payer fut l'introduction d'une nouvelle source d'incertitude, i.e. celle liée au fait que l'on a remplacé le vrai modèle par une approximation. La figure 2A) illustre ce problème : la partie droite n'est pas bien approximée à cause du manque d'information (aucune simulation réalisée dans cette partie). L'impact de cette erreur sur les résultats de VBSA a été discuté en associant un intervalle de confiance aux indices de sensibilité via un traitement du problème d'apprentissage du méta-modèle dans le cadre Bayésien. Bien qu'il faille souligner la complexité de mise en oeuvre ainsi que la sensibilité aux hypothèses (en particulier aux lois de probabilité a priori), cette stratégie "méta-modèle-VBSA-traitement Bayésien" nous a permis d'apporter des éléments de réponse à la question de l'influence des sources d'incertitudes paramétriques en un temps limité et raisonnable (quelques jours incluant les simulations et la construction du méta-modèle) et avec un nombre restreint de simulations (ici une trentaine).

0.2 Limitation n°2 : gérer des paramètres variant dans l'espace et le temps

La seconde limitation est liée à la nature des paramètres dans le domaine des aléas géotechniques : ce sont souvent des fonctions complexes du temps et/ou de l'espace et non pas simplement des variables scalaires. En d'autres termes, ces paramètres sont souvent représentés par des vecteurs de grande dimension (typiquement 100-1000). Un exemple sont les séries temporelles des déplacements horizontaux (Fig. 3B) et C)) simulés à La Frasse lors de la variation de la pression de pore en pied de glissement de terrain (Fig. 3A)) : ces séries sont obtenues à tous les noeuds du maillage et sont discrétisées sur 300 pas de temps. Un autre exemple est la carte hétérogène des conductivités hydrauliques d'un sol.

Une première démarche consisterait à évaluer un indice de sensibilité à chaque pas de temps en utilisant les techniques décrites ci-avant. Cependant, cette démarche serait difficilement réalisable avec des vecteurs de très grande dimension et ne permettrait pas de prendre en compte la corrélation qui peut exister (dans l'exemple des séries temporelles, une valeur à un temps donné a un lien avec celles d'avant et celles d'après). Dans un premier temps, nous nous sommes focalisés sur le cas des séries temporelles des déplacements horizontaux dans

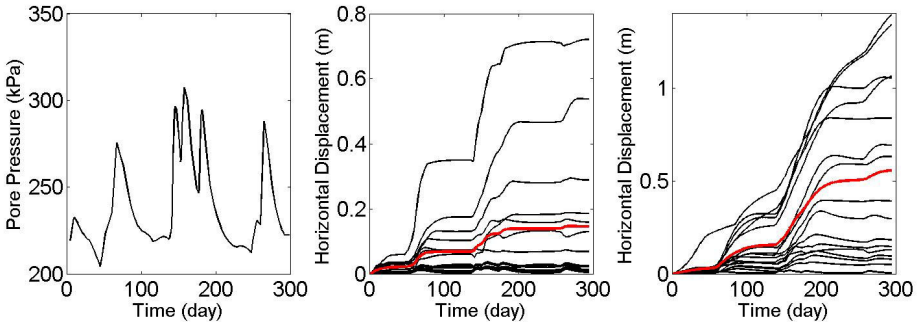

FIGURE 3 – Illustration des séries temporelles qui sont en entrée (pression de pore A)) et en sortie du modèle numérique simulant le glissement de La Frasse (glissements horizontaux en tête B) et en pied de glissement C)). La courbe rouge correspond à la moyenne temporelle.

le cas du glissement de La Frasse. Nous avons alors proposé une stratégie combinant :

- des techniques de réduction de la dimension, en particulier l'analyse des composantes principales ;
- un méta-modèle pour surmonter le coût calculatoire des indices de sensibilité.

La première étape permet de résumer l'information temporelle en décomposant les séries temporelles en un nombre restreint de paramètres (<3), qui correspondent aux composantes principales. Une analyse plus "physique" de ces composantes principales est faite en les interprétant comme une perturbation de la moyenne temporelle (courbe rouge sur la figure 3B) et C)) : cela permet alors d'identifier les principaux modes de variation temporelles pouvant être vus comme un enchaînement de phases de déstabilisation ou de stabilisation du glissement. Les indices de sensibilité calculés via le méta-modèle sont alors associés à chaque composante principale, i.e. à chaque mode de variation temporelle. Une telle démarche permet par exemple d'identifier les propriétés incertaines les plus importantes au regard de l'occurence d'une phase d'accélération lors du glissement sur une période donnée.

L'applicabilité de ce type de stratégie a aussi été discutée pour les paramètres d'entrée fonctionnels en se focalisant sur ceux spatialisés. Dans ce cas, la stratégie se trouve être limitée : 1. le nombre de composantes dans la décomposition reste important (plusieurs dizaines), i.e. assez grand pour rendre difficile la phase d'apprentissage (construction) du méta-modèle, qui est directement liée à ce nombre ; 2. le niveau auquel la décomposition peut être tronquée est décidé avant d'avoir pu lancer les simulations, i.e. avec peu de possibilité de savoir en amont si une partie de l'information laissée de côté pourrait avoir un impact sur la qualité d'approximation du méta-modèle ; 3. l'interprétation physique de chaque mode de variation peut être difficile. Sur cette base, des pistes de recherche ont été identifiées.

0.3 Limite n°3 : gérer le manque de connaissance

Enfin, une troisième limite est liée à l'hypothèse de base sur la représentation de l'incertitude. Par construction, VBSA repose sur les outils du cadre probabiliste avec l'hypothèse que la variance capture de façon satisfaisante toute l'incertitude sur la variable d'intérêt. Or, cette approche peut être limitée surtout dans le domaine des aléas géotechniques, pour lesquels les données / informations sont souvent imprécises, incomplètes voire vagues. Dans ce contexte de connaissance, l'utilisation systématique des probabilités peut être discutable. Une revue des principales critiques est faite et l'applicabilité d'un outil alternatif pour la représentation de l'incertitude est étudiée, à savoir les ensembles flous. La figure 4A) donne l'exemple d'un ensemble flou (définissant formellement une distribution de possibilités), qui permet de représenter une information d'expert du type : "je suis sûr de trouver la vraie valeur du paramètre incertain dans l'intervalle [a ;d] (support), mais l'intervalle [b ;c] (coeur) a plus de vraisemblance." A partir de ces deux informations, un ensemble d'intervalles emboîtés associés à un degré de confiance (α-coupe) est construit. Son interprétation dans le domaine des probabilités est donnée sur la figure 4B). Une telle représentation correspond à l'ensemble de distributions cumulées de probabilités dont les limites hautes et basses (Π et N) sont construites à partir des informations sur le coeur et le support.

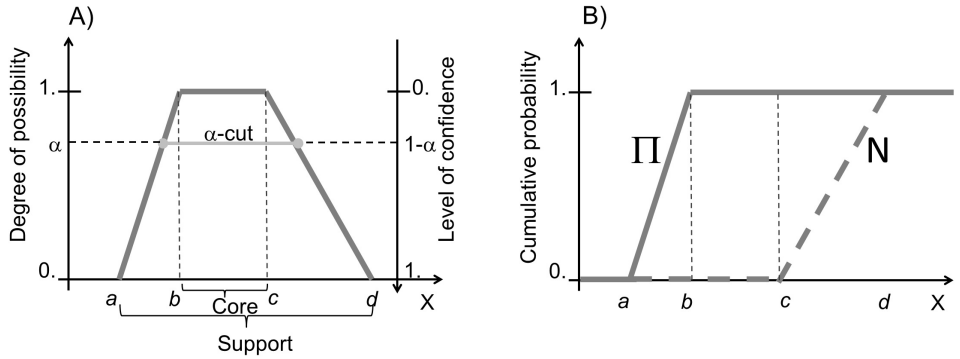

FIGURE 4 – A) Exemple d'un intervalle flou pour représenter l'imprécision d'un paramètre incertain à partir de l'information d'expert sur le support et le coeur. B) Interprétation dans le domaine des probabilités via les 2 distributions cumulées, haute Π et basse N.

Un ensemble flou est un outil très flexible pour traiter plusieurs formes d'incertitudes épistémiques :

- la représentation du caractère vague de l'information est abordée pour l'évaluation de la susceptibilité de présence de cavités à l'échelle régionale ;
- le raisonnement à partir de concepts qualitatifs vagues est traité dans le cas de l'imprécision liée à l'inventaire des éléments à risque pour un scénario de risque sismique à l'échelle d'une ville ;

- l'imprécision sur la valeur numérique d'un paramètre (comme illustrée par la figure 3) est abordée pour l'évaluation du coefficient d'amplification représentant les effets de site lithologique en risque sismique ;
- l'imprécision sur les valeurs des paramètres d'une courbe de décision probabiliste est abordée dans le domaine sismique.

Sur cette base, les principales procédures pour combiner représentation hybride des incertitudes (e.g., via probabilités et ensembles flous) et analyse de sensibilité sont étudiées. Une limitation majeure a été identifiée, à savoir le coût calculatoire : alors que la propagation dans le cadre purement probabiliste peut se baser sur des méthodes d'échantillonnage aléatoire exigeant basiquement de simuler différentes configurations des paramètres d'entrée, les nouvelles théorie de l'incertain exige souvent de manipuler des intervalles et donc de résoudre des problèmes d'optimisation. Une possible réponse à ce problème a été proposée en développant un outil graphique pour les analyses de stabilité. Une caractérisque intéressante est que cet outil n'est construit qu'à partir des simulations nécessaires à la propagation d'incertitude, donc sans coût calculatoire supplémentaire. Cet outil permet de placer sur le même niveau de comparaison, incertitude aléatoire et épistémique et donc d'identifier les contributions de chaque type d'incertitude à l'évaluation d'une probabilité de défaillance. Cette approche est appliquée à trois cas pour lesquels l'utilisation des probabilités est discutable pour représenter des paramètres imprécis : i. le cas d'un glissement de terrain dont les caractéristiques géométriques sont mal connues ; ii. le cas de la rupture d'un pilier dans une carrière abandonnée dont le taux d'extraction est difficilement évaluable à cause de la configuration particulière de la carrière ; iii. le cas de la rupture d'un pilier en calcaire présentant de fines couches argileuses dont les propriétés sont difficilement évaluables in situ. Notons que le cas iii. a exigé de développer une approche basée sur les méta-modèles, car l'évaluation de la stabilité du pilier exigeait un code numérique coûteux en temps de calcul.

La figure 5 donne l'exemple d'un résultat de l'outil graphique pour l'évaluation de stabilité d'une falaise d'angle de frottement ϕ (supposée être un paramètre aléatoire représenté par une distribution de probabilités) et de hauteur H_t du pied de falaise (supposée être un paramètre imprécis représenté par une distribution de possibilités) : plus la courbe dévit de la diagonale, plus le paramètre a une grande influence : ici c'est l'angle ϕ. En pratique, ce résultat indique que les futures actions en matière de gestion des risques devraient reposées sur des mesures préventives et/ou de protection, car ce paramètre est associé à une variabilité naturelle : l'incertitude ne peut donc pas être réduite. De plus, la portion du graphique où la déviation est maximale indique la région des quantiles de l'angle ϕ et la région des α-coupes de la hauteur H_t pour laquelle le paramètre est le plus influent : ici, cela correspond respectivement à la région des quantiles inférieurs à 75% de l'angle ϕ et à celle proche du coeur de H_t. Ce simple exemple montre comment ce nouvel outil peut être utilisé pour décider des actions futures

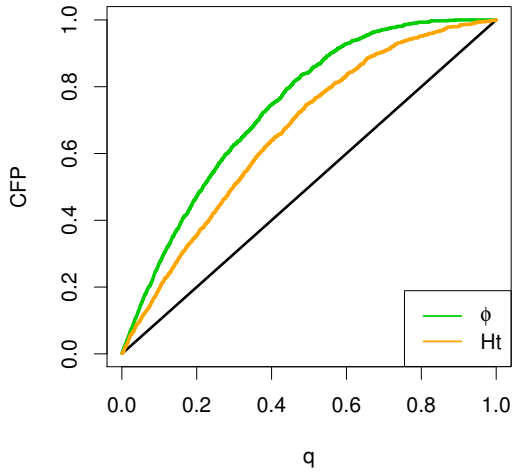

FIGURE 5 – Analyse de sensibilité avec l'outil graphique pour le cas d'une falaise dont le matériau a un angle de frottement aléatoire ϕ et de hauteur de pied de falaise H_t imprécis.

pour la gestion des incertitudes selon leur nature (épistémique ou aléatoire).

0.4 En résumé...

L'apport de cette thèse est avant tout d'ordre méthodologique. En se basant sur des techniques avancées dans le domaine de l'analyse statistique (indices de Sobol', techniques de réduction de dimension, ensembles flous, variables aléatoires flous, etc.), nous essayons d'apporter des réponses à des questions opérationnelles en matière de traitement des incertitudes dans l'évaluation des aléas géotechniques (glissement de terrain, séismes, cavités abandonnées, etc.), à savoir : quelle source d'incertitude doit être réduite en priorité ? Comment manipuler des codes de calcul avec un temps de calcul de plusieurs heures pour simuler de multiples scénarios (e.g., plusieurs centaines) ? Comment aborder la question des incertitudes lorsque les seules informations disponibles sont des opinions d'experts qualitatives et quelques observations quantitatives ? Ce travail repose soit sur une combinaison de plusieurs techniques, ou sur une adaptation de certaines d'entre elles. Un effort tout particulier a été fait pour étudier l'applicabilité de chaque procédure à l'aune de données sur des cas réels.

Le présent document a été rédigé sur la base des travaux de recherche que j'ai effectués au BRGM (Service géologqiue national) depuis 2010. La thèse repose sur quatre articles (trois en tant qu'auteur principal et un en tant que co-auteur), à savoir :

- Nachbaur, A., **Rohmer, J.**, (2011) Managing expert-information uncertainties for assessing collapse susceptibility of abandoned underground structures. Engineering Geology, 123(3), 166–178 ;

- **Rohmer, J.**, Foerster, E., (2011) Global sensitivity analysis of large scale landslide numerical models based on the Gaussian Process meta-modelling. Computers and Geosciences, 37(7), 917-927 ;

- **Rohmer, J.**, (2013) Dynamic sensitivity analysis of long-running landslide models through basis set expansion and meta-modelling. Natural Hazards, accepted, doi :10.1007/s11069-012-0536-3 ;

- **Rohmer, J.**, Baudrit, C., (2011) The use of the possibility theory to investigate the epistemic uncertainties within scenario-based earthquake risk assessments. Natural Hazards, 56(3), 613-632.

En plus de ces travaux, un étude originale a été effectuée en collaboration avec mon directeur de thèse, Thierry Verdel, et a été intégrée dans le présent manuscrit. Ce travail a également été publié récemment.

- **Rohmer, J.**, Verdel, T. (2014) Joint exploration of regional importance of possibilistic and probabilistic uncertainty in stability analysis. Computers and Geotechnics, 61, 308–315.

Lors de l'écriture, un effort tout particulier a été fait afin que le manuscrit ne soit pas un « simple » résumé de ces travaux, mais que le document ait une cohérence d'ensemble et forme un « tout ». Dans cette optique, plusieurs parties ont été réécrites par rapport aux articles originaux. Par ailleurs, des détails techniques ont été ajoutés dans un souci de clarté de l'énoncé.

Foreword

The present thesis summarises the research I have undertaken at BRGM (French geological survey) since 2010 on uncertainty treatment in various natural hazards: landslide, earthquake and abandonned underground structures. The thesis is strongly based on four peer-reviewed publications (three as first author and one as co-author):

- Nachbaur, A., **Rohmer, J.**, (2011) Managing expert-information uncertainties for assessing collapse susceptibility of abandoned underground structures. Engineering Geology, 123(3), 166–178;
- **Rohmer, J.**, (2013) Dynamic sensitivity analysis of long-running landslide models through basis set expansion and meta-modelling. Natural Hazards, accepted, doi:10.1007/s11069-012-0536-3;
- **Rohmer, J.**, Foerster, E., (2011) Global sensitivity analysis of large scale landslide numerical models based on the Gaussian Process meta-modelling. Computers and Geosciences, 37(7), 917-927.
- **Rohmer, J.**, Baudrit, C., (2011) The use of the possibility theory to investigate the epistemic uncertainties within scenario-based earthquake risk assessments. Natural Hazards, 56(3), 613-632.

In addition to these studies, an original research work has been conducted with my thesis advisor, Thierry Verdel, and added to the thesis. This was publised recently:

- **Rohmer, J.**, Verdel, T. (2014) Joint exploration of regional importance of possibilistic and probabilistic uncertainty in stability analysis. Computers and Geotechnics, 61, 308–315

A special attention has been paid to go beyond a “simple” summary of the main contents of the different research papers: some parts were re-written for sake of coherency of the whole document and additional technical details were added for sake of clarity.

Contents

Acknowledgements i

Technical Summary iii

Résumé étendu vii

0.1 Limitation n°1 : gérer le temps de calcul viii

0.2 Limitation n°2 : gérer des paramètres variant dans l'espace et le temps x

0.3 Limite n°3 : gérer le manque de connaissance xii

0.4 En résumé... xiv

Foreword xvii

List of figures xxiii

List of tables xxvii

1 Introduction 1

1.1 Hazard, Risk, uncertainty and decision-making 1

1.2 Aleatory and Epistemic uncertainty 3

1.3 Epistemic uncertainty of type "parameter" 5

1.4 A real-case example 6

1.5 Objectives and structure of the manuscript 7

2 A probabilistic tool: variance-based global sensitivity analysis 11

2.1 Global sensivity analysis 11

2.2 Variance-based global sensivity analysis 13

2.3 Application to slope stability analysis 15

2.3.1 Local sensitivity analysis 16

2.3.2 Global sensitivity analysis 18

2.4 Limitations and links with the subsequent chapters 20

3 Handling long-running simulators **23**
3.1 A motivating real case: the numerical model of the La Frasse landslide 23
3.1.1 Description of the model 23
3.1.2 Objective of the sensitivity analysis 24
3.2 A meta-model-based strategy 25
3.2.1 Principles 25
3.2.2 Step 1: setting the training data 26
3.2.3 Step 2: construction of the meta-model 27
3.2.4 Step 3: validation of the meta-model 28
3.3 A flexible meta-model: the kriging model 29
3.4 An additional source of uncertainty 31
3.5 Application to an analytical case 34
3.6 Application to the La Frasse case 36
3.7 Concluding remarks of Chapter 3 43

4 Handling functional variables **45**
4.1 Problem definition for functional outputs 45
4.2 Reducing the dimension 47
4.2.1 Principles 47
4.2.2 Principal Component Analysis 48
4.2.3 Interpreting the basis set expansion 49
4.3 Strategy description 51
4.3.1 Step 1: selecting the training samples 52
4.3.2 Step 2: reducing the model output dimensionality 52
4.3.3 Step 3: constructing the meta-model 52
4.3.4 Step 4: validating the meta-model 54
4.4 Application to the La Frasse case 55
4.4.1 Construction of the meta-model 55
4.4.2 Computation and analysis of the main effects 56
4.5 Towards dealing with functional inputs 58
4.5.1 Strategy description 58
4.5.2 Case study 59
4.5.3 Discussion 60
4.6 Concluding remarks of Chapter 4 62

5 A more flexible tool to represent epistemic uncertainties **67**
5.1 On the limitations of the systematic use of probabilities 67
5.2 Handling vagueness 71

5.2.1 A motivating real-case: hazard related to abandoned underground structures . . . 71
5.2.2 Membership function . . . 72
5.2.3 Application . . . 73
5.3 Reasoning with vagueness . . . 75
5.3.1 A motivating real-case: the inventory of assets at risk . . . 75
5.3.2 Application of Fuzzy Logic . . . 76
5.4 Handling imprecision . . . 79
5.4.1 Possibility theory . . . 79
5.4.2 A practical definition . . . 80
5.4.3 Illustrative real-case application . . . 81
5.5 Handling probabilistic laws with imprecise parameters . . . 82
5.5.1 A motivating example: the Risk-UE (level 1) model . . . 82
5.5.2 Problem definition . . . 83
5.5.3 Use for an informed decision . . . 84
5.6 Concluding remarks of Chapter 5 . . . 86

6 Sensitivity analysis adapted to a mixture of epistemic and aleatory uncertainty 89
6.1 State of the art of sensitivity analysis accounting for hybrid uncertainty representations . . . 89
6.2 A graphical-based approach . . . 93
6.2.1 Motivation . . . 93
6.2.2 Joint propagation of randomness and imprecision . . . 94
6.2.3 Contribution to probability of failure sample plot . . . 97
6.2.4 Adaptation to possibilistic information . . . 98
6.3 Case studies . . . 99
6.3.1 Simple example . . . 99
6.3.2 Case study n°1: stability analysis of steep slopes . . . 100
6.3.3 Case study n°2: stability analysis in post-mining . . . 103
6.3.4 Case study n°3: numerical simulation for stability analysis in post-mining 106
6.4 Concluding remarks for Chapter 6 . . . 111

7 Conclusions 115
7.1 Achieved results . . . 115
7.2 Open questions and Future developments . . . 118
7.2.1 Model uncertainty . . . 118
7.2.2 Use of new uncertainty theories for practical decision-making . . . 121

A Functional decomposition of the variance: the Sobol´ indices **123**

B Universal kriging equations **127**

C Key ingredients of a bayesian treatment of kriging-based meta-modelling **131**
C.1 Principles of Bayesian Model Averaging 131
C.2 Monte-Carlo-based procedures 132
C.3 Bayesian kriging 133
C.4 Deriving a full posterior distribution for the sensitivity indices 135

D Brief introduction to the main uncertainty theories **137**
D.1 Probability 137
D.2 Imprecise probability 138
D.3 Evidence theory 138
D.4 Probability bound analysis 139
D.5 Possibility theory 140

E Fuzzy Random Variable **141**

Bibliography **158**

List of Figures

1 Exemple d'un résultat dérivant dune analyse globale de sensibilité à partir des indices de Sobol' de 1er ordre et indices totaux. . . . viii
2 Approximation d'un modèle 1d par un méta-modèle de type krigeage construit à partir des configurations indiquées par des points rouges. . . . ix
3 Illustration des séries temporelles qui sont en entrée et en sortie du modèle numérique simulant le glissement de La Frasse. . . . xi
4 Exemple d'un intervalle flou pour représenter l'imprécision d'un paramètre incertain à partir de l'information d'expert sur le support et le coeur. . . . xii
5 Analyse de sensibilité avec l'outil graphique pour le cas d'une falaise dont le matériau a un angle de frottement aléatoire ϕ et de hauteur de pied de falaise H_t. xiv

1.1 Example of a uncertainty classification. . . . 6
1.2 La Frasse landslide topview and model. . . . 7
1.3 Generic framework for uncertainty treatment . . . 9

2.1 Infinite slope model. . . . 16
2.2 Uncertainty assumptions for infinite slope model. . . . 17
2.3 Application of OAT on the infinite slope model considering different cases of shifting parameters and reference values. . . . 18
2.4 VBSA for infinite slope model. . . . 19
2.5 Scatterplots for two parameters of the infinite slope model. . . . 20

3.1 Random generation of variables versus LHS (maximin criterion) . . . 27
3.2 Illustration of kriging meta-modelling technique with the infinite slope model. 35
3.3 Horizontal displacements at two observation points of the La Frasse model. . . 37
3.4 Cross-validation of the meta-model for the La Frasse case. . . . 38
3.5 Traceplots for three parameters of the kriging meta-model of the La Frasse model. 39
3.6 Temporal evolution of the main effects of the 1^{st} and 2^{nd} most important parameters in the La Frasse case. . . . 40
3.7 Mean of the main effects at different time instants. . . . 41

4.1 Time series in the La Frasse landslide model. 46
4.2 PCA applied to the La Frasse landslide model (upper part). 50
4.3 Comparison between observed and PCA-reconstructed time series. 51
4.4 PCA applied to the La Frasse landslide model (lower part). 52
4.5 Validation of the functional meta-model of La Frasse. 56
4.6 VBSA applied to the La Frasse case using PCA. 57
4.7 Spatially varying input: example and approximation through PCA. 60
4.8 Differences between the original and the approximated maps. 61
4.9 Choice of the optimal number of PCs in the decomposition of the spatially-varying input. 62
4.10 Maps of the PCs related to the expansion of a spatially-varying input. 63

5.1 Illustration of the limitations of using uniform probability distributions in situations of lack of knowledge. 69
5.2 Graphical representation of the set A under the classical Boolean theory and under the fuzzy set theory. 73
5.3 Construction of the membership function associated to the criterion for cavity susceptibility assessment. 74
5.4 Situation of the French city of Lourdes - Assets at risks and geotechnical zonation. 76
5.5 Methodology for inventory imprecision assessment adapted from the approximate reasoning of [Zadeh, 1975]. 78
5.6 Illustration of a possibility distribution associated with an imprecise parameter (A) and definition of the measure of possibility and necessity (B). 80
5.7 Illustration of the family of probabilistic damage curves associated to the α-cuts of the imprecise parameter r_D. 84
5.8 Synthesis in a pair of probabilistic indicators of all the possible alternatives for the probabilistic damage curves associated with the α-cuts of the imprecise parameter r_D. 85
5.9 Mapping of the lower and upper probabilistic indicator of the event: “exceeding damage grade D4”. 86

6.1 Example of pinching transformation of the fuzzy set assigned to the vulnerability index Vi of the M4 class of vulnerability. 91
6.2 Epistemic indicator as function of the filter width V_f applied to the fuzzy set of M4 vulnerability class. 92

6.3 Main steps of the joint propagation of a random variable represented by a cumulative probability distribution and an imprecise variable represented by a cumulative possibility distribution using the independence Random Set procedure of [Baudrit et al., 2007b]. 95
6.4 CFP curves associated with the random variable ϵ and the imprecise variable X considering different cases of "randomness strength r". 100
6.5 Schematic description of the failure geometry applied in [Collins and Sitar, 2010] to assess stability of steep slopes in cemented sands. 101
6.6 Plausibility and Belief functions resulting from the joint propagation of variability and imprecision in the slope stability analysis. 102
6.7 CFP curves associated with the random and imprecise variables of the slope stability analysis. 103
6.8 Schematic representation of the underground quarry of Beauregard (adapted from [Piedra-Morales, 1991]). 104
6.9 CFP curves associated with the random and imprecise variables of the mine pillar stability analysis. 105
6.10 A) Model geometry and boundary conditions for evaluating the stress evolution at mid height of pillar during loading; B) Map of plastic shear strain at the end of the loading; C) Typical stress evolution during loading. 107
6.11 Leave-One-Out Cross-Validation procedure applied to both kriging-type metamodels for the lower and upper bounds of the average stress. 110
6.12 Upper and Lower probability distribution bounding the true probability assigned to the average stress at the end of loading. 111
6.13 CFP derived from the joint propagation of imprecision and randomness for the pillar case. 112

E.1 Illustration of a fuzzy random variable 142

List of Tables

2.1 Assumptions for uncertainty representation of the infinite slope model. 16

3.1 Range of values for the slip surface properties of the La Frasse landslide. 25
3.2 Main steps of the meta-modelling strategy for global sensitivity analysis using computationally intensive numerical models. 26
3.3 Comparison between the “true” and the estimates of the main effects for the infinite slope analytical model. 36

4.1 Main steps of the meta-modelling strategy for dynamic sensitivity analysis using PCA. 53

5.1 Numerical choices for imprecision assessment considering the inventory of assets at risk at Lourdes. 77
5.2 Logical rules for imprecision assessment considering the inventory of assets at risk at Lourdes. 77
5.3 Imprecision assessment in the district n°18. 79

6.1 Description of the main steps and methods of the joint exploration procedure of possibilistic and probabilistic uncertainty.. 98
6.2 Assumptions on the uncertainty representation for the properties of the rock materials composing the mine pillar. 108

1 Introduction

The present PhD thesis focuses on the treatment of uncertainty in the assessments of "geo-hazards". These types of hazard (see Sect. 1.1 for a definition) are related to geological or geotechnical phenomena like earthquake, landslide, sinkhole, etc. Such hazards are generally categorized as natural hazards, but their origin can be of anthropogenic nature as well, an example being mining subsidences [Deck and Verdel, 2012]. When it comes to making informed choices for the management of geo-hazards (through risk reduction measures, land use planning or mitigation strategies, etc.), the issue of uncertainty is of primary importance (see [Hill et al., 2013] and references therein). Uncertainties should be recognized (i.e. identified) and their implications should be transparently assessed (i.e. propagated), honestly reported and effectively communicated as underlined by [Hill et al., 2013]. In Sect. 1.1, the notion of uncertainty is clarified and defined through its relationship with risk and decision-making. In Sect. 1.2, two facets of uncertainty are outlined, namely "aleatory" and "epistemic uncertainty". The latter facet is at the core of the present work, and more specifically the lack of knowledge, designated as "parameter uncertainty" in the following (Sect. 1.3). This type of epistemic uncertainty is illustrated with the real-case of the La Frasse landslide (Sect. 1.4). On this basis, the major research questions are raised, which constitute the lines of research of the present work (Sect. 1.5).

1.1 Hazard, Risk, uncertainty and decision-making

Natural risks can be understood as the combination of hazard and of vulnerability. A hazard can be defined as "a potentially damaging physical event, phenomenon or human activity that may cause the loss of life or injury, property damage, social and economic disruption or environmental degradation" [UN/ISDR, 2004]. Vulnerability can be defined as "the conditions determined by physical, social, economic, and environmental factors or processes, which

increase the susceptibility of a community to the impact of hazards" [UN/ISDR, 2004].

However, it should be underlined that the definition of risk is not unique and the glossary on components of risk provided by the United Nations University [Thywissen, 2006] is composed of more than 20 definitions. In all those definitions, the situation of risk is always understood relative to the situation of uncertainty using terms that can, in a broad sense, be related to the concept of uncertainty (expectation, probability, possibility, unknowns, etc.). Besides, it is interesting to note that the ISO standard ISO 31000:2009 on risk management defines risk as the "effect of uncertainty on objectives".

[Knight, 1921] provided an original vision on the relationship between risk and uncertainty by formally distinguishing both concepts as follows: in a situation of risk, the probability of each possible outcome can be identified, whereas in a situation of uncertainty, the outcome can be identified, but not the corresponding probabilities ([Knight, 1921], quoted by [Bieri, 2006]). Traditionally, rational decision-making under uncertainty is based on probabilities using the Independence Axiom introduced by [Von Neumann and Morgenstern, 1944] and extended by [Savage, 1954]. Under very general conditions, the independence axiom implies that the individual objective is linear in probabilities. This leads to the subjective expected utility theory under which decision support should only be guided by the values of probabilities.

Yet, within this formalism, the nature and quantity of information that have led to the estimate of the probability values does not influence the decision. As underlined by [Paté-Cornell, 2002], according to this theory, the rational decision maker is indifferent to two sources of information that result in the same probabilistic distribution of outcomes, i.e. regardless of whether they result from experiments based on flipping a coin, or following a "rain tomorrow" approach (let say, based on the "weatherman's opinion"). [Keynes, 1921] originally provided a view on this issue by distinguishing between probability and "weight of evidence", so that probability represents the balance of evidence in favour of a particular option, whereas the weight of evidence represents the quantity of evidence supporting the balance. According to this view, people should be more willing to act if the probability of an outcome is supported by a larger weight of evidence, i.e. the situation is less "ambiguous". The work of [Ellsberg, 1961] has provided experimental evidence that people do not behave in the same way in the face of two uncertain environments with the same probabilities, but with different weights of evidence (i.e., different degrees of ambiguity). In their well-known classic experiment, subjects prefer to take a chance on winning a prize with draws from an urn with a specified mixture of balls as opposed to taking a chance with a subjective probability that is equivalent, but ambiguous. Since these original works, extensive work have been carried out to better understand the complex relationship between uncertainty, ambiguity and information and how this affects

decision-making (see, e.g., [Cabantous et al., 2011] and references therein). Consequently, recent studies on risk analysis have outlined that the relationship between risk and uncertainty through probabilities may be too restrictive. For instance, [Aven and Renn, 2009] define risk in a broader sense as the "uncertainty about and sensitivity of the consequences (and outcomes) of an activity with respect to something that humans value".

1.2 Aleatory and Epistemic uncertainty

Giving a single "fit-to-all" definition for uncertainty remains difficult, because uncertainty can be interpreted differently depending on the discipline and context where it is applied, as outlined for instance by [Ascough(II) et al., 2008] in environmental and ecological studies. Therefore, several authors [van Asselt and Rotmans, 2002, Rogers, 2003, Walker et al., 2003, Baecher and Christian, 2005, Cauvin et al., 2008], among others, adopt a less ambitious (but more practical) approach by defining uncertainty through classification. Such an approach presents the appealing feature of enabling the risk practitioners to differentiate between uncertainties and to communicate about them in a more constructive manner.

Though differing from one classification to another, they have all in common to distinguish two major facets of uncertainty, namely "aleatory uncertainty" and "epistemic uncertainty". In the domain of natural hazards, the benefits of distinguishing both facets have been outlined for geo-hazards by [Deck and Verdel, 2012], and more specifically for seismic risk by [Abrahamson, 2000], for rockfall risk by [Straub and Schubert, 2008], for volcano risk by [Marzocchi et al., 2004].

- The first facet corresponds to aleatory uncertainty/variability (also referred to as randomness). The physical environment or engineered system under study can behave in different ways or is valued differently spatially or/and temporally. The aleatory variability is associated with the impossibility of predicting deterministically the evolution of a system due to its intrinsic complexity. Hence, this source of uncertainty represents the "real" variability and it is inherent to the physical environment or engineered system under study, i.e., it is an attribute/property;
- The second facet corresponds to epistemic uncertainty. This type is also referred to as "knowledge-based", as the latin term *episteme* means knowledge. Contrary to the first type, epistemic uncertainty is not intrinsic to the system under study and can be qualified as being "artificial", because it stems from the incomplete/imprecise nature of available information, i.e., the limited knowledge of the physical environment or engineered system under study. Epistemic uncertainty encompasses a large variety of forms of uncertainty, which are clarified in Sect. 1.3;

It is worth stating that both sources of uncertainty can be inter-connected so that the study of a stochastic system is by its nature pervaded by randomness, but the resources to measure and obtain empirical information on such a stochastic system can be limited, i.e., uncertainty can exhibit both variability and a lack of knowledge (epistemic uncertainty). Conversely, this aspect should not exclude the situation where knowledge with regard to deterministic processes can also be incomplete [van Asselt and Rotmans, 2002], e.g., a wellbore has been drilled, but its depth has not been reported so that it remains imprecisely known.

It should be recalled that the objective here is not to reopen the widely discussed debate on the relevance of the separation of uncertainty sources (see e.g., [Kiureghian and Ditlevsen, 2009]). Here, the scope is narrower and it is merely underlined that the efforts to separate both sources, though appearing as a "pure modelling choice", should be seen from a risk management perspective as discussed by [Dubois, 2010]:

- Aleatory uncertainty, being a property of the system under study, cannot be reduced. Therefore, concrete actions can be taken to circumvent the potentially dangerous effects of such variability. A good illustration is the reinforcement of protective infrastructures such as the height of dykes to counter in a preventive fashion the temporal variations in the frequency and magnitude of storm surges. Another option can be based on the application of an additional "safety margin" for the design of the engineered structure;
- Epistemic uncertainty, being due to the capability of the analyst (measurement capability, modeling capability, etc.), can be reduced by, e.g., increasing the number of tests, improving the measurement methods or evaluating calculation procedure with model tests. In this sense, this type of uncertainty is referred to as "knowledge-based" [Kiureghian and Ditlevsen, 2009]. Given the large number of uncertainty sources, the challenge is then to set priorities, under budget/time constraints, on the basis of the identification of the most influential/important sources of uncertainty: this is at the core of the present PhD thesis.

From this viewpoint, it should be kept in mind that even in situations where a lot of information is available, uncertainty can still prevail, either because the system under study is by essence random (although epistemic uncertainty may have vanished through the increase of knowledge) or because new knowledge has illuminated some “not-yet-envisaged” complex processes, of which our understanding is still poor. This can be illustrated with the earthquake on January 26^{th} 2011 at Christchurch (New Zealand): an extreme shaking of 2.2 g was recorded whereas the magnitude is moderate of M_w=6.2 [Holden, 2011]. Hence, the treatment of uncertainty is not only a matter of knowledge-gathering, but “the fundamental imperfection of knowledge is the essence of uncertainty” ([Shackle, 1955] quoted by

[van Asselt and Rotmans, 2002]). Addressing uncertainty should therefore guide actions for risk management by identifying and ranking issues worthy to be tackled within the risk assessment procedure, which may concretely consist of collecting new data, but may also involve rethinking the assessment procedure, e.g., by comparing different points of views / experts' judgments.

1.3 Epistemic uncertainty of type "parameter"

Epistemic uncertainty encompasses too many aspects to be practically used as a unique concept, so that for the purposes of a risk analysis, [Cauvin et al., 2008] have suggested distinguishing between four classes of uncertainty (Fig. 1.1 adapted from [Deck and Verdel, 2012]).

- Resources uncertainty deals with knowledge about both the general scientific context of the study and its local particularities. More specifically, it concerns the existence of information about the processes being investigated and the objects being studied;
- Expertise uncertainty is related to all the choices, actions or decisions that can be made by the expert to carry out the risk study. It mainly relies on his/her particular experience as an individual, on his/her subjectivity and on the way he/she represents and interprets the information he/she has gathered;
- Model uncertainty is basically induced by the use of tools to represent reality and is related to the issue of model representativeness and reliability. This type of uncertainty is also named structural uncertainty defined as the failure of the model to represent the system even if the correct parameters are known [Hill et al., 2013];
- Data uncertainty represents both the natural variability existing in the data, the lack of knowledge about their exact values and the difficulty of clearly evaluating them.

The incomplete knowledge pervading the parameters of the models supporting geo-hazard assessment is at the core of the present work: this corresponds to the fourth category "Data uncertainty" of [Cauvin et al., 2008]. As recently described by [Hill et al., 2013], this type of epistemic uncertainty both encompasses parametric uncertainty (incomplete knowledge of the correct setting of the model's parameters) and input uncertainty (incomplete knowledge of true value of the initial state and the loading). For sake of simplicity, this type of epistemic uncertainty is designated with the generic term "parameter uncertainty" in the present work. The term "parameter" is indifferently used to refer to the system's initial state (e.g., initial stress state at depth), to the loading/forcing acting on the system (e.g., changes of groundwater table) and to the system's characteristics (e.g. soil formation's property).

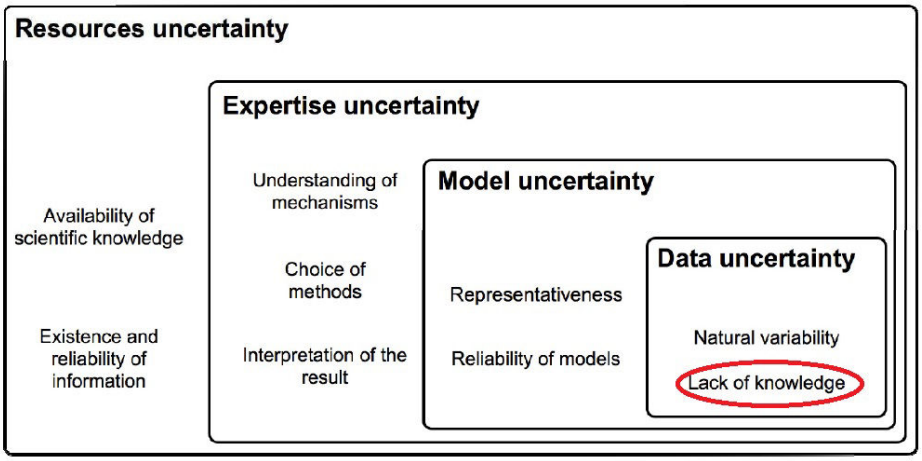

Figure 1.1: Uncertainty classification in geo-hazard assessments as proposed by [Cauvin et al., 2008]. The "lack of knowledge" of the category "Data uncertainty" is at the core of the present work. It is referred to as "parameter uncertainty" in the present work.

1.4 A real-case example

To further clarify the concept of parameter uncertainty, let us consider the landslide of "La Frasse" (Swiss), which has been studied by Laloui and co-authors [Laloui et al., 2004, Tacher et al., 2005]. This landslide with active mass of ≈ 73 million m^3, is located in the Pre-alps of the Canton of Vaud in Switzerland (at ≈ 20 km east from Lake Geneva) and has experienced several crises in the past, during which a maximum observed velocity of 1 m/week could be observed in the lower part of the landslide. An overview of the landslide is provided in Fig. 1.2A). The evolution of the groundwater table is considered to be at the origin of the sliding and the instabilities were mainly observed during the 1994 crisis (over a period of nearly 300 days). Therefore, in order to assess the effect of the hydraulic regime on the geomechanical behaviour of the landslide, finite-element simulations considering a 2D cross-section through the centre of the landslide were performed by [Laloui et al., 2004] using the finite element program GEFDYN by [Aubry et al., 1986]. The model is composed of 1,694 nodes, 1,530 quadrangular elements, and six soil layers derived from the geotechnical investigations. Figure 1.2B) gives an overview of the model, as well as the boundary conditions used for analysis. Instabilities observed in 1994 were triggered by pore pressure changes occurring at the base of the slide (see [Laloui et al., 2004] for further details).

The numerical model used for predicting the hydro-mechanical behaviour of the landslide involves a large variety of assumptions. The model involves model uncertainties related to:

- the system's geometry: use of a two-dimensional cross section, spatial location of the slip surface, definition of six soil formations;

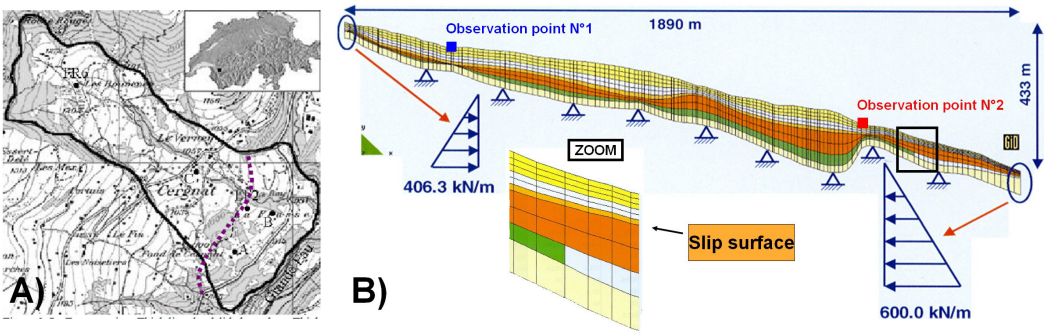

Figure 1.2: A) Topview of the La Frasse landslide in Switzerland. B) Overview of the two-dimensional finite-element model used for assessing the hydro-mechanical behaviour of the La Frasse landslide (adapted from [Laloui et al., 2004]).

- the soil formations' behaviour: spatially homogeneous properties, choice in the constitutive law (Hujeux for the slip surface's material [Hujeux, 1985] and Mohr Coulomb for the others);

The model involves parameter uncertainties related to:

- the loading/forcing conditions of the system: the temporal evolution of the water level, location of the flow changes, nature of the boundary conditions (e.g., nil normal displacements and nil flow at the bottom);
- the properties' values related to: the density, the initial stress state, the elastic behaviour (Young's modulus, Poisson's ratio), the plastic behaviour (internal friction angle, cohesion, dilatancy angle) and the flow behaviour (porosity, horizontal and vertical intrinsic permeability) of the six soil formations.

This real-case provides an example of an advanced numerical model used to support geo-hazard assessment. This shows that such models can involve a large number of sources of uncertainty. Considering only the input parameters, the La Frasse landslide model involves more than 50 parameters, i.e. more than 50 sources of uncertainty (not to mention the model uncertainties). A challenge for an efficient uncertainty treatment is then to reduce the number of uncertain parameters.

1.5 Objectives and structure of the manuscript

In the light of the afore-described real-case, the following question can be raised: among all the sources of parameter uncertainty, which of them have the greatest influence on the uncertainty associated to the results of the geo-hazard assessment? Which sources of uncertainty are the most important and should be taken into account in priority in the analysis?

How to rank the uncertain parameter in terms of importance? Conversely, which sources of uncertainty can be treated as insignificant and thus can be neglected in the analysis, i.e. how to simplify the model? Finally, on what parameters should the characterization effort be put in priority (additional lab tests, in site experiments, numerical simulations)? How to optimize the allocation of the resources for hazard assessment? Thus, the central research question is the importance ranking of parameter uncertainties (epistemic).

Addressing this question is of great interest in situations where the resources (time and budget) for hazard and risk assessments are generally limited. An example is the development of Risks Prevention Plans, which are powerful operational and statutory tools [MATE, 1999], but their practical implementation can be tedious as they impose to work "in the state of knowledge" and "according to expert opinion", i.e. with no other resources that those available at the time of the study, which, in practice, are generally limited as outlined by [Cauvin et al., 2008].

From a methodological perspective, this question falls into the goal for quantitative uncertainty assessment termed as "Understand" [de Rocquigny et al., 2008]: "To understand the influence or rank importance of uncertainties, thereby guiding any additional measurement, modelling or research efforts with the aim of reducing epistemic uncertainties". This question is then related to the step of sensitivity analysis of the generic framework for uncertainty treatment (Fig. 1.3). The description of this generic framework is mainly based on the recent best practices for uncertainty assessment ([de Rocquigny et al., 2008] in collaboration with the European Safety Reliability Data Association). While sensitivity analysis (step 2') focuses on the study of "how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input" [Saltelli et al., 2008], the related practice of "uncertainty analysis" (step 2) focuses on the quantification of uncertainty in the model output. Step 2 and 2' are usually run in tandem.

Several techniques exist in the literature to address the question of sensitivity analysis. Chapter 2 first provides a brief overview of the main techniques for sensitivity analysis, and then, focuses on the most commonly-used and most advanced tools for conducting sensitivity analysis in a global manner (Global Sensitivity Analysis GSA), namely techniques relying on the decomposition of the variance in a probabilistic setting VBSA [Saltelli et al., 2008]. Using a simple analytical model for landslide hazard assessment, this first chapter highlights how VBSA can be useful to answer the question related to importance ranking. On the other hand, this first chapter also highlights and discusses the main constraints for its practical implementation in the context of geo-hazard assessments, namely:

1. Though providing very valuable information, the drawback of VBSA is its computational

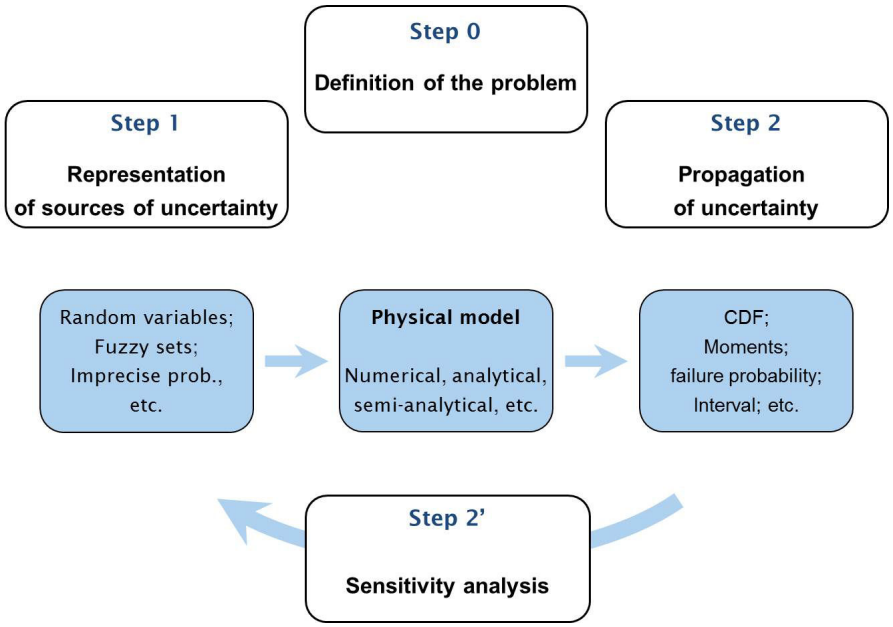

Figure 1.3: Main steps of the generic framework for uncertainty treatment (adapted from [de Rocquigny et al., 2008]).

efficiency: it requires running the model (numerical or otherwise) supporting the geo-hazard assessment a large number of times (> 1,000). This can pose difficulties in practice when using long-running numerical code like the one presented for the La Frasse landslide: it has a computation time cost of about four days. The issue of handling such computationally intensive hazard assessments is addressed in Chapter 3 based on the work described in [Rohmer and Foerster, 2011];

2. The inputs and outputs of model for geo-hazard assessments can be complex in the sense that in most cases they are not only scalar (i.e. that can hold only one value at a time), but they can vary in time (e.g. the surface displacements at the lower part of the La Frasse landslide), in space (e.g. hydraulic conductivities of the soil formations like in the La Frasse case, see [Tacher et al., 2005] or both (e.g. spatial displacements along the surface of the La Frasse landlslide). This second issue is addressed in Chapter 4 and is based on the work described in [Rohmer, 2013];

3. A major pillar of VBSA is the probabilistic setting: uncertainty is represented using probability distributions and the sensitivity measures are defined using variance. Yet, in many cases and particularly in the field of geo-hazards (e.g. [Karimi and Hüllermeier, 2007, Cauvin et al., 2008], the data available for hazard assessment can be particularly scarce and is often associated with imprecision and incompleteness, which is often due to spatial and financial constraints. It is in this context that the expert's role is essential. Based

on his/her experience and his/her regional knowledge, the expert synthesizes and interprets the commonly imprecise, if not vague, information obtained from the inventory and the geological and historical-economic contexts. See an example for susceptibility assessment of abandoned underground structures by [Nachbaur and Rohmer, 2011]. Again, this can be illustrated with the Risk Prevention Plans, which imposes to work "in the state of knowledge" and "according to expert opinion". In these situations, the systematic use of probabilities for uncertainty representation can be questioned. Such limitations are discussed in Chapter 5. The applicability of an alternative mathematical tool, namely the use of Fuzzy Sets [Zadeh, 1965], for representing and processing the uncertain information is investigated and demonstrated through real-case examples derived from [Rohmer and Baudrit, 2011, Nachbaur and Rohmer, 2011]. Finally, Chapter 6 addresses the question of how to conduct sensitivity analysis for mixed uncertainty mathematical representations, namely using probabilities and interval-valued tools (like Fuzzy sets). This chapter is mainly based on the work described in [Rohmer and Verdel, 2014].

In summary, the present PhD thesis addresses three difficulties for importance ranking of parameter uncertainties using models supporting geo-hazard assessments: 1. the large computation time cost; 2. the functional (complex) nature of the parameters; 3. the knowledge situation characterised by information imprecision, if not vagueness, and data scarcity.

2 A probabilistic tool: variance-based global sensitivity analysis

In this chapter, the problem of importance ranking of parameter uncertainties is handled with the probabilistic tools of global sensitivity analysis GSA. In this view, the basic concepts of GSA are firstly introduced (Sect. 2.1) and the most advanced procedures, namely variance-based methods VBSA are then described (Sect. 2.2). These are then applied on a simple analytical model used for slope stability analysis (Sect. 2.3). On this basis, limitations for the direct application of such methods to model supporting geo-hazard assessments are then discussed (Sect. 2.4). This discussion allows defining three main research questions at the core of the present work, which will be respectively addressed in Chapters 3 to 6.

2.1 Global sensivity analysis

Sensitivity analysis aims at studying of "how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input" [Saltelli et al., 2008]. This step has clearly been outlined in numerous regulatory frameworks such as the Impact Assessment Guidelines of [European Commission, 2009], which clearly specifies that "Sensitivity analysis can be used to explore how the impacts of the options you are analysing would change in response to variations in key parameters and how they interact".

There is a large variety of different methods to address this question in the literature: see the review provided by the European Commission Joint Research Centre JRC-IPSC (Italy).[1] and [Saltelli et al., 2008, Pappenberger et al., 2010, Iooss, 2011].

The main existing techniques are:

- One-factor-At-a-Time (described below);

[1] Available online at: http://ipsc.jrc.ec.europa.eu/?id=755)

- Experimental designs [Kleijnen, 2005];
- Morris screening method [Morris, 1991];
- Regression-and correlation-based sensitivity measures [Saporta, 2011];
- Sampling and scatter-plot-based techniques [Helton et al., 2006b];
- Functional decomposition of variance: Sobol' indices [Sobol', 1990].

The most widespread method remains the "One-factor-At-a-Time" (OAT) approach (also named local sensitivity analysis). This consists in analysing variations from a base model, i.e. exploring changes in the results by varying in turn the input parameters or considering different scenarios. Applications of OAT to investigate the sensitivity to input parameters related to soil / rock properties (e.g., density, cohesion, angle of internal friction, etc.) or to slide characteristics (e.g., sizes, failure mechanisms, etc.) is illustrated, for instance, by [Gorsevski et al., 2006] for landslide susceptibility models. An example in the field of seismic risk is provided by [Crowley et al., 2005].

Though the implementation of OAT is simple and can rapidly provide valuable information for importance ranking, OAT presents several shortcomings as pointed out in the statistical literature (see [Saltelli and Annoni, 2010] and references therein):

- the sensitivity measures are valid for a specific reference (base) case. Modifying this reference case obviously influences the ranking of the uncertain parameters. Thus, OAT-based sensitivity measures only provide "local" information, but do not provide information regarding the rest of the domain of variation of the other input parameters. More specifically, [Saltelli and Annoni, 2010] used a geometric argument to demonstrate the low efficiency of the method to explore the inputs' space: the region partially explored by OAT rapidly decreases to zero with increasing the number of input parameters. Besides, they showed that the points of the OAT design remain "stuck" in the neighbourhood of a central point. Finally, they outlined that the higher the number of input parameters, the less OAT is capable of following the shape of the real cumulative probability distribution (CDF) of the model output;
- the parameters' values are deterministically modified following a shifting procedure (e.g., input parameter n°1 is shifted by 10 percent of its initial value). In addition to the problem of arbitrarily choosing the shift value, the method also hardly accounts for information on the probability distributions associated to the parameters (form, shape, etc.), hence neglects that "some values are more likely to come up than others";
- the parameters' values are modified in turn, hence making it impossible to capture any possible interactions between the parameters. A simple example is the two-dimensional linear model with cross-terms products of the form: $x_1 + x_2 + x_1 \times x_2$. Methods based

on experimental designs (e.g., two-level factorial design or more sophisticated ones), as carried out by [Abdallah, 2009], can overcome such a limitation. Yet, such methods show good performance when an a priori idea on the nature of the considered model (i.e. its mathematical structure like linearity, monotonicity, etc.) is available. These methods are said to be "model-dependent".

A more flexible class of methods is thus desirable as advocated by [Saltelli et al., 2008] in order to: i. incorporate the effect of the range of the input variation and of the nature of the probability distribution (normal, uniform, etc.); ii. explore the sensitivity over the whole range of variation (i.e. in a global manner) of the input random variables; iii. to fully account for possible interactions between them, and to provide a quantitative measure of sensitivity without introducing a priori assumptions on the model's mathematical structure (i.e. model-free). This message is clearly outlined by the American Environmental Protection Agency EPA [Environmental Protection Agency, 2009]: "[Sensitivity Analysis] methods should preferably be able to deal with a model regardless of assumptions about a model's linearity and additivity, consider interaction effects among input uncertainties, [...], and evaluate the effect of an input while all other inputs are allowed to vary as well". This objective can be fulfilled by Global Sensitivity Analysis GSA relying on variance-based sensitivity analysis VBSA as described by [Saltelli et al., 2008].

2.2 Variance-based global sensivity analysis

In this section, we introduce the basic concepts underlying variance-based sensitivity analysis VBSA, which can be considered among the most advanced methods to fulfill the requirements of GSA. The basic concepts of VBSA are first briefly introduced in the present section. For a more complete introduction, the interested reader can refer to Appendix A.

Let us define f as the model (numerical, analytical) used to support geo-hazard assessment. Considering the n-dimensional vector $\boldsymbol{X}$ as a random vector of independent random variables X_i with $i = 1, 2, \ldots, n$, then the output $Y = f(\boldsymbol{X})$ is also a random variable (as a function of a random vector). VBSA aims at determining the part of the total unconditional variance $\mathrm{Var}(Y)$ of the output Y resulting from the variation of each input random variable X_i, here assumed to be independent. Considering that the variance can adequately capture the uncertainty, this analysis relies on the functional analysis of variance (ANOVA) decomposition of f based on

which the Sobol' indices (ranging between 0 and 1) can be defined:

$$\begin{aligned} S_i &= \frac{Var(E(Y|X_i)}{Var(Y)} \\ S_{ij} &= \frac{Var(E(Y|X_i,X_j)}{Var(Y)} - S_i - S_j \end{aligned} \tag{2.1}$$

The first-order S_i is referred to as "the main effect of X_i" and can be interpreted as the expected amount of Var(Y) (i.e. representing the uncertainty in Y) that would be reduced if it was possible to learn the true value of X_i. This index provides a measure of importance useful to rank in terms of importance the different input parameters within a "factors' prioritizing setting" [Saltelli et al., 2008]. The second order term S_{ij} measures the combined effect of both parameters X_i and X_j. Higher order terms can be defined in a similar fashion. The total number of sensitivity indices reaches $2^n - 1$. In practice, the sensitivity analysis is generally limited to the pairs of indicators corresponding to the main effect S_i and to the total effect S_{T_i} of X_i (Saltelli et al., 2008). The latter is defined as follows:

$$S_{T_i} = 1 - \frac{Var(E(Y|X_{-i})}{Var(Y)} \tag{2.2}$$

where $X_{-i} = (X_1, ..., X_{i-1}, X_{i+1}, ..., X_n)$. The total index corresponds to the fraction of the uncertainty in Y that can be attributed to X_i plus its interactions with all other input parameters. $S_{T_i} = 0$ means that the input factor X_i has no effect so that X_i can be fixed at any value over its uncertainty range within a "factors' fixing" setting (as described in [Saltelli et al., 2008]).

Different algorithms are available for the estimation of the Sobol' indices like (extended) Fourier Amplitude Sensitivity Test (E)FAST [Saltelli et al., 1999], algorithms based on Monte-Carlo sampling as the one developed by [Sobol', 1990], and more advanced techniques like the one described by [Saltelli, 2002, Saltelli et al., 2010]. A more extensive introduction is provided in the fourth chapter of [Saltelli et al., 2008].

More recently, it has been shown how this technique improves insight into the nature of the considered model through the notion of effective dimension [Kucherenko et al., 2011], which can be understood as the number of dominant input parameter in a given model. The relationship between the main and total effects helps exploring the model complexity. Several cases have been formalized by [Kucherenko et al., 2011]. In particular, if $S_i \approx S_j$ for any i and j and $S_i \approx S_{T_i}$ then the model has equally important variables, but with "weak" interaction

among them (recall that two input parameters are said to interact if their effect on the model output cannot be expressed as a sum of their single effects). Conversely, if $S_i \approx S_j$ for any i and j and $S_i \ll S_{T_i}$ then the model has equally important variables, but with "strong" interaction among them. Furthermore, if the sum of the main effects equals one, this indicates that the model is additive, in the sense that the model output y can be decomposed as a sum of one-dimensional functions f_i of the input parameters x_i as follows:

$$f(x) = \sum_{j=1}^{n} f_j(x_j) \tag{2.3}$$

where the uni-variate (one-dimensional) functions f_i can be linear or of greater "complexity" (non linear) such as polynomial, or splines, etc. This means that the different input parameters of f do not interact. This feature can be of great value especially when using a model for geo-hazard assessment in a "black-box" fashion (see discussion provided by [Bommer et al., 2006] for loss models in the field of seismic risk).

2.3 Application to slope stability analysis

Let us consider a simple application of VBSA using a commonly-used model to assess landslide susceptibility, namely the infinite slope analytical model (e.g. [Hansen, 1984]). The stability of the infinite slope model as depicted in Fig. 2.1 is evaluated by deriving the factor of safety SF, which corresponds to the ratio between the resisting and the driving forces acting on the slope (Eq. 2.4). If SF is lower than 1.0 the potential for failure is high.

$$SF = \frac{C + (\gamma - m \cdot \gamma_w) \cdot z \cdot \cos(\theta) \cdot \tan(\phi)}{\gamma \cdot \sin(\theta) \cdot \cos(\theta)} \tag{2.4}$$

The model parameters (as indicated in Fig. 2.1) correspond to C, the cohesion of the soil material; ϕ, the friction angle; θ, the slope angle; γ, the soil unit weight; z, the thickness of slope material above the slip plane; and m, the ratio between thickness of superficial saturated slope material and z. The water unit weight γ_w is considered constant at 9.81 kN/m^3.

Assumptions on the representation of the uncertain input parameters are summarised in Table 2.1 and the corresponding cumulative probability distributions are depicted in Fig. 2.2.

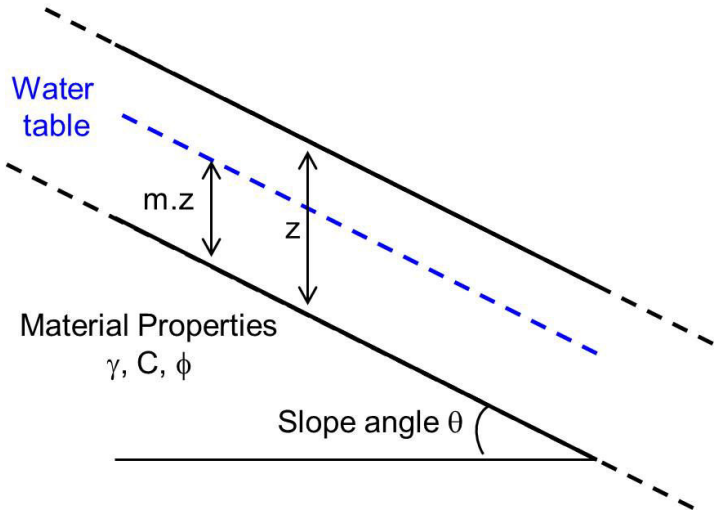

Figure 2.1: Schematic representation of the infinite slope model (adapted from [Hansen, 1984]).

Table 2.1: Assumptions for uncertainty representation of the infinite slope model.

N°	Parameter	Symbol	Probability law	Parameters	Unit
1	Cohesion	C	Uniform	[5;15]	kPa
2	Friction angle	ϕ	Gaussian	Mean=35,Variance=2	°
3	Unit weight	γ	Gaussian	Mean=22,Variance=2.5	kPa
4	Slope angle	θ	Log-normal	Mean=30,Variance=1.35	°
5	Ratio of thicknesses	m	Uniform	[30;90]	%
6	Thickness	z	Uniform	[10;25]	m

2.3.1 Local sensitivity analysis

Let us first apply the most simple procedure for sensitivity analysis, namely OAT as described in Sect. 2.1. In this view, it is necessary to define a reference case and how each parameter's value is shifted. Four cases are considered:

- Case A: the reference case is defined using the mean value of each parameter in Table 2.1. Each parameter are shifted by +/- one standard deviation for normally (or log-normally) distributed parameters or by the half the width of the support for uniformly distributed parameters;
- Case B: same reference case and shifting parameters as Case A, except for parameter n°4 (slope angle θ) with reference value at 25°;
- Case C: same reference case and shifting parameters as Case A, except for parameters with non-uniform probabilistic law, i.e. parameters n°2 to 4, whose range of variation is defined by twice their standard deviation;

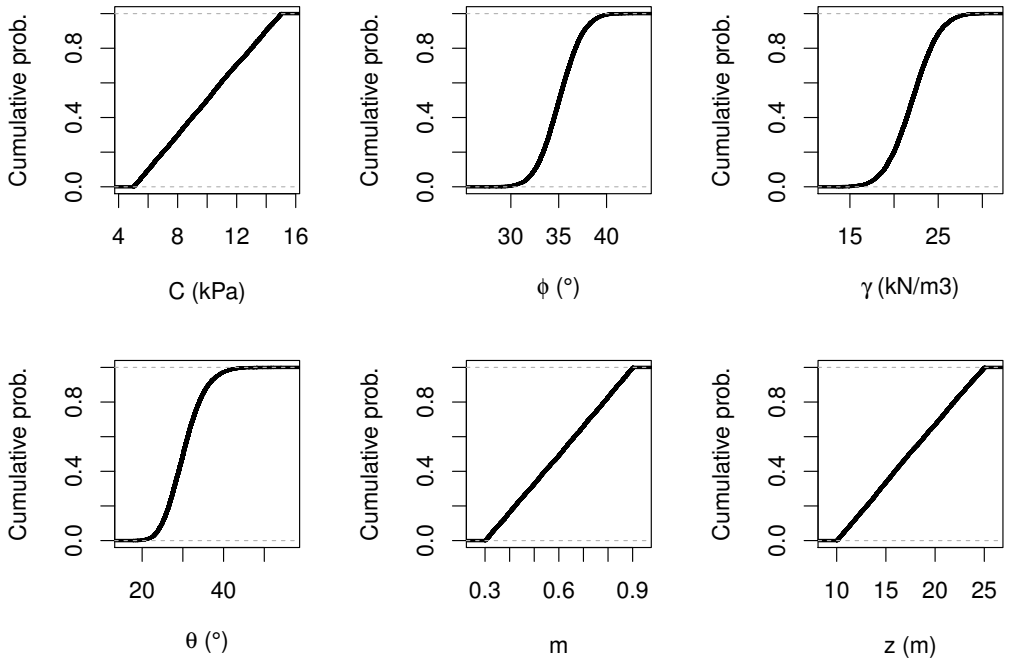

Figure 2.2: Cumulative probability distribution assigned to each uncertain input parameter of the infinite slope model.

- Case D: same reference case and shifting parameters as Case A, except for parameters n°2 to 4, whose range of variation is defined by half their standard deviation.

The results of the OAT-based sensitivity analysis are depicted in Fig. 2.3 in terms of differences with the reference case. The most important parameters are outlined by the horizontal lines with the largest length. Case A indicates that the most important parameters are θ and m (respectively parameter n°4 and n°5) with very close values of sensitivity measures. The ranking appears to be in agreement with VBSA (see Fig. 2.4), but the value of the sensitivity measure assigned to m appears to be overestimated by OAT. Case B shows that the value of the sensitivity measure is dependent on the reference value: by modifying the reference value of parameter n°4, its sensitivity is increased compared to Case A. Cases C and D illustrate the difficulty in defining the shifting parameters when the parameters have infinite support: OAT only works when the support are bounded, hence imposing to arbitrarily define an interval in such a situation. On the one hand, case C indicates that parameter n°2 and n°5 may have equivalent influence and that the most important parameter is the fourth one. On the other hand, Case D indicates that the most important parameter is the fifth one with a sensitivity measure larger than the one of parameter n°2. This illustrates the sensitivity of OAT to the definition of the shifting parameters and of the base case, which can lead to different ranking

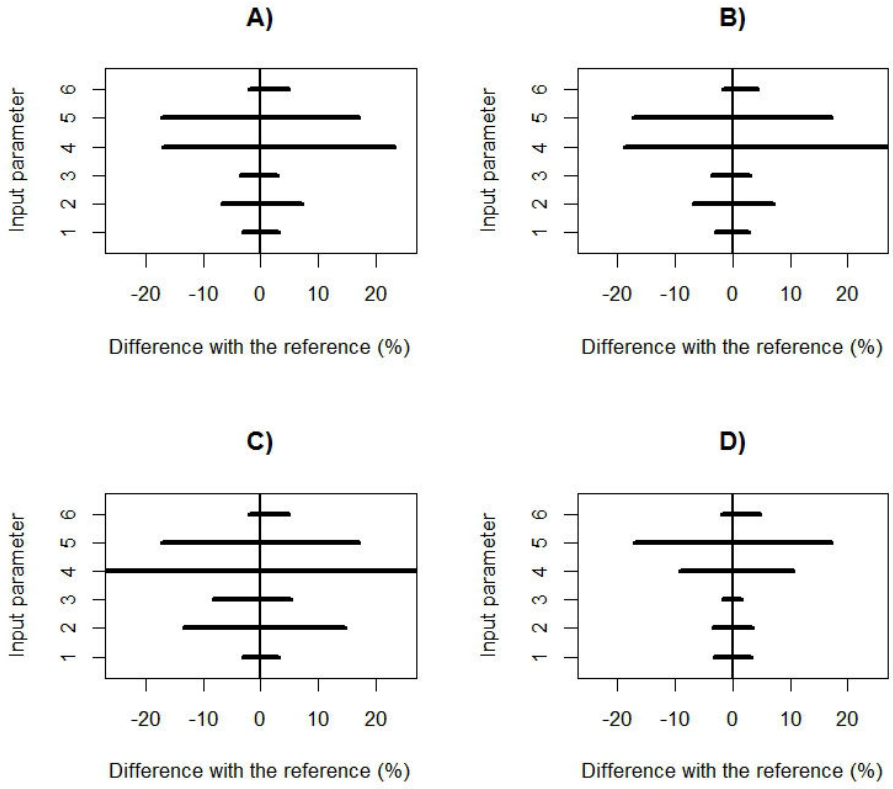

Figure 2.3: Application of OAT on the infinite slope model considering different cases of shifting parameters and reference values. The number of the input parameters are indicated in Table 2.1. See Sect. 2.3.1 for details.

or quantitative sensitivity measures.

2.3.2 Global sensitivity analysis

Main and total effects are computed using the Monte-Carlo-based algorithm developed by [Saltelli, 2002], which requires $N \cdot (n+2)$ model runs (N is the number of Monte-Carlo samples and n is the number of input parameters). Sampling error, due to the Monte-Carlo evaluation of the variances in the definition of Sobol' indices (Eqs. 2.1 and 2.2), are estimated through a confidence interval calculated using 100 bootstrap samples as proposed for instance by [Archer et al., 1997]. Preliminary convergence tests showed that N=20,000 both yields satisfactory convergence of the sensitivity measures to one decimal place and non-overlapping confidence intervals defined at a level of 90 %. To illustrate, results are depicted in Fig. 2.4 for respectively 2,500 and 20,000. In practice the package "sensitivity" [2] of the R software

[2] available at: http://cran.r-project.org/web/packages/sensitivity/index.html

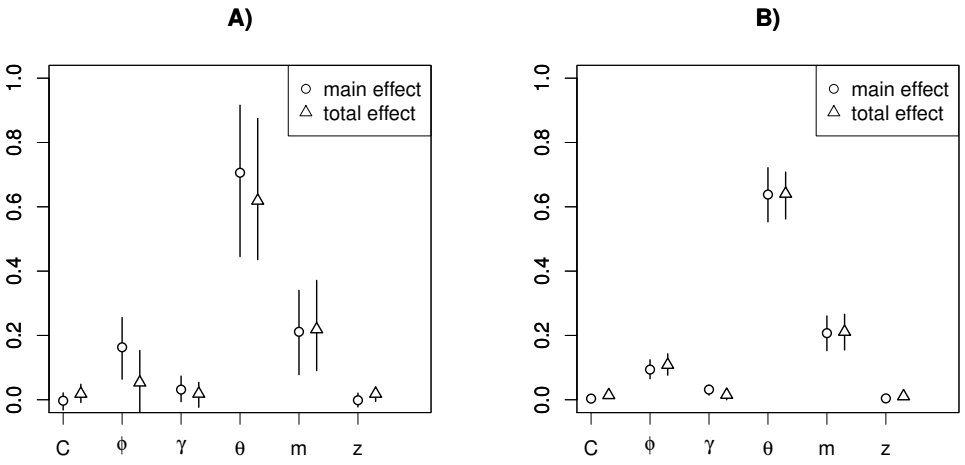

Figure 2.4: Main and total effects estimated for the six uncertain parameters of the infinite slope model considering two sets of Monte-Carlo samples with size: A) 2,500; B) 20,000.

[R Core Team, 2014] was used.

Several observations can be made:

- the total variance on the safety factor reaches 0.045: this represents the whole uncertainty on SF given the uncertainty on the input parameters (Fig. 2.1);
- the total number of model runs reaches $20,000 \times (6+2) = 160,000$. The convergence of the total effects was more difficult than the one for the main effects;
- considering the value of the main effects, θ appears to be the dominant input parameter with a sensitivity measure exceeding 60 %, meaning that more than 60 % of the uncertainty on SF (here its variance) are driven by this property;
- the ranking of the epistemic uncertainties are: θ, m, and ϕ. These are the input parameters, which contribute the most to the variance of SF. Priority should be given to them for the characterisation studies;
- considering the total effects, the negligible input parameters (i.e. with total effects of almost zero), correspond to: C, γ and z. These parameters can be fixed at any value over their ranges of variation with little effect on the model output. The negligible effect of such parameters should not be understood in an absolute manner, but relatively to the influence of the three other parameters and given the assumptions made on their uncertainty representation and variation. Interestingly, [Sobol et al., 2007] demonstrated that the magnitude of the error when fixing a given input parameter at its nominal value is greater or equal to its total index. If the specific input parameter is randomly and

uniformly distributed, then the expectation of this error is exactly twice its total index. For instance, if z was fixed, the resulting error would reach around 6 %;

- the sum of the main effects approximately reach ≈ 98%: this indicates that the interactions between parameters are very weak (here of the order of 2 %). This is also shown by the fact that the main effects nearly equal the total indices. The model is nearly additive, i.e. it could reasonably be approximated by a sum of uni-variate functions as described by Eq. 2.3. For instance, the relationship between SF and respectively m and θ is shown on the scatterplot of Fig. 2.5, constructed using 1,000 random samples of the input parameters. A linear regression model and a loess (local polynomial regression fitting) smoother model are fitted to each these scatterplots (respectively green and red-coloured lines in Fig. 2.5). The non-linear relationship between the safety factor and the slope angle is clearly outlined.

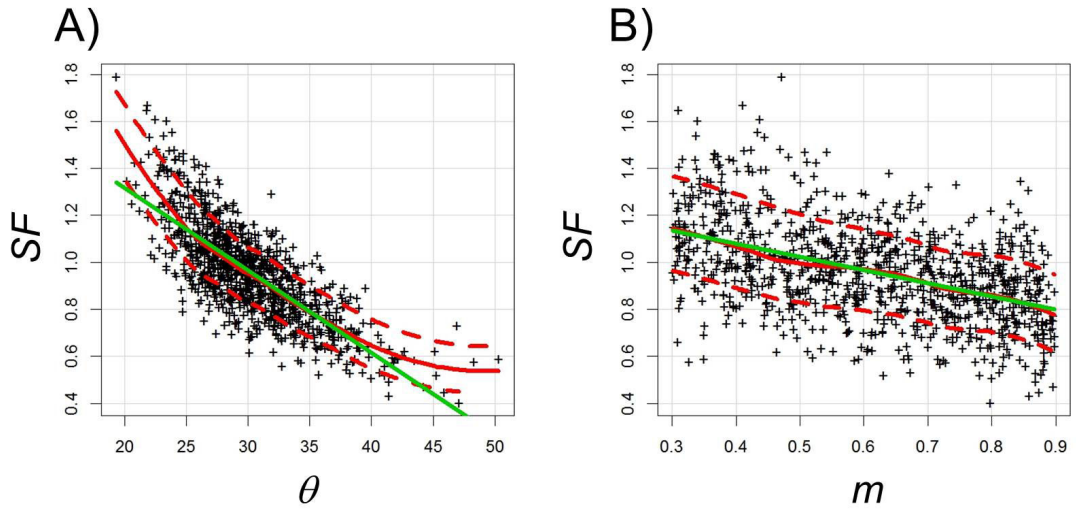

Figure 2.5: Relationship between the safety factor SF and: A) the slope angle θ ; B) the parameter m. The trend is either modelled using a linear model (green-coloured line) or a loess (local polynomial regression fitting) smoother with associated confidence intervals at 95 % (red-coloured lines).

2.4 Limitations and links with the subsequent chapters

The application of VBSA on the infinite slope analytical model outlines the richness of the information, which can be provided through the calculation of the Sobol' indices, whether for importance ranking and for deeper insight in the model behaviour. Yet, to the author's best knowledge, this kind of analysis has rarely been conducted in the field of geo-hazards, except for the study by [Hamm et al., 2006]. This can be explained by the specificities of the domain

of geo-hazard assessments, which impose considering several constraints.

- Despite the extensive research work on the optimization of the computation algorithms (e.g., [Saltelli et al., 2010] and references therein), VBSA remains computationally intensive, as it imposes to run a large number of simulations. In the example described in Sect. 2.3, the number of necessary model runs for the Sobol' indices to converge reaches 160,000. If a single model run had a low computation time (CPU time), say of 1 second, the application of VBSA would require about 44 hours (about 1.8 days) of calculation, which is achievable using a single computer unit (CPU). If the CPU time of a single model run was 1 minute, the application of VBSA would require more than 111 days of calculation, which is achievable using a computer cluster (e.g., [Boulahya et al., 2007]) of limited number of CPU (10 to 20). If the CPU time of a single model run was 1 hour, the application of VBSA would require more than 6,666 days (about 18 years) of calculation. To achieve VBSA with one week of calculation, the computer cluster should be composed of at least 1,000 CPU. Few research team or engineering companies can afford such large computer clusters. Nevertheless, most numerical models supporting geo-hazard assessments fall in the third category, either because they are large-scale or because the underlying processes are difficult to solve numerically. The application of slope stability analysis at the spatial scale of a valley, [Olivier et al., 2013, Baills et al., 2013], illustrates the first case, whereas the model of the La Frasse landslide illustrates the second case (it has a CPU time of about 4 days because it involves a complex elastoplastic model describing the complex behaviour of the slip surface, see further details in the next chapter). In those situations, the direct application of VBSA is obviously not achievable. Chapter 3 discusses this issue and proposes to rely on meta-modelling techniques, which basically consists in replacing the long-running simulator by a costless-to-evaluate mathematical approximation (see an overview by [Storlie et al., 2009]) to overcome such a difficulty;
- The second limitation is related to the nature of the parameters (input or output) that VBSA deals with: they are scalar. Yet, in the domain of geo-hazard, parameters are often functional, i.e. they are complex functions of time or space (or both). This means that parameters can be vectors with possible high dimension (typically 100 - 1,000). In the infinite slope case, the inputs parameters describing the soil properties can be spatially varying, for instance due to the presence of heterogeneities at the scale of the slope thickness (like clay patches embedded within some sandy soil formation). Besides, the water table can temporally vary, for instance due to time-varying rainfall infiltration. In the La Frasse case, the outputs are not scalar but temporal curves of the displacements (discretized in 300 steps) at any nodes of the mesh, i.e. the outputs are vectors of size 300 at any location. Another example is the spatial distribution of hydraulic conductivities

of a soil formation (see an example provided by [Tacher et al., 2005]). Chapter 4 further discusses this issue and describes a possible strategy to both overcome the computation burden and the high dimensionality of model outputs. The case of functional inputs is also addressed;

- Finally, a third limitation is related to the way uncertainty is mathematically represented. By construction, VBSA is based on the assumption that the variance can capture the main features of the uncertainty. This assumption has been shown not to be valid in cases of heavy tailed or multi-modal distributions [Auder and Iooss, 2008]. Besides, the emphasis is on a particular moment of the distribution, which may be too restrictive for efficient decision-making, because a decision-maker/analyst state of knowledge on a parameter or on a model output is represented by the entire uncertainty distribution [Borgonovo, 2007] or a given probability of exceedance like for the stability analysis [Morio, 2011]. Alternatives to VBSA have then been proposed in the statistical community, either based on the entire probability distribution [Borgonovo, 2007] or on the use of the statistical entropy [Auder and Iooss, 2008] in cases when the variables are deterministic but not known exactly. Yet, in the domain of geo-hazard assessments, data are often scarce, incomplete or imprecise, which add more difficulties. In the infinite slope case, data can be derived either from literature data or from laboratory tests conducted on soil / rock samples (but these are usually of small number). In particular, water table's height is related to water circulations, which are known to be complex at the scale of a slope and suffer from a lack of information and studies (e.g., [Winckel et al., 2004] for the French Basque coast). Systematically resorting to the probabilistic framework in such situations can be debatable. In Chapter 5, an in-depth discussion is provided. The applicability of an alternative tool (namely Fuzzy sets originally introduced by [Zadeh, 1965]) is then explored to mathematically represent epistemic uncertainty in a more flexible manner. Its integration in a sensitivity analysis is explored in Chapter 6.

3 Handling long-running simulators

As shown in Chapter 2, variance-based methods VBSA for Global Sensitivity Analysis GSA are powerful tools for importance ranking of parameter uncertainties. Yet, as discussed in Sect. 2.4, several difficulties for practical application to models supporting geo-hazard assessments exist. In this chapter, the first one is addressed, namely the implementation of VBSA when using complex large-scale computationally intensive numerical simulations. This chapter is primarily based on the work described in [Rohmer and Foerster, 2011]. In Sect. 3.1, a motivating example is used: the numerical simulations supporting hazard assessment related to the La Frasse landslide. An approach relying on meta-modelling is then introduced (Sect. 3.2), and in particular, a focus is given on kriging-type meta-models (Sect. 3.3). Since such techniques basically correspond to an approximation of the true simulator, this introduces an additional source of uncertainty (i.e. meta-model error, [Janon et al., 2014]): the impact of such uncertainty is discussed in Sect. 3.4. The meta-model-based strategy is then illustrated with application on the infinite slope model (Sect. 3.5) as described in Chapter 2 and finally the application is done on the real case of the La Frasse landslide (Sect. 3.6).

3.1 A motivating real case: the numerical model of the La Frasse landslide

3.1.1 Description of the model

The general setting and context of the La Frasse landslide are described in the introduction (Sect. 1.4, see also Fig. 1.2). The general behaviour of the landslide is strongly correlated to the properties of the slip surface. The complex behaviour of the slip surface material was modelled using the Hujeux elastoplastic multi-mechanism constitutive model [Aubry et al., 1982, Hujeux, 1985, Lopez-Caballero et al., 2007] and the Mohr-Coulomb crite-

rion was assumed for the other soil materials.

The Hujeux constitutive model permits coverage of a large range of deformation and takes into account: (1) the influence of confinement and stress path on the moduli; (2) the effects of over-consolidation; and (3) the influence of the void ratio. It can be used for granular as well as for clayey soil behaviours. It is based on a Coulomb type failure criterion and the critical state concept. The volumetric and deviatoric hardening regimes implemented in the Hujeux model lead to a dependence on the consolidation pressure as in the Cam-Clay family models, and to the evolution of the plastic yield surface with the deviatoric and volumetric plastic strains. Moreover, the model accounts for dilatancy/contractance of soils and non-associated flowing behaviour with evolution of the plastic strain rate through a Roscoe-type dilatancy rule. As outlined by [Laloui et al., 2004], the main parameters for the slip surface materials are: (1) the bulk (K) and shear (G) elastic moduli, which are assumed to depend on the mean effective stress through a power-type law of exponent n_e (referred to as the non-linear coefficient); (2) the critical state (linked with the initial critical pressure p_{co}) and plasticity parameters, essentially the friction angle ϕ at perfect plasticity, the plastic compressibility β (which determines the evolution of the yield surface depending on plastic strain by altering p_{co}); and (3) the dilatancy angle ψ, appearing in the flow rule and defining the limit between soil dilatancy and contractance.

Note that these parameters can be directly measured from either in situ or laboratory test results [Lopez-Caballero et al., 2007, Lopez-Caballero and Modaressi Farahmand-Razavi, 2008]. The other Hujeux law parameters, appearing in the flow rule, the hardening and the domains' thresholds are categorized as "not-directly measurable" and are estimated through numerical calibration techniques between the observed/experimental data and the simulated ones.

3.1.2 Objective of the sensitivity analysis

The sensitivity analysis was primarily focused on the seven properties of the slip surface, which primarily controls the hydro-mechanical behaviour of the landslide [Laloui et al., 2004]. The other parameters (categorized as "not-directly measurable" by [Lopez-Caballero et al., 2007]) were kept constant. The properties of the six other soil layers were assumed to be constant as well. An in-depth exploration of all these sources of uncertainty (a total number > 50) is not tackled here and can be addressed by using a screening method in a preliminary step (like the Morris method as proposed for instance by [Campolongo et al., 2011]).

Here, the objective was to understand the influence of the seven input parameters of the Hujeux constitutive model on the horizontal displacement calculated at two observation

points, namely in the upper (observation point 1, Fig. 1.2), and lower parts of the landslide (observation point 2, Fig. 1.2). The sensitivity analysis was carried out in a dynamic manner at each time step of the 300 days long crisis period (decomposed into 300 time steps). Put it in other words, the objective was to identify which properties drive the most the overall uncertainty in the temporal evolution of the surface displacements. In this view, the main effects (Sobol' indices of first order) were calculated within a factors' prioritization setting (as described by [Saltelli et al., 2008]). Assuming a situation where the same "level of uncertainty" is assigned to the parameters of the Hujeux model, a 25 % variation around the original values identified by Laloui and co-authors was affected to each of the seven input parameters (Table 3.1). A uniform probability distribution was assigned to each of these input parameters.

Table 3.1: Range of values for the slip surface properties of the La Frasse landslide (a variation in a range of 25 % around the original values given in [Laloui et al., 2004] is assumed).

Input factor	Symbol	Unit	Lower Value	Upper value
Bulk modulus	K	MPa	180	300
Shear modulus	G	MPa	83.25	138.75
Non-linearity coefficient	n_e	-	0.225	0.375
Internal friction angle	ϕ	°	19.125	31.875
Dilatancy angle	ψ	°	14.25	23.75
Plastic compressibility	β	-	20.625	34.375
Initial critical pressure	p_{co}	MPa	0.375	0.625

For a given inputs' configuration, a simulator run required nearly 96 hours (4 days) on a computer unit (CPU) with a 2.6 GHz dual core processor and 1 GB of RAM. The main and total effects were calculated using the sampling strategy of [Saltelli, 2002]. As an illustration, let us assume that the convergence of the Sobol' indices can be achieved with 1,000 random samples, the computation of main and total effects would require $N \times (n+2) = 1,000 \times (7+2) = 9,000$ model evaluations. The total computation time would reach $9,000 \times 96 = 864,000$ hours (36,000 days), which is obviously not achievable even when using a computer cluster with a larger number of CPU (see e.g., [Boulahya et al., 2007]).

3.2 A meta-model-based strategy

3.2.1 Principles

Let us consider f the complex large-scale computationally intensive numerical model. The meta-modelling technique consists in replacing f by a mathematical approximation referred to as "meta-model" (also named "response surface", or "surrogate model"). This corresponds to a function constructed using a few computer experiments (i.e. a limited number of time

Table 3.2: Main steps of meta-modelling strategy for global sensitivity analysis using computationally intensive numerical models

Step	Description
1	Generate n_0 different values for the input parameters $\boldsymbol{x}$ of the landslide model (for instance using a LHS technique combined with maximin criterion);
2	Evaluate the corresponding landslide model outputs y by running the computationally intensive model. The pairs $(x_i; y_i)$ with $i = 1, \ldots, n_0$ constitute the training data, based on which the meta-model can be constructed;
3	Assess the approximation and the predictive quality using cross-validation procedure;
4	Using the "costless-to-evaluate" meta-models, compute the Sobol' indices and analyse the importance of each of the input parameters.

consuming simulations), and aims at reproducing the behaviour of the "true" model in the domain of model input parameters (here the parameters of the Hujeux rheological law) and at predicting the model responses (here the horizontal displacements at two observation points) with a negligible computation time cost. An overview of the different meta-modelling techniques is provided by [Storlie et al., 2009]. In this manner, any approach relying on intensive multiple simulations, such as global sensitivity analysis, in-depth uncertainty analysis or parametric analysis involving multiple of scenarios, are made achievable at a "reasonable" computation time cost. The main steps of the methodology are summarized in Table 3.2.

3.2.2 Step 1: setting the training data

The approximation is constructed by running f given a limited number n_0 of different scenarios for the values of the model input parameters $\boldsymbol{x}$ (here the seven parameters of the Hujeux law), named training samples. The objective is to create a mapping (named training data) between $\boldsymbol{x}$ and the quantity of interest, namely the horizontal displacement y at the observation points. A trade-off should be found between maximizing the exploration of the inputs' domain of variation and minimizing the number of simulations, i.e. a trade-off between the accuracy of the approximation and the computation time cost. To fulfil such requirements, training samples can be randomly selected by means of the Latin Hypercube Sampling (LHS) method [McKay et al., 1979].

LHS is a stratified random sampling method (also named quasi-random), which can lead to convergence with smaller sample size than simple random samples. Let us take the example of two independent variables to which a uniform law is assigned. LHS proceeds as follows.

The range of each variable x_j (with j from 1 to n, the total number of random variables) is exhaustively divided into n_0 disjoint intervals of equal probability and one value is selected at random from each interval. The n_0 values thus obtained for x_1 are paired at random without replacement with the n_0 values obtained for x_2. These n_0 pairs are combined in a random manner without replacement with the n_0 values of x_3 to form n_0 triples. This process is continued until a set of n_0 n-tuples (n-dimensional vectors) is formed. This procedure can be further improved to maximise the domain exploration by using the "maxi-min" space filling design criterion of [Koehler and Owen, 1996], i.e. by maximizing the minimum distance between points. Figure 3.1A) provides the example of 15 points randomly generated with standard uniform law. We can notice that there are regions where some points are lacking (for instance over the domain $[0.4;0.7] \times [0.;0.5]$), whereas some points are very close (at $X_1 \approx 0.8$ and $X_2 \approx 0.3$): this outlines the lack of efficiency for the exploration of the domain. On the other hand, Fig. 3.1B) provides the example of 15 points randomly generated using LHS combined with maximin criterion: this shows that the repartition over space is more homogeneous with approximately equivalent distance from one point to another so that the exploration of the space tends to be well optimised.

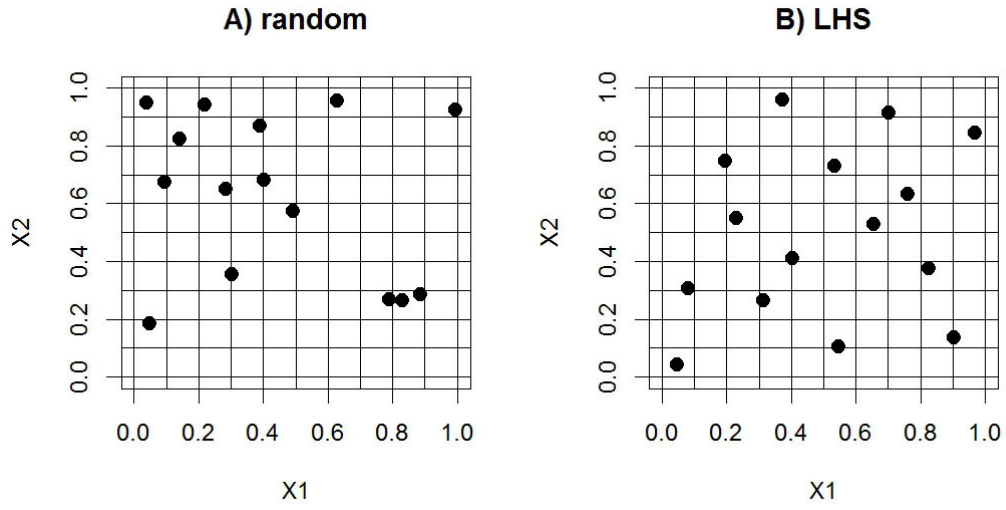

Figure 3.1: A) Random generation of two variables using standard techniques; B) Generation with Latin Hypercube Sampling combined with maximin criterion.

3.2.3 Step 2: construction of the meta-model

Using the training data, the approximation can be carried out relying on several types of meta-models, either using simple polynomial regression techniques (see e.g., [Abdallah, 2009]), non-

parametric regression techniques [Storlie et al., 2009], kriging modelling [Forrester et al., 2008], artificial neural networks [Papadrakakis and Lagaros, 2002], polynomial chaos expansions [Ghanem and Spanos, 1991], etc. The choice of the meta-model type is guided by the a priori non-linear functional form of the simulator, as well as the number of input parameters. Here, the focus is given on kriging meta-modelling. Full description is provided in Sect. 3.3. This non-parametric technique presents several attractive features.

- It is flexible to any kind of functional (mathematical) form of the simulator. In particular, it introduces less restrictive assumptions on the functional form of the simulator than a polynomial model would imply;
- It is an exact interpolator, which is an important feature when the simulator is deterministic;
- It provides a variance estimate of the prediction, the latter being very useful to guide the selection of future training samples according to the target of the optimization problem (see e.g., [Jones et al., 1998, Gramacy and Lee, 2009]).

Yet, it should be noted that these advantages come at the expense of a parameter learning stage (see e.g., [Langewisch and Apostolakis, 2010]), which may pose additional computational difficulties. An important issue to keep in mind is that these techniques introduce an additional source of uncertainty, namely the approximation uncertainty (aka meta-model error), which is further discussed in Sect. 3.4.

3.2.4 Step 3: validation of the meta-model

The third step aims at validating the meta-model quality. Two issues should be addressed: 1. the approximation quality, i.e., to what extent the meta-model manages to reproduce the simulated model outputs, and 2. the predictive quality, i.e., to what extent the meta-model manages to predict the model outputs for "yet-unseen" input parameter configurations. The approximation quality can be assessed using the differences between approximated and "true" model outputs (i.e., the residuals) and by computing the coefficient of determination, R^2. The latter can be done as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n_0} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n_0} (\overline{y} - y_i)^2} \tag{3.1}$$

where the y_i correspond to the observed model outputs (i.e., to the model outputs which were simulated using the long-running geomechanical simulator), $\overline{y}$ corresponds to the mean

of the observed model outputs, and $\hat{y}_i$ correspond to the approximated model outputs (i.e., the outputs which were estimated using the meta-model). A coefficient R^2 close to one indicates that the meta-model has been successful in matching the observations. Regarding the second quality issue, a first approach would consist of using a test sample of new data. Although the most efficient, this approach might often be impractical because additional numerical simulations are costly to perform. A possible option to overcome such a situation, coined by [Tukey, 1954] as "uncomfortable science", relies on q-fold cross-validation procedures (see, e.g., [Hastie et al., 2009]). This technique involves: 1. randomly splitting the initial training data into q equal subsets (q is typically between 5 and 10); 2. removing each of these subsets in turn from the initial set and fitting a new meta-model using the remaining q-1 sub-sets; 3. using the subset removed from the initial set as a validation set and estimating it using the new meta-model. Using the residuals computed at each iteration of this procedure, a coefficient of determination (denoted R^2_{CV}) can be computed using a formula similar to Eq. 3.1. For small training sets, the cross validation procedure with $q = 1$ is usually used corresponding to the so-called "leave-one-out" cross validation procedure. A typical threshold above 80 % is commonly used to qualify the predictive quality as "satisfactory" (e.g. [Marrel et al., 2008, Storlie et al., 2009]). Once validated, the meta-model can replace the long-running flow simulation to conduct VBSA (see Chapter 2).

3.3 A flexible meta-model: the kriging model

The basic concepts of kriging meta-modelling are introduced in this section. This type of meta-model can be viewed as an extension to computer experiments of the kriging method used for spatial data interpolation and originally developed by [Krige, 1951] for mining applications. For a more complete introduction to kriging meta-modelling and full derivation of equations, the interested reader can refer to [Sacks et al., 1989, Jones et al., 1998, Forrester et al., 2008].

The kriging model considers the deterministic (i.e. not random) response of the simulator $y = f(\boldsymbol{x})$ as a realization of a Gaussian stochastic process F so that $f(\boldsymbol{x}) = F(\boldsymbol{x}, \omega)$ where ω belongs to the underlying probability space Ω. In the following, we use the notation $F(\boldsymbol{x})$ for the process and $F(\boldsymbol{x}, \omega)$ for one realization. The process F results from the summation of two terms:

- $f_0(\boldsymbol{x})$, the deterministic mean function, which is usually modelled by a constant or a linear model and represents the trend of f;
- $Z(\boldsymbol{x})$, the Gaussian centred stationary stochastic process (with zero mean and covariance described below), which describes the deviation (i.e. departure) of the model from its underlying trend f_0.

The stochastic process Z is characterized by the covariance matrix $\boldsymbol{C}$, which depends on the variance σ_Z^2 and on the correlation function $\boldsymbol{R}$, which governs the degree of correlation through the use of the vector of length-scale parameters $\boldsymbol{\theta}$ between any input vectors. The covariance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is then expressed as $C(\boldsymbol{u},\boldsymbol{v}) = \sigma_Z^2 \cdot R(\boldsymbol{u},\boldsymbol{v})$, where $\boldsymbol{u} = (u_1; u_2; ...; u_n)$ and $\boldsymbol{v} = (v_1; v_2; ...; v_n)$ are two input vectors of dimension n.

A variety of correlation (and covariance) functions have been proposed in the literature (see e.g., [Stein, 1999]). The commonly used model is the Gaussian correlation function defined as follows:

$$R(\boldsymbol{u},\boldsymbol{v}) = \exp(-\sum_{i=1}^{n} \frac{\|u_i - v_i\|^2}{\theta_i}) \tag{3.2}$$

where the term θ_i determines the rate at which the correlation decreases as one moves in the i^{th} direction (with i from 1 to n). Intuitively, if $\boldsymbol{u} = \boldsymbol{v}$ then the correlation is 1, whereas if the distance between both vectors tends to infinity, then the correlation tends to 0.

Let us define $\boldsymbol{X}_D$ the design matrix composed of the input vectors $\boldsymbol{x}$ selected in step 1 of the methodology (i.e. the training samples) to be simulated so that $\boldsymbol{X}_D = (\boldsymbol{x}^{(1)}; \boldsymbol{x}^{(2)}; ...; \boldsymbol{x}^{(n_0)})$ and $\boldsymbol{y}_D$ the vector of simulated landslide displacements associated with each selected training samples so that $\boldsymbol{y}_D = (y^{(1)} = f(\boldsymbol{x}^{(1)}); y^{(2)} = f(\boldsymbol{x}^{(2)}); ...; y^{(n_0)} = f(\boldsymbol{x}^{(n_0)}))$.

Under the afore-described assumptions, the distribution of the horizontal displacement for a new input vector of the Hujeux law $\boldsymbol{x}^*$ follows a Gaussian distribution conditional on the design matrix $\boldsymbol{X}_D$ and of the corresponding simulated horizontal displacements $\boldsymbol{y}_D$ with mean value given by the kriging predictor $\hat{y}(\boldsymbol{x}^*)$ for the new configuration $\boldsymbol{x}^*$ and the variance by the kriging variance s^2 respectively defined by Eq. 3.3 and Eq. 3.4:

$$\hat{y}(\boldsymbol{x}^*) = \bar{y} + r(x^*) \cdot \boldsymbol{R}_D^{-1} \cdot (\boldsymbol{y}_D - \boldsymbol{I} \cdot \bar{y}) \tag{3.3}$$

with the constant: $\bar{y} = (\boldsymbol{I}^T \cdot \boldsymbol{R}_D^{-1} \cdot \boldsymbol{I})^{-1} \cdot (\boldsymbol{I}^T \cdot \boldsymbol{R}_D^{-1} \cdot \boldsymbol{y}_D)$

$$s(\boldsymbol{x}^*)^2 = \bar{\sigma}^2 \cdot ((1 - r(\boldsymbol{x}^*))^T \cdot \boldsymbol{R}_D^{-1} \cdot r(\boldsymbol{x}^*) + \frac{(1 - \boldsymbol{I}^T \cdot \boldsymbol{R}_D^{-1} \cdot r(\boldsymbol{x}^*))^2}{\boldsymbol{I}^T \cdot \boldsymbol{R}_D^{-1} \cdot \boldsymbol{I}}) \tag{3.4}$$

with the constant: $\bar{\sigma}^2 = \frac{(\boldsymbol{y}_D - \boldsymbol{I}\cdot\bar{y})^T \cdot \boldsymbol{R}_D^{-1} \cdot (\boldsymbol{y}_D - I\cdot\bar{y})}{n_0}$

where $r(\boldsymbol{x}^*)$ is the correlation vector between the test candidate $\boldsymbol{x}^*$ and the training samples; $\boldsymbol{R}_D$ is the correlation matrix of the training samples $\boldsymbol{X}_D$ and $\boldsymbol{I}$ is the unit matrix of size $n_0 \times n_0$

Equation 3.3 is used as a predictor, i.e. the "best estimate". The variance in Eq. 3.4 can be used to estimate the mean square error of the predictor and to deduce a confidence interval associated to the prediction. The regions of the input parameters' space where few data are available are underlined with higher variance, so that Eq. 3.4 also provides a local indicator of the prediction accuracy useful to guide sampling effort as extensively discussed by [Jones et al., 1998].

The above equations Eq. 3.3 and 3.4 are categorized as "ordinary kriging" and are the most common version of kriging used in engineering [Forrester et al., 2008]. A more general form of kriging equations exists, known as "universal kriging", and allows computing the deterministic mean function f_0 as a polynomial regression (generally of low-order) with unknown coefficients. For sake of clarity, these equations are presented in Appendix B. Note that in case of data measurements errors or non-deterministic computer code (for instance stochastic code in Fluid Mechanics), a constant regularization term referred to as "nugget effect" may be defined, hence introducing a white noise (e.g., [Gramacy and Lee, 2012]).

As afore-mentioned, the parameters of the kriging model corresponding to the constant (or the coefficients of regression) used to model the mean function f_0, the variance and the length scale parameters $\boldsymbol{\theta}$, are determined through a parameter learning stage. The most commonly-used method relies on a maximum likelihood estimation procedure (details can be found in Appendix B). For instance, this is implemented in the GEM-SA software [O'Hagan, 2006] and the MATLAB toolbox DACE [Lophaven et al., 2002]. However, the optimisation algorithms used for the parameters identification may show limitations related to numerical instabilities, multi-modality and dimensionality issues. Recently, a particular attention has been paid to this issue by developing efficient likelihood maximization algorithms like the one of [Park and Baek, 2001]: it is implemented in the package named "DiceKriging" developed by [Roustant et al., 2012] of the R software [R Core Team, 2014].

3.4 An additional source of uncertainty

The meta-model uses a limited number of simulator runs, i.e. input-output pairs (corresponding to the training data), to infer the values of the complex simulator output given a "yet-unseen" input configuration. Such an approximation introduces a new source of un-

certainty referred to as “code uncertainty” associated with the meta-model [O’Hagan, 2006] also referred to as “meta-model error” by [Janon et al., 2014], so that the sensitivity measures computed with the meta-model are “uncertain”.

One major advantage of kriging-like metamodels is the ability to provide an estimate of the uncertainty due to the use of a metamodel, hence of the metamodel error, through the use of the kriging variance defined by Eq. 3.4. Yet, the determination of the values of the covariance structure (correlation lengths, variance of the process, etc.) and of the regression coefficients in the trend of Eq. 3.4 is commonly performed using maximum likelihood maximisation techniques (see Appendix B): this might underestimate the “true kriging variance” [Helbert et al., 2009], because it neglects the whole set of possible values for these kriging parameters. By construction, they are chosen as the ones maximising the likelihood, but other values, though of smaller likelihood (i.e. they are said to be less likely), may also provide a good fit to the data (interpolation of the observations by the kriging meta-model). Besides, estimating covariance parameters by maximum-likelihood-based procedures with few data has shown to produce very dispersed results [Ginsbourger et al., 2009]. In some situations, the likelihood can be very flat, hence resulting in difficulties in choosing the "correct" kriging parameters, hence introducing large uncertainties [Helbert et al., 2009]. Several studies [Handcock and Stein, 1993, Zimmerman and Cressie, 1992] showed the interest of accounting for uncertainty in the estimates of the parameters of the covariance structure as well: this can be properly tackled under the Bayesian formalism.

Under this framework, the simulator is treated as an “unknown” function in the sense that the simulator output for any yet-unseen input configuration is unknown until the simulator is actually run for the considered configuration [Oakley and O’Hagan, 2004]. Metamodels are seen as emulators corresponding to statistical approximation so that a prior probability distribution is assigned to the simulator outputs (this corresponds to the background knowledge of the expert on the model’s behaviour, see Appendix D). A Gaussian Process (kriging metamodel) is chosen as the prior model for the simulator. Given some data (here the long-running simulations), the Bayes rule (as given by Eq. 3.6) can be applied to obtain a posterior distribution for these unknowns, which should be seen as an update version of the uncertainty representation of the model’s behaviour accounting for both the prior and the data (Appendix C). This framework can also incorporate uncertainty in the parameters of the kriging model by accounting for priors on them. Formally, the posterior probability density function of the variable of interest y given the observations (the scenarios which have already been simulated using the long running model) y_D, and the parameters of the kriging model

denoted $\psi = (\boldsymbol{\beta}, \sigma_Z, \boldsymbol{\theta})$ for the new configuration $\boldsymbol{x}^*$ can be expressed as:

$$p(y(\boldsymbol{x}^*)|\boldsymbol{y}_D) = \int p(y(\boldsymbol{x}^*, \boldsymbol{\psi})|\boldsymbol{y}_D) d\boldsymbol{\psi} = \int p(y(\boldsymbol{x}^*)|\boldsymbol{y}_D, \boldsymbol{\psi}) p(\boldsymbol{\psi}|\boldsymbol{y}_D) d\boldsymbol{\psi} \tag{3.5}$$

The first conditional probability in the right integral of Eq. 3.5 can be formulated using the expressions of the simple kriging given that: $(y(\boldsymbol{x}^*)|\boldsymbol{y}_D, \boldsymbol{\psi})$ follows a Gaussian random variable by construction (see Sect. 3.3). The second conditional probability is the posterior function of the parameters of the kriging model given the observations. It directly results from the Bayes rule as follows:

$$p(\boldsymbol{\psi}|\boldsymbol{y}_D) = \frac{p(\boldsymbol{y}_D|\boldsymbol{\psi}) p(\boldsymbol{\psi})}{p(\boldsymbol{y}_D)} \tag{3.6}$$

where $p(\boldsymbol{\psi})$ is the prior assigned to the kriging parameters, and $p(\boldsymbol{y}_D|\boldsymbol{\psi})$ is the likelihood. With little a priori knowledge, a flat prior (see e.g., [Gelman et al., 2013]) is usually placed on the linear regression coefficients in the trend part of the kriging model, including a conjugate inverse-gamma prior for the variance of the process. Gibbs samples can be obtained for these parameters. Priors also need to be placed on the parameters of the correlation structure. In the case of La Frasse, little is known in advance about the process, and non-informative Bayes priors are used [Berger et al., 2001]. They can be sampled using the Metropolis-Hastings algorithm. The likelihood expression derives from the assumption that y is normally distributed. Full derivation of the equations of the posterior distributions and details on the Monte-Carlo-based sampling scheme to solve the problem of approximation (and of inference) under the Bayesian formalism can be found in [Gramacy, 2005, Neal, 1997]. More details can also be found in the Appendix C. In practice, the tgp package [Gramacy, 2007, Gramacy and Taddy, 2012] of the R software [R Core Team, 2014] was used.

This approach returns the most likely value for the output given any input configuration (conditional mean of the kriging model of Eq. 3.3) as well as an entire probability distribution [O'Hagan, 2006]. Additional useful information for risk management purposes is the level of confidence (or accuracy) related to the sensitivity analysis based on the kriging model by defining a confidence interval with bounds corresponding to the 5% and to the 95% quantile of the full posterior distribution of the sensitivity measures. This confidence interval can be used to summarize the "code uncertainty" [O'Hagan, 2006] and is useful when the predictive quality of the meta-model is not high due to a small training data (see e.g., [Marrel et al., 2008,

Marrel et al., 2009]). In Sect. 3.6, it is shown through application on the La Frasse case how this additional information can be useful to interpret the results of the VBSA based on only a few simulations, i.e. to guide future characterization efforts even when the knowledge on the true simulator is limited.

3.5 Application to an analytical case

The application is first focused on the analytical (costless-to-evaluate) infinite slope model (as described in Chapter 2). Two input parameters were considered: the thickness z and the slope angle θ, namely a parameter with great influence (the latter one) and a second one with less importance (see Fig. 2.4). The other input parameters were assumed to be constant: C=10 kPa, ϕ =25°, $\gamma = 22kN.m^{-3}$, $\gamma_w = 9.81kN.m^{-3}$ and m=90 %.

The objective was to identify whether z or θ contributes the most to the SF variability within a "factors' prioritisation setting". It is assumed that very sparse data are available to characterize the uncertainty on these input factors so that z uniformly varies between 5 and 25 m and θ uniformly varies between 25° and 35° (step 1). Two different training set with respectively 6 and 20 samples of the form $\{z;\theta;SF\}$ were generated using the LHS-maximin approach (step 2). For each training sample, a kriging meta-model was then constructed with linear trend, exponential correlation (separable power), no nugget. An hierarchical scheme for Bayesian update (see details in Appendix C) with flat (non-informative) prior was assumed for the regression coefficient of the trend. The confidence intervals were computed using a MCMC with 4,000 rounds (iterations) of stochastic approximation (including 1,000 burn-in rounds).

Figure 3.2 (top) shows the comparison between the values of the safety factor SF obtained from direct simulations on a grid of 10×10 in the input factors' domain $[5\ ;\ 25]\times[25\ ;\ 35]$ (straight line) and from the prediction on the same grid using the kriging model (dashed line) for both training data of size 6 and 20. The coefficient of determination (Eq. 3.1) estimated using the whole grid respectively equals to 90.9 % for the first training sample and to 98.8 % for the second one. This shows a very good match for both meta-models. For illustration purposes, a "leave-one-out" cross validation procedure (step 3) was conducted and indicated a "high" predictive quality in both cases with coefficients of determination of 96.2 % for the first training set and 99.7 % for the second one. The estimated SF using both kriging meta-models (Fig. 3.2, middle) were compared to the "true" observed SF. The closer the dots to the straight black line, the better the approximation.

The results for the computation of the main and total effects required (step 4) $N \times (n+2) = 2,500 \times (2+2) = 10,000$ model evaluations using the sampling strategy of [Saltelli, 2002]. Recall

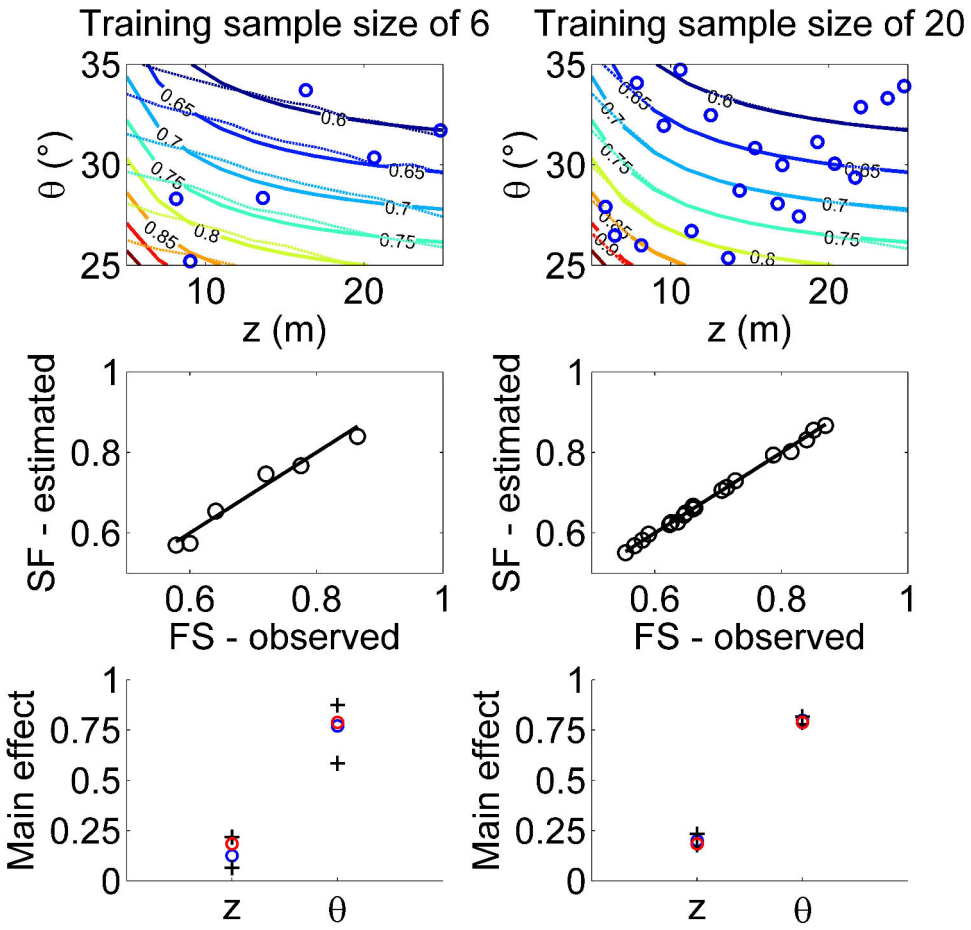

Figure 3.2: Top: comparison between the true values (continuous lines) and the estimates (dashed lines) of the factors of safety SF using the kriging model constructed with a training sample size of 6 (left) and of 20 (right). The training input configurations are represented by blue dots. Middle: comparison between the observed SF and the estimates within a "leave-one-out" cross validation procedure. Bottom: comparison between the true values (red dots) and the estimates of the main effects for the slope thickness z and of the slope angle θ (blue dots). The bounds of the confidence intervals associated with both kriging models are represented by black cross-type markers.

that the primary objective here was importance ranking so that we focused on the use of the main effects. The most likely values of the main effects calculated with both kriging meta-models (blue dots in Fig. 3.2, bottom) were compared to the main effects obtained from direct simulations (red dots on Fig. 3.2, bottom) by means of the R package "sensitivity" (using the same sampling strategy, same number of random of samples). These results are summarised in Table 3.3.

Table 3.3: Comparison between the "true" and the estimates of the main effects for the infinite slope analytical model. μ corresponds to the mean of the main effect computed with the kriging model. CI corresponds to the confidence interval defined by the 5 % and the 95 % quantile computed with the kriging model (within a Bayesian setting).

Input parameter	Thickness z (m)	Slope angle θ (°)
True Model	18.41 %	78.76 %
Metamodel constructed with 6 training samples	$\mu = 12.41\%$, $CI =$ [6.48;21.73] %	$\mu = 77.01\%$, $CI =$ [58.41;87.43] %
Metamodel constructed with 20 training samples	$\mu = 20.02\%$, $CI =$ [16.90;23.29] %	$\mu = 79.77\%$, $CI =$ [77.90;81.68] %

Differences are larger for the kriging model constructed with fewer training samples. However, the "true" values for the main effects still lie within the confidence interval bounded by the 5 % and the 95 % quantile (black cross-type marker in Fig. 3.2, bottom). Not surprisingly, increasing the number of training samples (i.e. our knowledge of the true function) decreases the range of "code uncertainty" as well as the differences between the true and estimated values (Table 3.3).

3.6 Application to the La Frasse case

A total number of 30 input parameters' configurations of the Hujeux parameters were generated. The resulting horizontal displacements computed over the crisis period are shown on Fig. 3.3 for the observation points 1 and 2. Note that the orders of magnitude of the displacements in the lower part (observation point 2) is larger than the ones in the upper part (observation point 1), because the pressure perturbation was primarily applied at the base of the landslide. The training data were generated using a grid computing architecture (computer cluster) composed of 30 CPU, so that all simulations were performed in parallel.

At each time step of the 1994 crisis period (300 time instants of 1 day), a kriging-type meta-model was constructed using the 30 training data to approximate the horizontal displacements at the observation points 1 and 2. The trend of the kriging meta-model is linear with exponential-type correlation function and no nugget. The model response were first pre-processed so that they had zero mean and standard deviation of 1.0. The hierarchical Bayesian update scheme (Appendix C) assumed a flat (non-informative) prior distribution for the linear regression coefficients of the trend. A "leave-one-out cross-validation" procedure was carried out for each step in order to assess the predictive quality of the meta-models. Fig. 3.4 depicts the temporal evolution of the coefficient of determination R^2 for the cross-validation procedure.

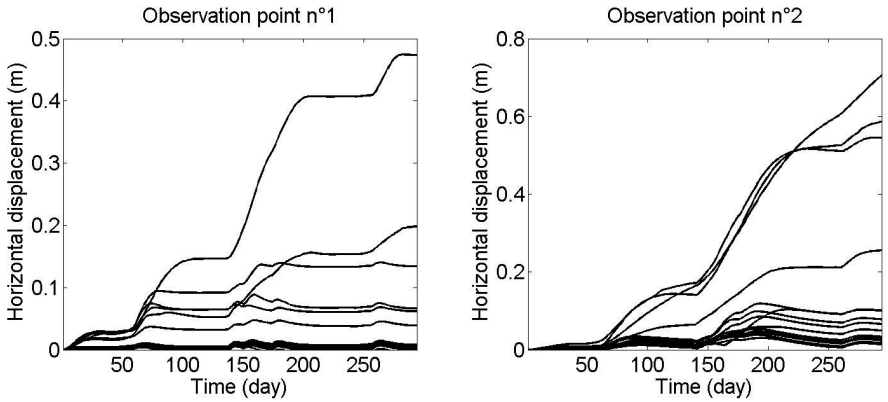

Figure 3.3: Temporal evolution of the training samples corresponding to the horizontal displacements (m) calculated for 30 different input configurations of the Hujeux law parameters (at the observation point 1 in the upper part of the landslide (left) and at the observation point 2 in the lower part of the landslide (right)).

During the first half of the crisis period (first 150 days), R^2 decreases over time for both observation points between 99.9 % and ≈95 %, hence indicating that the predictive quality is "high" over this period. During the second half of the crisis period, the quality is still satisfactory if we consider observation point 2 (R^2 varying between ≈ 80% and ≈ 95%), whereas it can be qualified as "low to moderate" for observation point 1 (R^2 steeply decreasing from $\approx 95\% to \approx 62\%$), hence indicating a possible moderate level of meta-model error.

The main and total effects were calculated using the sampling strategy of [Saltelli, 2002], hence requiring $N \times (n+2) = 1,000 \times (7+2) = 9,000$ meta-model evaluations. As the primary purpose of this study was the importance ranking of the parameters, the analysis was focused on the values of the main effects. Preliminary convergence tests were carried out for N=250, 500, 1,000 and 2,000: they showed that N=1,000 yields satisfactory convergence of the sensitivity measures to two decimal places (+/− 0.025). The confidence intervals were calculated using a MCMC inference procedure with 5,000 stochastic rounds (iterations) plus 1,000 burn-in rounds. Figure 3.5 provides the traceplots at 150 days for the linear coefficients and the range parameters (length-scales) for three input parameters: the elastic coefficient n_e, the plastic compressibility β, and initial critical pressure p_{co}. These demonstrate good "mixing" in the sense that they traverse its posterior space relatively rapidly, and they can jump from one remote region of the posterior to another in relatively few steps: these indicate that the MCMC chain has converged to its stationary distribution for the assumed number of rounds (other parameters behave similarly).

The total computation time of the kriging-based sensitivity analysis reached a total of 108 hours (4.5 days), including the generation of the training data (≈4 days), the construction of a

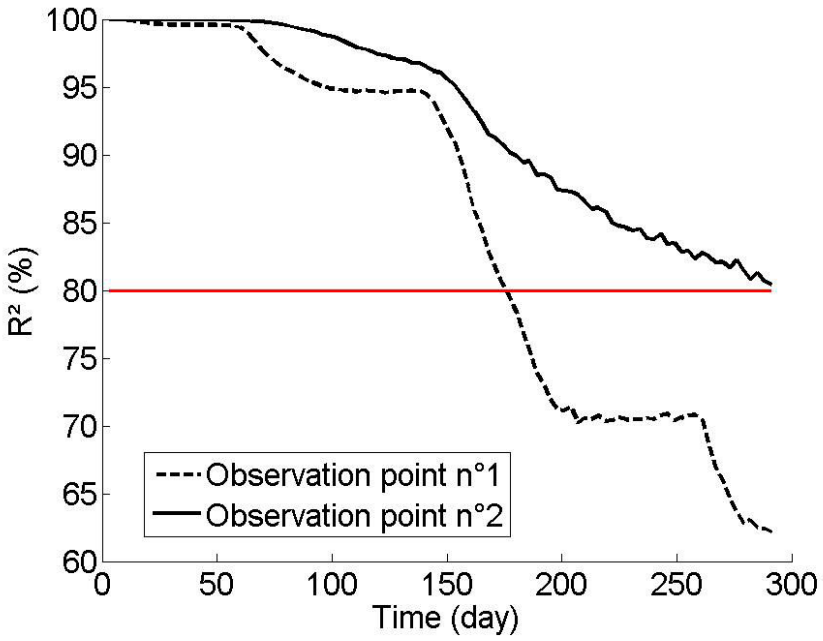

Figure 3.4: Temporal evolution of the coefficient of determination R^2 for the "leave-one-out" cross validation procedure of the kriging models constructed at each time instant of the crisis period at the observation point 1 in the upper part of the landslide (black dashed line) and at the observation point 2 in the lower part of the landslide (black straight line). The threshold of 80 % indicates a "satisfactory" predictive quality and is outlined by a horizontal red straight line.

meta-model at each step of the crisis period (≈3 hours) and the cross-validation procedure (≈3 hours). If the same analysis had been undertaken by direct simulations, the total computation time would have reached $9,000 \times 30 \times 96 = 28,800$ hours (1,200 days) using the same 30 CPU cluster. To achieve a computation time of 108 hours, a computer cluster composed of 8,000 CPU would have been required.

Figure 3.6 (top) depicts the temporal evolution of the first most important input parameter (straight green line) at the observation point 1, in the upper part of the landslide (Fig. 3.6, left), and at the observation point 2, in the lower part of the landslide (Fig. 3.6, right). Similarly, Fig. 3.6 (bottom) provides the temporal evolution of the second most important input parameter. The input parameters (Table 3.1) were ranked in terms of importance based on the mean of the main effect (blue straight line, Fig. 3.6) computed with the kriging models constructed at each instant of the crisis period.

This preliminary ranking of the input parameters was re-assessed in a second step taking into account the range of uncertainty associated to the sensitivity measures i.e. using the 5% and the 95 % quantile of the posterior probability distribution associated to the main effects (black

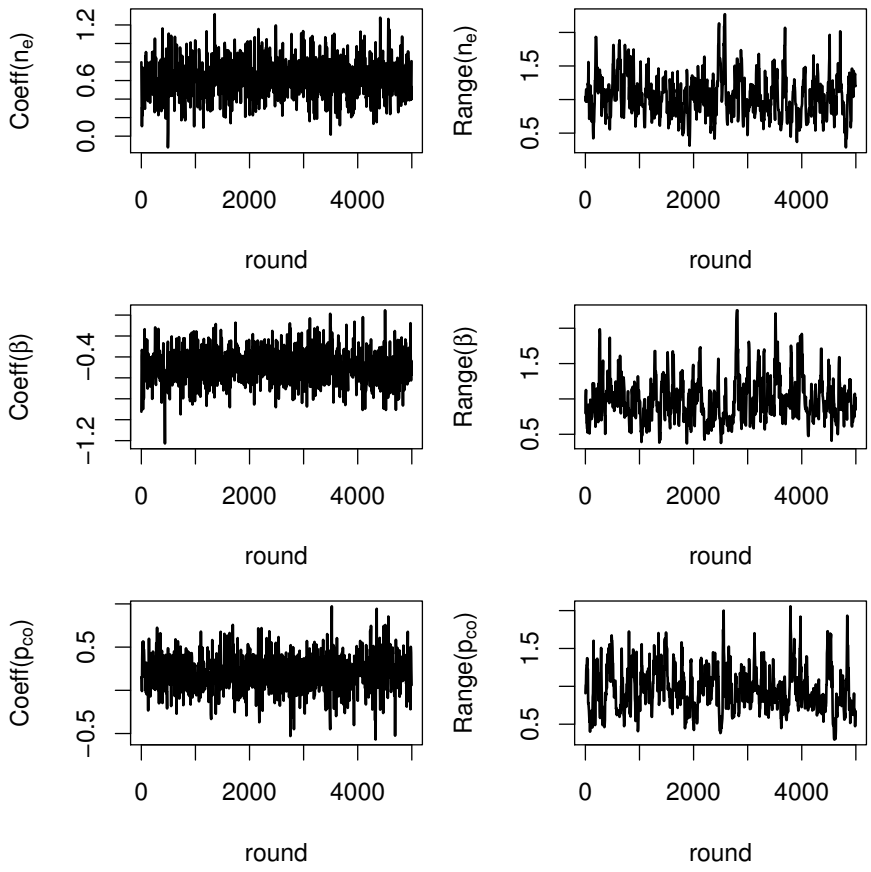

Figure 3.5: Traceplots for the linear coefficients and the range parameters (length-scales) for the elastic non-linear coefficient n_e (top), the plastic compressibility β (middle), and the initial critical pressure p_{co} (bottom).

dashed line, Fig. 3.6). In a second step, we qualified the kriging model as "unsure" with respect to the sensitivity measures in regions where the confidence intervals of the first and second most important input factors intersect (i.e. where the confidence intervals overlap).

Considering observation point 1, Fig. 3.6 (left) shows that for the first 150 days, coefficient n_e can be identified as the first most important input factor with a mean of the main effect constant at $\approx$ 20%, whereas the dilatancy angle ψ can be identified as the second most important parameter with a mean of the main effect constant at almost 10 %. For the second crisis period, the confidence intervals intersect and the ranking can be qualified as "unsure": the meta-model error prevents use from a confident ranking.

Figure 3.7 (left) gives the mean of the main effects and the associated confidence intervals at three different steps of the crisis period, namely 30 days (Fig. 3.7, top), 150 days (Fig. 3.7, middle) and 210 days (Fig. 3.7, bottom). At 30 days, n_e can clearly be identified as the first

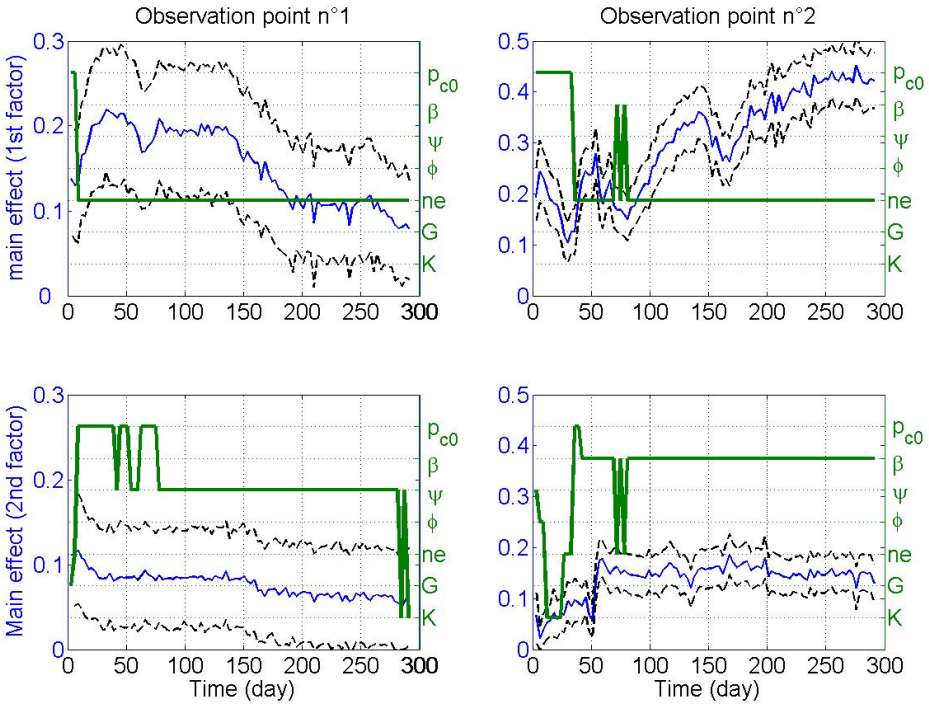

Figure 3.6: Temporal evolution during the crisis period of the mean of the main effects (blue straight line) at the observation point 1 in the upper part of the landslide (left) and at the observation point 2 in the lower part of the landslide (right) for the first (green straight line, top) and the second (green straight line, bottom) “most important” input factor. The black dashed lines represent the 5% and the 95% quantile.

most important input factor, but the ranking of the others is hardly feasible considering the intersecting confidence intervals. Over time (at 150 and 210 days), the confidence intervals for all input factors intersect so that the ranking is “unsure”. This result is in agreement with the low coefficient of determination of the cross-validation procedure over the second half of the crisis period (Fig. 3.4, black dashed line). The sum of main effects (i.e. a measure of the interaction in the model, see Sect. 2.2) decreases from around 50% down to 25%, hence indicating that the "complexity" of the relationship between the horizontal displacements and the input factors increases over the crisis period. This means that the approximation of the complex relationship becomes harder over time due to the increase of the contribution of the interactions among the input parameters. Physically, this is related to the location of the observation point 1, which is far from the pore pressure perturbation located in the lower part of the landslide so that the correlation of the mechanical response with the hydraulic response is expected to increase over time with increasing pore pressure. As a conclusion, the knowledge on the “true” simulator should be increased (i.e. additional long-running simulations) for the second crisis time period in order to increase the predictive quality of the

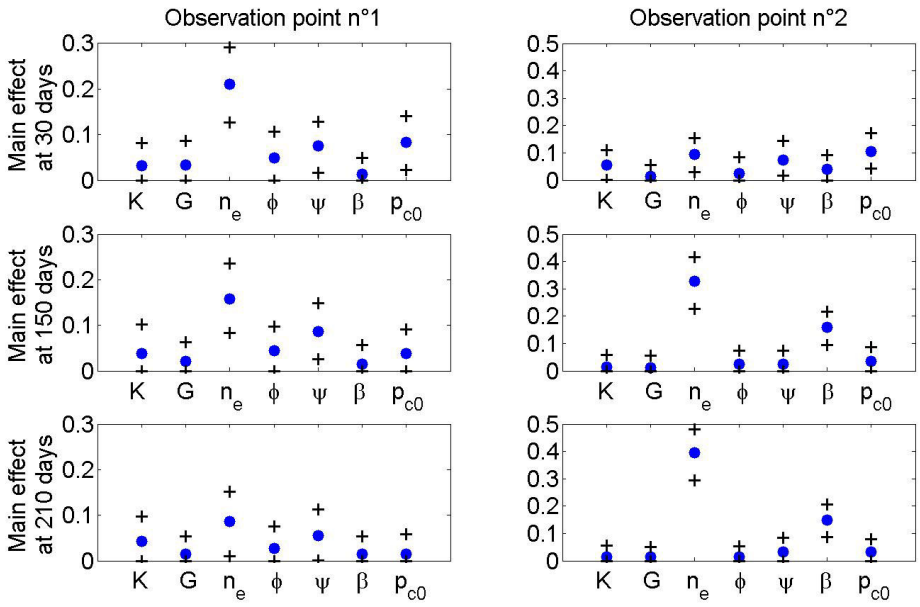

Figure 3.7: Mean of the main effect (blue dots) for each input factor of the slip surface constitutive law at different instants of the crisis period (30 days (top), 150 days (middle) and 210 days (bottom)) at the observation point 1 in the upper part of the landslide (left) and at the observation point 2 in the lower part of the landslide (right). The bounds of the confidence intervals (5% and 95% quantile) are outlined by black cross-type markers.

kriging meta-model, hence to narrow the width of the confidence interval in the upper part of the landslide.

Considering the observation point 2, the coefficient n_e can also be identified as the first most important parameter, over the time period after 50 days, with a mean of the main effect increasing from $\approx 20\% to \approx 45\%$, whereas the plastic compressibility β can be identified as the "second most important" input factor with a mean of the main effect approximately constant and equal to 15 % (Fig. 3.6). Before the time instant of 50 days, the ranking is made more difficult since the confidence intervals appear to intersect. This is clearer in Fig. 3.7 (right), which gives the mean of the main effects and the associated confidence intervals for steps 30 days (Fig. 3.7, top), 150 days (Fig. 3.7, middle) and 210 days (Fig. 3.7, bottom). It shows that over time, n_e and β can be identified "with certainty", but the ranking of the others remains hardly feasible. The analysis of the sum of the main effects reveals that the interactions among the input parameters becomes weaker over time by increasing from nearly 41% to 65%. Yet, the interaction term can here still be considered large.

Despite the limited number of simulator runs of 30 i.e. the limited knowledge on the "true" simulator, several conclusions can still be drawn to guide future investigations. The sensitivity

analysis based on the kriging meta-model emphasizes coefficient n_e as the most important source of parameter uncertainty i.e. as the input factor requiring in priority further investigations over the crisis period, whatever the part of the landslide (upper or lower). This result can have strong implications in practice, because the estimation of this parameter can be tedious, since it is strongly dependent on the availability of lab tests at small strains, where the behaviour is truly elastic (e.g. strains lower than 10^{-4}): this condition is known to be hardly realised for classical triaxial tests where the accuracy is not better than 10^{-3} (e.g. [Biarez et al., 1994]). This parameter is then usually deduced using standard values estimated for analogous types of soil. Nevertheless, such an analogy-based approach can be expected to be hard to achieve in the La Frasse landslide case since the considered soil material, being on the slip surface, is inherently heterogeneous.

On the other hand, the sensitivity analysis also outlines the plastic compressibility β as important for further investigations in the lower part of the landslide i.e. where the evolution of pore pressures was the most important. This indicates that the yield surface in the lower part evolves (this parameter affects the pre- consolidation stress, which determines the major axis of the elliptic yield envelope) during the pore pressure perturbation. In practice, this parameter can be obtained from oedometer tests. No further conclusions can be drawn without increasing the knowledge on the "true" simulator, for the third (or lower) most important parameter due to the uncertainty on the meta-model together with the Monte-Carlo sampling error.

These conclusions are valid for the considered illustrative case especially regarding the assumptions on the range of uncertainty assigned to all input factors (variation in a range of 25 % around the original values). It should be noted that more sophisticated situations for uncertainty representation can be considered. For instance, in a calibration setting, the uncertainty on each slip surface property should be adequately represented making use of any kind of information related to the measurement procedure (number of samples, measurement error, and possibility to construct empirical probability distribution, etc.), but also to the model inadequacy to fit the observations (i.e. differences between observed and simulated curves). Accounting for these multiple sources of information in a calibration procedure when using long running simulations can be tackled within a Bayesian framework, relying on the combination of basis set expansion (described in Chapter 4) and of kriging meta-models [Higdon et al., 2008].

3.7 Concluding remarks of Chapter 3

Motivated by the real-case of the La Frasse Swiss landslide, VBSA was made possible via the combination with kriging meta-models despite the high computation time cost of the numerical simulator (one single simulation took 4 days of computation). To the author's best knowledge, the application of such kinds of technique is original in the domain of landslide risk assessment. The computation burden was largely alleviated: the first order sensitivity indices were derived using only 30 different simulations so that the analysis was carried out within a week time period. Yet, applying meta-modelling techniques should not give "false impression of confidence": the approximation and predictive quality should be carefully studied, in particular by performing cross-validation analysis. Besides, the impact of meta-model error (i.e. the additional uncertainty introduced because the true simulator was replaced by an approximation) should be accounted for in the presentation of the final results. Here, we have treated the problem under the Bayesian formalism: this allowed assigning confidence intervals to the derived sensitivity indices and the importance ranking could be done accounting for the limited knowledge on the "true" simulator. In this manner, the zones where the meta-model was "unsure" could be highlighted and the confidence in the results of sensitivity analysis could be improved.

Yet, it should be underlined that the Bayesian treatment of the metamodel learning phase, though having been improved recently [Le Gratiet et al., 2014], may show limitations in practice, since it can lead to an increase in the computation time cost (compared to a learning phase relying on maximisation of the likelihood function). Further on the practical level, such an approach can be very "technical" especially regarding the convergence of the MCMC chain, which can be tedious and can take some time. Besides, it should be underlined that the results can be very dependent on the choice of the type of priors assigned to the kriging parameters [Helbert et al., 2009]. Regarding the first issue, other methods have been proposed in the literature: i. the bootstrap-based approach by [Storlie et al., 2009] and ii. the derivation of the lower and upper bounds on the meta-model errors in the framework of reproducing kernel Hilbert space interpolation (equivalent to kriging) as proposed by [Janon et al., 2014]. Regarding the issue of prior dependency, it is worth noticing the study by [Helbert et al., 2009], who showed the great interest in using informative priors (when it is appropriate to do so) to both improve the predictor accuracy and the kriging variance. Deriving an informative prior can be based on proxy simulations of the landslide (simulations with degraded physics but with lower CPU time): an example can be the use of simple reservoir models for flow behaviour assessment, see e.g., [Matti, 2008].

4 Handling functional variables

This chapter addresses the issue of dealing with functional variables in sensitivity analysis. The term "functional" is used to refer to variables, which are not scalar (i.e. they do not take a single value), but they are complex functions of time or space (or both). These variables can either correspond to the outputs (e.g. spatial distribution of surface deformation induced by mining activities) or to the inputs of the geo-hazard assessments (e.g., spatial distribution of the hydraulic properties of a soil formation, see an example in [Tacher et al., 2005]). In the following, the problem of handling functional outputs are first analysed (Sect. 4.1). Then a method relying on dimension reduction techniques is proposed (Sect. 4.2) and is applied to the La Frasse landslide case (Sect. 4.3). Sections 4.1-4.3 are based on the work described by [Rohmer, 2013]. Finally, an additional study investigating the applicability of the proposed strategy to the case of functional inputs is described (Sect. 4.4).

4.1 Problem definition for functional outputs

This section describes the difficulties related to the fact that the outputs of the geo-hazard assessments are generally functional. Let us consider again the landslide model outputs of La Frasse (described in the Chapter 3): these correspond to time series of landslide displacements at given spatial locations.

As described in Chapter 3, the objective of GSA applied to this case was to understand the influence of the seven input parameters of the Hujeux constitutive model on the temporal evolution of the horizontal surface displacements calculated in the upper part and in the lower part of the landslide, i.e. to identify which properties drive the most the overall uncertainty in the temporal evolution of the surface displacements. Note that the analysis is restricted to time-dependent landslide model outputs by using the maximum displacements in the lower

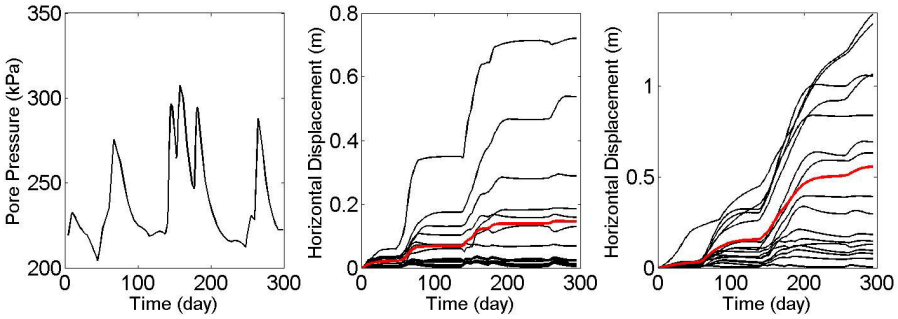

Figure 4.1: A) Example of pore pressure changes imposed at the base of the slide; here shown at node 292; Maximum time-dependent horizontal displacements computed for 30 different values of the slip surface properties: B) in the upper part and C) in the lower part; The red-coloured curve corresponds to the mean temporal function.

and upper part of the landslide. Sensitivity of spatio-temporal outputs constitutes a direction for further works using for instance the approach proposed by [Antoniadis et al., 2012].

We used the 30 input parameters' configurations of the seven slip surface properties (see description in Sect. 3.1), which were randomly generated using the procedure described in Sect 3.2. A set of 30 time-dependent displacements were then computed for the upper and lower parts of the landslide (see Fig. 4.1B and C). Each landslide model output was discretised in ≈ 300 time steps, each of them representing a time interval of one day. The total time duration corresponds to the one of the changes in the pore pressure (see an example at the bottom of the landslide, at node 292 of the model in Fig. 4.1A). Further details on the assumptions underlying the Landslide model can be found in [Laloui et al., 2004].

Formally, these model outputs can be represented by T–dimensional vectors (i.e. curves) where T corresponds to the number of calculation time steps (here T=300). To investigate the "dynamic" sensitivity analysis of the the time-dependent model output, a possible approach (named time-varying GSA) could consist in calculating the Sobol' indices separately at each time step as proposed in Chapter 3. Yet, this approach presents several disadvantages (see also discussions in [Campbell et al., 2006, Auder et al., 2012]), because it may become intractable for long time series (T typically ranging from 100 to 1,000), and it introduces a high level of redundancy, because of the strong relationship between outputs from successive time steps. Put in other words, information on the correlation between successive time steps may be lost in this approach.

An alternative is to compute the generalized sensitivity index proposed by [Lamboni et al., 2011], which allows explaining the influence of each input parameter on the overall functional output variability. Though this approach proves to be very efficient, it may miss important dynamic

features of the output, i.e. the general structure (or form) of the temporal evolution as outlined by [Campbell et al., 2006]. In other words, the questions of primary interest for dynamic sensitivity analysis are: What shifts the temporal evolution up or down? What makes a possible peak wider or narrower? What reverses the temporal evolution? What accelerates the behaviour? etc., i.e. the influence of the input parameters on the dominant modes of temporal evolution of the model output. In this view, the strategy proposed by [Campbell et al., 2006] and recently applied in the field of system biology modelling by [Sumner et al., 2012] should be tested in the La Frasse case: this consists in the reduction of the dimensionality by expanding the functional model output in an appropriate functional coordinate system (with lower dimension than T) followed by VBSA of the coefficients of the expansion. To alleviate the computation time cost associated with the calculation of the Sobol' indices, meta-modelling techniques can then be applied to the coefficients of the expansion [Auder et al., 2012]. Each ingredient of this strategy (summarised in Table 4.1) is detailed in the following.

4.2 Reducing the dimension

In this section, the objective is to reduce the dimension of the functional (time-dependent) output of the landslide model. In the present study, the most common case in landslide modelling is considered, namely functional data discretised on a regular grid of time points, i.e. vectors of large but finite dimension.

4.2.1 Principles

Formally, consider a set of n_0 functional model outputs, $\boldsymbol{y}_i(t)$ (with $i = 1, ..., n_0$), and discretised in T-vectors, i.e. the time step t takes finite values in the set (1,2,..., T). Define $\boldsymbol{Y}$ the $n_0 \times T$ matrix so that each row is composed of a model output.

In the La Frasse case, the set of functional model outputs corresponds to n_0=30 vectors of horizontal displacements of dimension T=300 (number of time steps). The objective of the basis set expansion is then to reduce the set of temporal curves to scalar values of finite number $d << 300$ so that they describe the key features of the temporal evolution of the calculated displacements, i.e. their dominant modes of time variations. This can be achieved by expanding the functional model output in an appropriate functional coordinate system, i.e. in terms of some basis functions of time $\phi_k(t)$ (with $k = 1, ..., d$). The basis set expansion of the

set of centred temporal curves $\boldsymbol{y}_i^C(t)$ reads as follows:

$$\boldsymbol{y}_i^C(t) = \boldsymbol{y}_i(t) - \bar{\boldsymbol{y}}(t) \approx \sum_{k=1}^{d} h_{ik} \cdot \phi_k(t) \tag{4.1}$$

where $\bar{\boldsymbol{y}}$ the mean temporal function is computed as the mean of the $\boldsymbol{y}(t)$ at each time step t. The scalar expansion coefficients h_{ik} indicate the "weight" (contribution) of each of the d basis function in each of the n_0 temporal curves. Usually, the dimension d is chosen so that most information is concentrated in the d first basis functions. See below for an example using Principal Component Analysis.

The basis functions can be of various forms, such as pre-defined Legendre polynomials, trigonometric functions, Haar functions, or wavelet bases [Ramsay and Silverman, 2005]. The disadvantage is to give beforehand an idea of the modes of variations. Alternatives are adaptive basis functions, which determine the basis functions from the data. The classical data-driven method is the multivariate Principal Component Analysis, denoted PCA [Jolliffe, 2005], which can be applied to the time-dependent model outputs viewed as vectors of finite dimension. A continuous form of the method, the functional principal component analysis exists as well. Note that more advanced methods of basis set expansion combined with clustering techniques may be used (see e.g. [Auder et al., 2012]), when the structure of the model outputs are very complex (e.g., highly non-linear).

4.2.2 Principal Component Analysis

The PCA decomposition is a multivariate statistical procedure aiming at reducing the dimensionality of a data set while minimizing the loss of information. The basic concepts are here introduced. For a more complete introduction and full derivation of equations, the interested reader can refer to [Jolliffe, 2005, Ramsay and Silverman, 2005].

Let us denote the variance-covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{Y}_C^T \cdot \boldsymbol{Y}_C / n_0$ of the columns of $\boldsymbol{Y}_C$, i.e. the matrix composed of the set of discretised centred temporal curves $\boldsymbol{y}_i^C(t)$. The PCA decomposition truncated at level d is based on the expansion of $\boldsymbol{\Sigma}$ as follows:

$$\boldsymbol{\Sigma} \approx \sum_{k=1}^{d} \lambda_k \cdot \boldsymbol{v}_k \cdot \boldsymbol{v}_k^T \tag{4.2}$$

where $\lambda_1 > \lambda_2 > ... > \lambda_d$ are the eigenvalues of $\boldsymbol{\Sigma}$ and $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_d$ are normalised and mutually orthogonal eigenvectors associated to these eigenvalues. Then, the principal components PCs $\boldsymbol{h}_k$ (with $k = 1, 2, ..., d$) are the mutually orthogonal linear combinations of the columns of $\boldsymbol{Y}_C$ and of the eigenvectors $\boldsymbol{v}_k$ so that $\boldsymbol{h}_k = \boldsymbol{y}^C \cdot \boldsymbol{v}_k$ and $\|\boldsymbol{h}_k\|^2 = \lambda_k$. The inertia defined as the trace of $\boldsymbol{\Sigma}$ measures the total dispersion (variability) among the rows of $\boldsymbol{Y}$. By construction, it is the same as the trace of the matrix composed of the PCs. This means that the amount of variation described by the PCs declines as k increases so that $d << T$ can be chosen considering a given "level of explained variance", let say, 99.0%.

4.2.3 Interpreting the basis set expansion

The application of PCA to the set of time-dependent displacements in the upper part of the La Frasse landslide (Fig. 4.2B) shows that the two first principal components PC1,2 respectively account for 99.0% and ≈ 1% of the variation in the set. Figures 4.2A and 4.2C show the temporal evolution of both PCs during pore pressure evolution at the base of the landslide. The first principal component PC1 is negative throughout the whole time duration of the crisis and alternately evolves between phases of steep decreases (approximately corresponding to the time of peaks of pore pressure) and phases of constant evolution (approximately corresponding to the time-interval between pore pressure peaks). The second PC (denoted PC2) decreases from 0 to a negative constant and then steeply increases above zero after the second major period of pore pressure peaks (at the time instant of ≈ 175 days).

Figure 4.3A shows two examples of time series (for inputs' configuration N°27 and N°28), here referred to as the observations, computed using the long-running simulator. These are compared to the time series re-constructed (approximated) via the PCA procedure, i.e. using the afore-described PCs (here referred to as approximations). This shows that the "projection" leads to minor error and the time series can reasonably be summarised using only the two afore-described PCs.

To get a better "physical" picture, [Campbell et al., 2006] advocate plotting the mean temporal function plus and minus some multiple of the PC (this multiplicative constant is chosen as 0.5 in Fig. 4.2B and 4.2D). This allows interpreting the PCs as perturbations from the mean temporal function, i.e. deviations from the "average" temporal behaviour of the landslide. In the upper part (see red-coloured curve in Fig. 4.1B), this average behaviour corresponds to successive phases of sharp increases in the horizontal displacements ("destabilized" phase) and of quasi- horizontal evolution ("stabilized" phase). Figure 4.2 shows that the first PC corresponds to a vertical up-down shift relatively to the mean function over the whole time duration, but with a magnitude of shift increasing with time. In other words, model runs which

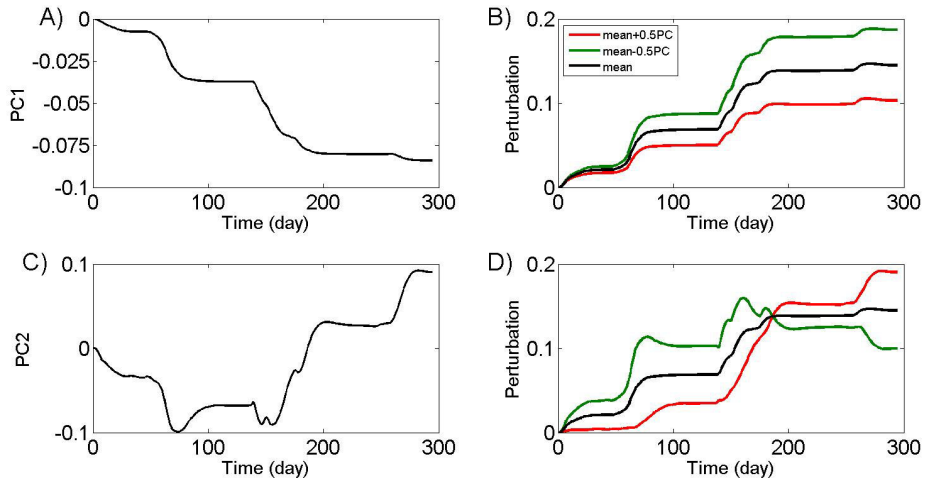

Figure 4.2: A) First principal component PC1 for the set of time-dependent horizontal displacements computed in the upper part of the landslide. B) Interpretation of PC1 as a perturbation of the mean temporal function (black curve) plus (red curve) and minus (green curve) a multiple of PC1 (here set at 0.5); C) Second principal component PC2; D) Interpretation of PC2 as a perturbation of the mean temporal function.

have negative scores for PC1 will have higher than average displacement values across the whole time duration. From a risk assessment perspective, those model simulations might lead to an increase over time of the horizontal displacements, i.e. this mode of temporal behaviour can be viewed as the overall most unstable ones. The second PC accounts for the same behaviour as PC1 before the time instant of 175 days. After this date, the behaviour is reversed, i.e. a model run with negative scores for PC2 will have lower than average displacement, i.e. this can be viewed as a stabilization mode.

The application of PCA to the set of time-dependent displacements in the lower part of the La Frasse landslide shows that 99.9% of the variation can be explained by three PCs (with contribution of respectively 98.3%, 1.4% and 0.3%). Figure 4.3B shows the comparison between observed time series and the PCA- re-constructed time series considering two simulation scenarios. This shows that the "projection" can lead to minor-to-moderate error even by using three PCs (in particular for case N°27). The impact of this source of error (uncertainty) is further discussed in Sect. 4.3.4.

The “average” behaviour captured by the mean temporal function (see red-coloured curve in Fig. 4.1C) is different to the one in the upper part and corresponds to a monotonically increasing function, hence showing that the average behaviour in the lower part of the landslide is “destabilized” over the whole time period (contrary to the upper part). The interpretation of the two first PCs is similar to the ones in the upper part (but here, relatively to the average

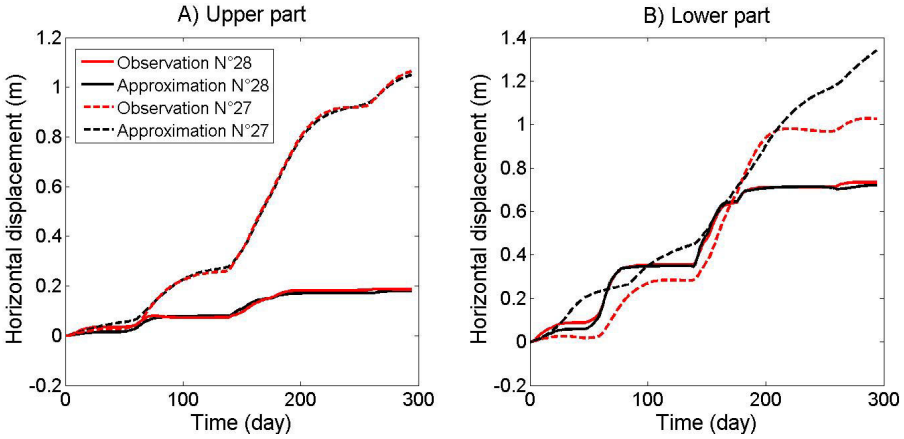

Figure 4.3: Comparison between observed (black) and PCA-reconstructed (red) time series, i.e. between the ones simulated using the long-running simulator and the ones approximated using PCA considering two simulation scenarios: A) in the upper part; B) in the lower part.

destabilized behaviour). The interpretation of the third PC (Fig. 4.4) is more complex and can be understood as a phase of acceleration, followed by a phase of deceleration and finally by a new phase of acceleration (considering model simulations with negative scores on PC3).

The basic idea of the dynamic sensitivity analysis through PCA analysis is then to assess the sensitivity of the scores of each PC to the input parameters. For instance, if the scores of PC1 are sensitive to a particular input parameter, this means that this parameter is important in producing the type of behaviour in the model output as afore-described. Sensitivity can be assessed using Sobol' indices, as proposed for instance by [Sumner et al., 2012]. Yet, as underlined in the introduction, the algorithms to compute such sensitivity indices require a large number of simulations, which may be impractical when using landslide model with CPU time of several hours (in the La Frasse case, the CPU time exceeds 4 days). This problem is further tackled in the following.

4.3 Strategy description

To overcome the computation challenge related to the estimation of the Sobol' indices using a long-running landslide model, this section describes the strategy relying on the meta-modelling technique for dynamic sensitivity analysis using PCA analysis. The main steps of the methodology are summarized in Table 4.1.

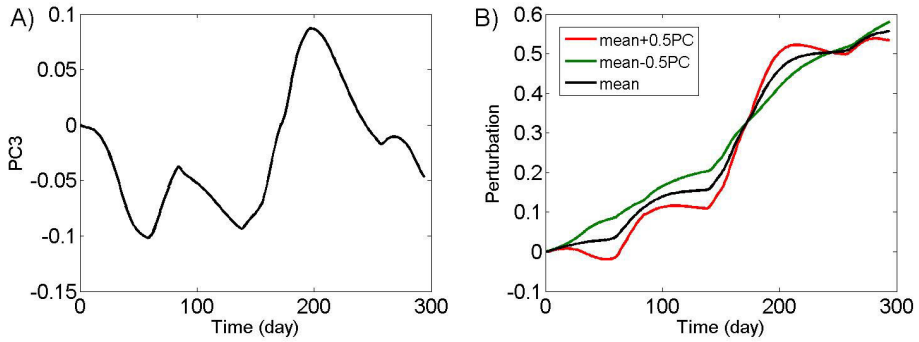

Figure 4.4: A) Third principal component PC3 for the set of temporal curves computed in the lower part of the landslide. B) Interpretation of PC3 as a perturbation of mean temporal function (black curve) plus (red curve) and minus (green curve) a multiple of PC3 (here set at 0.5). The two first PC are very similar to those computed for the upper part (as shown in Fig. 4.2).

4.3.1 Step 1: selecting the training samples

The first step consists in the selection of the training data using procedures similar to the ones described in Sect. 3.2. For each of the randomly selected training sample $\boldsymbol{x}_i$, a functional model output $\boldsymbol{y}_i(t)$ is calculated by running the computationally intensive landslide model. The set of n_0 pairs of the form $\{\boldsymbol{x}_i; \boldsymbol{y}_i(t)\}$, with $i = 1,2,...,n_0$, constitute the training data on which the meta-model is constructed.

4.3.2 Step 2: reducing the model output dimensionality

As the model output is functional, the procedure described in Sect. 4.2 is conducted to reduce the dimensionality of the functional model output. Step 2 results then in a set of n_0 pairs of the form $\{\boldsymbol{x}_i; h_{ik}\}$, with h_{ik} the weight of the k^{th} PC (i.e. the PC scores), with $k = 1,2,...,d$ and $i = 1,2,...,n_0$. See Sect. 4.2 for further discussion on the choice of d and on the interpretation of each PC.

4.3.3 Step 3: constructing the meta-model

Using the training data, the scores for each PC can then be approximated as a function of the input parameters $\boldsymbol{x}$, i.e. by a meta-model. Different types of meta-model exist as described in Sect. 3.2. Here, the meta-model of type Projection Pursuit was chosen. Only the basic concepts are described below. For a more complete introduction and full derivation of equations, the interested reader can refer to [Friedman and Stuetzle, 1981].

Table 4.1: Main steps of the meta-modelling strategy for dynamic sensitivity analysis using PCA.

Step	Description
1	Generate n_0 different values for the input parameters $\boldsymbol{x}$ of the landslide model using a Latin Hypercube Sampling technique;
2	Evaluate the corresponding time-dependent landslide model outputs $\boldsymbol{y}$(t); Perform the PCA analysis using these training data: • 2.1: choose the dimensionality reduction d by analysing the fraction of the variance accounted for by each Principal Components PC; • 2.2: analyse the main modes of variations by interpreting the d PC as perturbations around the mean temporal function;
3	Based on the training data and on PCA, construct a meta-model approximating the scores of each PC;
4	Assess the approximation and the predictive quality, e.g. using cross-validation procedure;
5	Using the "costless-to-evaluate" meta-models, compute the Sobol' indices and analyse the importance of each of the input parameters on each PC (seen as the main modes of variation of the landslide time-dependent output using interpretations conducted in step 2).

Let us define $\hat{f}$ as the meta-model and $\boldsymbol{x}$ the n-dimensional vector of input parameters. This non-parametric regression technique assumes that $\hat{f}$ takes the form:

$$\hat{f}(x) = \sum_{k=1}^{M} g_k(\boldsymbol{\alpha}_k \cdot \boldsymbol{x}) \tag{4.3}$$

where the n-dimensional vectors $\boldsymbol{\alpha}_k$ and $\boldsymbol{\alpha}_m$ are orthogonal for $k \neq m$; the term $\boldsymbol{\alpha}_m \cdot \boldsymbol{x}$ corresponds to a linear combination of the elements of $\boldsymbol{x}$; g_k is estimated using some flexible smoothing method. The vectors $\boldsymbol{\alpha}_k$, the function g_k and the dimension M are determined in an iterative manner (see algorithm described in [Friedman and Stuetzle, 1981]).

Basically, this type of meta-model takes the form of the basic additive model, which fits the projection of $\boldsymbol{\alpha}_k \cdot \boldsymbol{x}$ rather than $\boldsymbol{x}$, so that the $\boldsymbol{\alpha}$ directions are chosen to optimise the model fit. Thus, projection pursuit regression technique involves additive modelling and dimension reduction with M usually smaller than n. By using functions of linear combinations of the elements of $\boldsymbol{x}$, this technique allows accounting for variable interactions and non-linearity in the true numerical model f.

4.3.4 Step 4: validating the meta-model

To assess the approximation quality, the differences between the approximated and the true quantity of interest (i.e. the residuals) are usually used. On this basis, the coefficient of determination R^2 can be computed (similarly as Eq. 3.1): here the quantity of interest is the score for a given PC. However, estimating a coefficient of determination for each PC score may not be easily interpretable (what is the physical meaning of R^2 of 98 % on the first PC?). Besides, this prevents from assessing the impact of the error resulting from both the use of a meta-model (i.e. related to the fact the "true" simulator is replaced by it) and the use of PCA to approximate the time series (an illustration of this type of error is provided in Fig. 4.3B).

An alternative is then to compute the temporal evolution of R^2, which can be achieved by reconstructing the functional model output, i.e. by transforming the estimated scores for each PC in the “physical” domain of the functional model output (using Eq. 4.1 and 4.2). This procedure presents the advantage of clearly highlighting the time domain, where the approximation is of poor quality.

To assess the predictive quality, cross-validation procedures can be used (see Sect. 3.2). In the case of functional outputs, this technique can be performed as follows: 1. the initial training data are randomly split into q equal sub-sets; 2. a sub-set is removed from the initial set; the basis set expansion is performed using the $q-1$ remaining functional observations, so that the eigenvectors and eigenvalues are re- evaluated; 3. a new meta-model associated to the calculated PC scores is constructed; 4. the sub-set removed from the initial set constitutes the validation set; the PC scores of the validation set are estimated using the new meta-model; 5. the functional observations of the validation set are then “re-constructed” using the estimated PC scores; the time-dependent residuals (i.e. the residuals at each time step) are then estimated. This procedure is carried out in turn for each of the q sub-sets and corresponds to the “leave-one-out” cross validation, if each sub-set is composed of a single observation.

Using the time-dependent residuals (computed for the q iterations of the cross-validation procedure), the predictive quality can be assessed regarding the temporal evolution of the coefficient of determination. Once validated, the costless-to-evaluate meta-models can be used to estimate the PC scores at any “yet-unseen” values of the input parameters and can be used to conduct VBSA using the Sobol’ indices. As underlined in Sect. 4.2, if the considered PC component is sensitive to a particular input parameter, this means that this parameter is important in producing the type of behaviour in the model output as analysed in step 2.

4.4 Application to the La Frasse case

4.4.1 Construction of the meta-model

The afore-described methodology (see summary in Table 4.1) was applied to the La Frasse case. In practice, the packages of the R software [R Core Team, 2014], named "sensitivity" [1] and "modelcf" [2] developed by [Auder et al., 2012] for meta-modelling of functional model outputs were used.

A set of 30 time-dependent horizontal displacements were calculated for 30 different configurations of the slip surface properties (7 input parameters, see description in chapter 3), which were randomly chosen using a LHS technique. Using these training data, the PCA analysis as described in Sect. 4.2 was carried out and provides the scores for two PCs in the upper part and three ones in the lower part.

The scores for each PC were approximated using a meta-model of 7 input parameters. Different types of meta-model were tested (not shown here) and the approximation and the predictive quality were assessed following the procedure described in Sect. 4.3. For our case, the Projection Pursuit Regression technique presented a good trade-off between high levels of both approximation and predictive quality (see discussion below) and simplicity of the mathematical formulation of the meta-model.

Regarding the approximation quality, the coefficient of determination R^2 (Fig. 4.5A) steeply increases over time and rapidly reaches very high values (> 99.9%) after the time instant of 50 days for both parts of the landslide. Regarding the predictive quality, a leave-one-out cross-validation procedure was conducted. Figure 4.5B provides the temporal evolution of R^2_{CV}. This measure of predictive quality reaches high values (in average ≈ 98%) for the upper part of the landslide (straight line in Fig. 4.5A) after the time instant of 50 days, whereas the values appears to be lower (in average 85 %) for the lower part of the landslide, but can still be considered "satisfactory" (for instance, [Storlie et al., 2009] used a threshold at 80 %).

The analysis of the meta-model quality should interpreted with respect to the "heterogeneity" of the temporal outputs (Fig. 4.1). Though most temporal outputs cover the range 0 to 0.1m in the upper part (only six samples cover the range 0.1 to 0.7m, see Fig. 4.1B), the overall temporal evolution of the horizontal displacements appears to be well accounted for (both in terms of approximation and of prediction). On the other hand, the whole range of displacements' variation appears quite well covered in the lower part (see Fig. 4.1C), but the approximation

[1] available at http://cran.r-project.org/web/packages/sensitivity/index.html
[2] available at http://cran.r-project.org/src/contrib/Archive/modelcf/

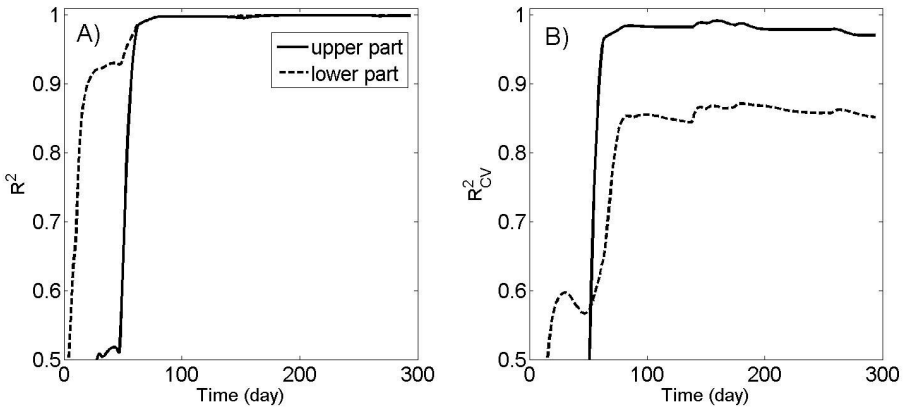

Figure 4.5: A) Temporal evolution of the coefficient of determination R^2 calculated based on the approximated scores of PC1,2 (upper part of the La Frasse landslide, straight line) and of PC1,2,3 (lower part of the La Frasse landslide, dashed line); B) Temporal evolution of R^2_{CV} resulting from the Leave-One-Out cross validation procedure in the upper (straight lines) and lower part (dashed lines) of the La Frasse landslide.

remains of lower quality than for the upper part (as illustrated by Fig. 4.3B). The explanation should be sought in the presence of a few specific steeply-increasing temporal curves (the ones reaching the values of displacements greater than 1.0 m at the end of the time period): their temporal mode of variation (related to PC3, see Fig. 4.4) is accounted for with greater difficulties as there are too few observations of this type. Improvements can be achieved through additional simulations potentially selected using adaptive sampling strategy (see e.g., [Gramacy and Lee, 2009]).

4.4.2 Computation and analysis of the main effects

The main effects for each PC component could then be computed using the costless-to-evaluate meta-model and using, in this case, the algorithm of Sobol' [Sobol', 1990]. Preliminary convergence tests were carried out and showed that using 5,000 Monte-Carlo samples yielded satisfactory convergence of the sensitivity measures to two decimal places. Confidence intervals computed using bootstrap techniques were very narrow so that they are not shown in the following. The total number of simulations reaches 40,000 (considering the number of input parameters of 7). Given the CPU time of nearly 4 days for one single simulation, the VBSA could obviously not have been achievable using the "true" numerical landslide model. By using the meta-model, the CPU time of the sensitivity analysis only corresponds to the CPU time needed for the computation of the 30 training data, (i.e. of 4 days using a computer cluster composed of 30 CPU) and for the validation of the approximation and predictive quality (CPU time < 1 hour).

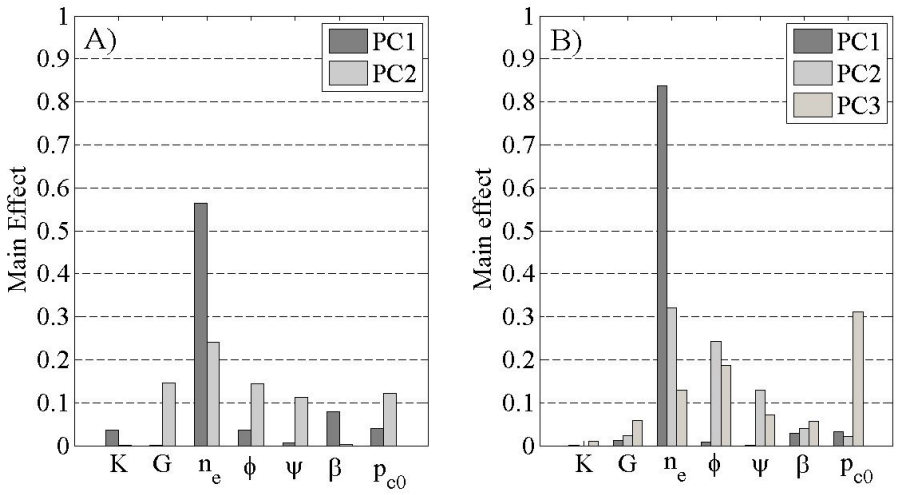

Figure 4.6: A) Main effects (Sobol' indices of first order) associated to each property of the slip surface computed for the two first principal components PC1,2 in the upper part of the landslide; B) Main effects for PC1,2,3 in the lower part of the landslide.

Figure 4.6 provides the results of the sensitivity analysis for the upper (Fig. 4.6A) and the lower part (Fig. 4.6B) of the landslide for each PC as analysed in Sect. 4.2. The analysis of these results can be conducted regarding the goal of the risk practitioner, which can be formulated as follows: "if the goal is to understand the global and major mode of temporal behavior, the analysis should primarily focus on PC1 for the La Frasse case to decide accordingly future investigations and characterization studies".

Considering the upper part of the landslide (Fig. 4.6A), the non-linearity coefficient n_e presents the greatest influence on the main mode of variation accounted for by PC1 (with a main effect almost reaching 60 %). This means that this slip surface property activates the overall vertical up-down shift behaviour of the time-dependent displacements as discussed in Sect. 4.2 i.e. a possible unstable mode of variation over the whole time duration. Interestingly, this result is in good agreement with the analysis of Chapter 3 using time-varying VBSA. This result can have strong implications for future characterization tests, because the measurement of n_e is known to be difficult as requiring lab tests at small strains (see discussion in Sect. 3.6).

Though contributing to a lesser extent to the variability of the time-dependent displacements (<1%, see Sect. 4.2), the analysis of PC2 can be of great interest if the risk practitioner aims at understanding the occurrence of an “acceleration” behaviour after the second pore pressure peaks. The VBSA analysis shows that the coefficient n_e also influences this mode of variation but with lower contribution (main effect of $\approx$ 25%). The second most influential parameters are the shear modulus G, the angle of friction ϕ, the angle of dilatancy ψ and the initial

critical pressure p_{c0} with a main effect ranging between 10 and 15%. Activation of the mode of variation represented by PC1 is primarily driven by one dominant parameter, whereas the activation of PC2 is dictated by five parameters with almost equivalent contribution (ranging between 10 and 20 %).

Considering the lower part of the landslide (Fig. 4.6B), the clear influence of n_e (with a main effect of $\approx 85\%$) is shown on the first mode of variation. The second mode is both influenced by n_e and ϕ with main effect of respectively 32% and 24.5%. The influence of the latter property may be related to the occurrence of irreversible deformation (plastic shear strain) in the lower part of the landslide during the pore pressure changes. If the risk practitioner aims at understanding the occurrence of an "acceleration" behaviour between the second and third pore pressure peak, the analysis of the third mode can be useful (see interpretation of PC3 in Sect. 4.2). Figure 4.6B shows that the third mode is highly influenced by the initial critical pressure p_{c0} (main effect of $\approx 30\%$). This parameter is linked with the initial void ratio and compaction ratio with a relationship depending on the type of soil (clays or sands, see e.g., [Lopez-Caballero et al., 2007]): it determines the initial size of the ellipse-like yield envelope. The second most important parameter is ϕ with a main effect of 20 %: this is also related to the size of the yield envelope (slope of the critical line). This analysis indicates that the complex mode of temporal variation represented by PC3 is highly dependent on the initial yield envelope.

4.5 Towards dealing with functional inputs

In this section, the issue of GSA applied to functional inputs is addressed. In the domain of geo-hazard, a usual situation corresponds to a soil formation characterized by spatially-varying hydraulic conductivities (see for instance [Tacher et al., 2005]). In the following, the analysis is restricted to this case.

4.5.1 Strategy description

In practice, the spatially-varying input is usually discretised onto a mesh model (composed of nodes or grid cells). An illustration is provided in Fig. 4.7 with a heterogeneous soil porosity discretised onto a 100 by 100 mesh cells. A possible approach to conduct VBSA would consist in estimating a sensitivity measure at each node of the mesh model following procedures described in Chapter 2. To deal with the computational burden of the computer code, meta-model could be constructed at each node following procedures described in Chapter 3. Yet, such an approach is hardly feasible in practice: 1. mesh models are usually characterized

by a large number of mesh nodes/cells (typically over 10,000): in the example of Fig. 4.7, it reached 10,000 mesh cells ; 2. if sensitivity analysis is independently conducted at each node, the information on the possible spatial correlation of the input is lost: the approach is said to be redundant. An alternative can then rely on a similar strategy as the one for functional output:

- Step 1: reduce the number of dimensions through a basis set expansion (for instance PCA), which basically consists in projecting the spatial input onto a space of lower dimensionality, while preserving its main statistical properties ;
- Step 2: construct a meta-model on each of these components (of limited number);
- Step 3: conduct VBSA on each component;
- Step 4: interpret the sensitivity similarly as for the case with functional outputs, namely in terms of main modes of spatial variation (see Sect. 4.2).

Steps 1 to 3 of the strategy are usually adopted in the domain of petroleum engineering, but the 4^{th} step is rarely conducted. For instance, [Busby et al., 2007] applied PCA and kriging-based meta-model for oil and gas reservoir modelling, and reduced the spatial fields using 30 components at the cost of 20 % of unexplained variance.

4.5.2 Case study

Let us consider a soil formation, whose porosity's spatial distribution can be represented by a Gaussian random field characterized by: a mean of 15%, a standard deviation of 1%, and an isotropic correlation function of type Gaussian with a range parameter (aka correlation length) of 1,000m and no nugget. Figure 4.7A provides an example of one stochastic realisation of the soil porosity. A total of 250 random maps were generated and were used to compute the principal components PC following the procedure described in Sect. 4.2.

Figure 4.9 shows the evolution of the proportion of explained variance (inertia) as a function of the number of PCs. Contrary to the temporal displacements in the La Frasse case, several tens of PCs are necessary to capture a "satisfactory" amount of information. For instance, 31 and 49 PCs are respectively necessary for the inertia to reach 95% and 99%. Figure 4.7B provides the map reconstructed using 31 PCs to be compared to the original map in Fig. 4.7A. Overall, the approximation appears to be satisfactory with slight differences of the order of 1% (in porosity unit). Considering the 250 samples, the difference can be calculated at each spatial location. Fig. 4.8 provides the differences averaged for the set of 250 samples (expressed in porosity unit in %) at each node with a maximum of 4.5% and a spatial mean of the order of 1 %.

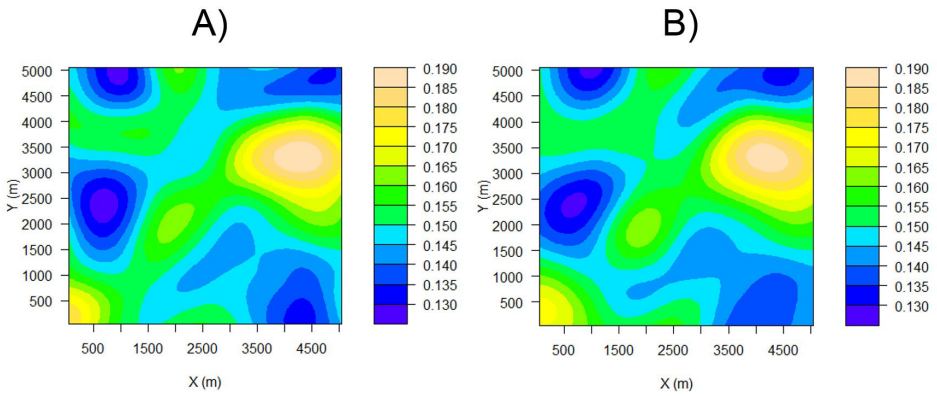

Figure 4.7: A) Example of one stochastic realisation of the soil porosity represented by a Gaussian random field (see text for details); B) Reconstructed map using 31 PCs.

4.5.3 Discussion

Figure 4.9A shows that when increasing the range parameter (length-scale) from 1,000 to 3,000m, the decomposition is "easier" in the sense that the number of PCs decreases. Fig. 4.9B shows that when increasing the anisotropy of the Gaussian correlation (ratio between the distribution along y and along x coordinate) from 1.1 to 2.0, the number of required PCs is increased. Therefore, the more complex the spatial distribution (larger anisotropy, smaller correlation lenght), the larger the number of necessary PCs.

To handle more complex geological structures, several authors have shown the limitations of PCA applied to geological structures. For instance, [Sarma et al., 2008] showed the efficiency of kernel-based (non-linear) PCA over the traditional (linear) PCA to reproduce the essential features of complex geological structures represented by multipoint geostatistics (e.g., [Caers and Zhang, 2004]), like complex channelized sedimentary bodies, and even for large models. Nevertheless, the number of dimensions can still remain large enough (of the order of a few tens) depending on the complexity of the geological setting, e.g., of 30 in the studies by both [Sarma et al., 2008] and [Ma and Zabaras, 2011]: this might pose difficulties for the training phase (i.e. the construction) of the meta-model by imposing a large number of training samples. This is referred to as the "curse of dimensionality" (as originally coined by [Bellman et al., 1961]). For instance, [Busby et al., 2007] and [Ma and Zabaras, 2011] respectively used 200 and 1,000 training samples to propagate spatial uncertainty. When using a computer code with high CPU time of several days like in the La Frasse case, running such a large number of simulations can be challenging even when using an appropriate grid computing architecture and the optimization of computing resources.

To further illustrate, the cost of a two-level full-factorial (each variable is set to either its

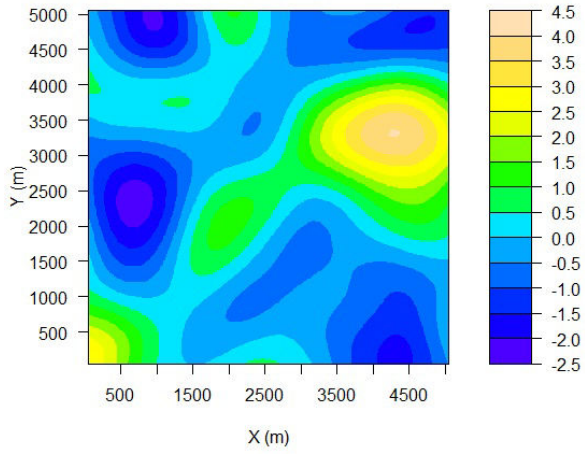

Figure 4.8: Average difference (in porosity unit) using 250 samples (see text for details) between the original and the approximated maps.

lower or upper bound) design of experiments grows as 2^d with d the dimensionality. With such a two-level full-factorial design, the increase of the dimensionality from d=5 to 20 implies the increase of the number of computer runs from 32 to more than one million. Recently, [Hemez and Atamturktur, 2011] performed numerical experiments by applying the well-known empirical rule of "10 runs per dimension" to construct the meta-model: they highlighted the danger of too sparse sampling for the quantification of uncertainty margin.

A second limitation relates to the choice in the truncation level of the expansion, i.e. how much information should be kept? Considering PCA, this is done using the level of total explained variance (Fig. 4.9). When applied to model outputs like in the La Frasse case, such an analysis is carried out using the whole data sets of functional model outputs, i.e. once the computationally intensive numerical simulations have been performed. On the contrary, this level should be decided a priori in the case of the functional model inputs, i.e. before any simulation has been carried out. Investigating whether leaving 5% of total explained variance can have a significant influence on the approximation quality comes then at the expense of additional long-running simulations.

Aside the computational challenge, the interpretation of the sensitivity related to the PCs can also pose problems in practice. In the La Frasse case, this was tackled using the mean temporal function plus and minus some multiple of the PC: this allowed highlighting phases of acceleration or stabilization of the landslide in relation with the most influential PCs (see Sect. 4.4). In the spatially-varying case, the first PC is easily interpretable as it is almost constant. The second one corresponds to two (approximately) anti-symmetric areas (Fig. 4.10A): a

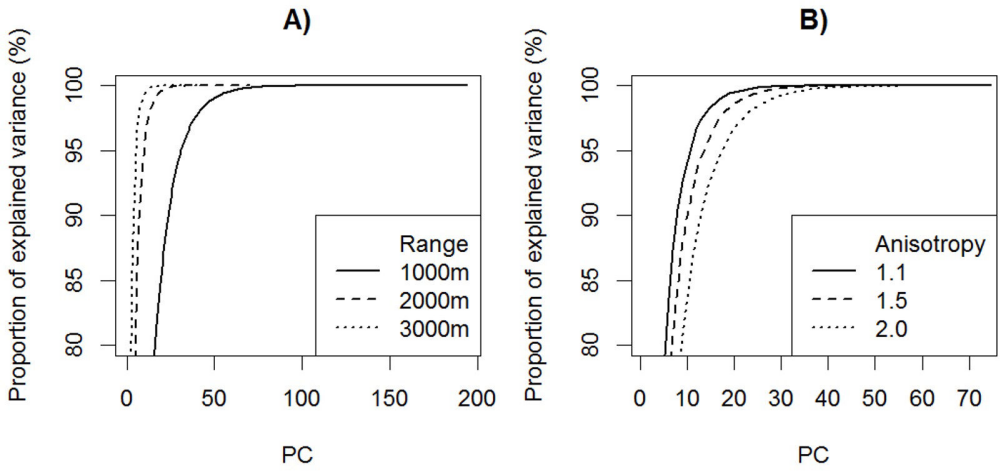

Figure 4.9: Evolution of the proportion of explained variance (inertia) as a function of the number of PC related to the decomposition of the 250 realisations of the soil porosity (see text for details). A) Three cases of range parameter (correlation length) are considered: 1,000; 2,000 and 3,000m; B) Three cases of anisotropy are considered: 1.1; 1.5; and 2.0.

left region presents a peak, whereas the other one presents a valley. The third PC is more complicated to interpret with a region of two peaks and one valley (Fig. 4.10B). The larger the order of the PCs, the larger the number of peaks/valleys (Fig. 4.10C and D). Here, it should be recalled that the decomposition has been carried out considering a "simple" isotropic Gaussian field: the PCs associated to more complex geological features are expected to be more complicated to interpret.

4.6 Concluding remarks of Chapter 4

Chapter 4 tackles the issue of applying VBSA considering two difficulties related to the use of models supporting geo-hazard assessments, namely: 1. these models can have high computation time cost; 2. these models have functional variables: they are usually not scalar, but can vary over time and space.

Using the case of the La Frasse landslide as an application example, the issue of handling functional outputs (here time-varying surface displacements during the landslide crisis of 1994) was first tackled. A methodology to carry out dynamic (global) sensitivity analysis of landslide models was described combining:

- Basis set expansion to reduce the dimensionality of the functional model output;
- Extraction of the dominant modes of variation in the overall structure of the temporal

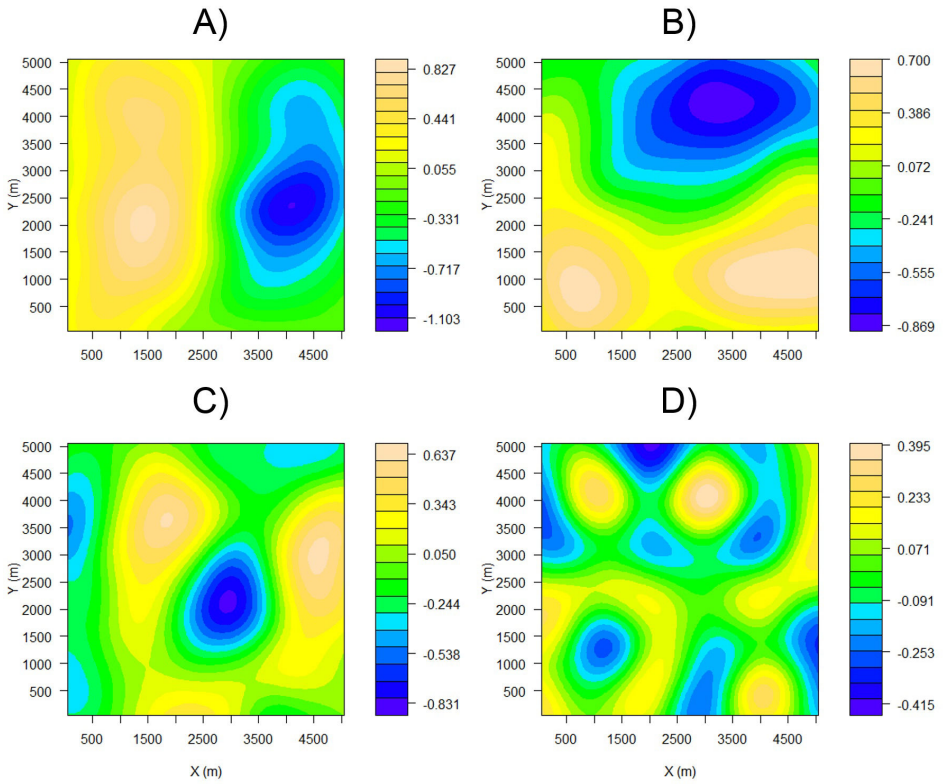

Figure 4.10: Maps of different spatial PCs derived from the decomposition of 250 realisations of the soil porosity (see text for details): A) PC n°2; B) PC n° 3; C) PC n°10 and D) PC n°25.

evolution;

- Meta-modelling techniques to achieve the computation, using a limited number of simulations, of the Sobol' indices assigned to each of the modes of variation.

It was shown how to extract useful information on dynamic sensitivity using a limited number (a few tens) of long running simulations. The analysis of the sensitivity measures assigned to the dominant modes of variation was interpreted by adopting the perspective of the risk practitioner in the following fashion: "identifying the properties, which influence the most the possible occurrence of a destabilization phase (acceleration) over the whole time duration or on a particular time interval".

However, it should be acknowledged that the difficulty of the methodology resides in the physical interpretation of the dominant modes of variation (even viewing them as perturbations of the mean temporal function), especially compared to the traditional time-varying VBSA (more easily interpretable, but also intractable for very long time series with >1,000 time steps). To better investigate this issue, other real-case applications of the methodology should be conducted in the future.

A second difficulty is related to the different uncertainty sources introduced not only by the use of a meta-model (and accounted for by using a cross validation procedure), but also by the basis set expansion procedure on a very limited number of simulations. As a future direction of work, the adaptation of the bootstrap methodology introduced by [Storlie et al., 2009] may be proposed to associate the sensitivity measures with confidence intervals reflecting both meta-model and projection (related to PCA) error.

Based on the study on the functional output, the applicability of the proposed methodology was also investigated for the case of functional inputs (for instance a soil with spatially varying hydraulic conductivities). Nonetheless, three limitations were outlined: 1. The number of necessary components in the basis expansion of the functional input can be very large (over several tens) depending on the complexity of the geological setting: this might impose a large set of training samples in order to reach a satisfactory level of approximation quality; 2. the level of truncation of the basis set expansion should be decided before carrying out any long-running simulations; 3. as afore-mentioned, the physical interpretation of the sensitivity measures assigned to the main modes of variation can be very tedious. Two options can be investigated in the future.

The first option should focus on the improvement of the construction of the meta-model itself to overcome the computational limitation imposed by the number of necessary components in the basis set expansion. Since all the different PCs are not expected to participate in the same manner to the prediction of the model response, the basic idea is to select the most important input variables during the fitting process. In such a situation, the meta-model should both handle a large number of input variables, but relatively small sizes of learning samples. This is a situation of high-dimensional regression. Traditional techniques (such as generalized linear modelling with stepwise selection, see discussion in [Hastie et al., 2009]: chapter 3) have shown limitations in such situations. More advanced techniques have then been proposed by relying on the idea of shrinkage for variable selection: this aims at retaining only a small number of input variables corresponding to the best variables for improving prediction accuracy, and discarding the rest. The interest is then to build a good (i.e. with high accuracy of prediction on future data) nearly unbiased regression model using only a small subset of such co-variates (sparse model).

A second option should focus on the improvement of the interpretability of VBSA applied to functional input. As discussed, using the sensitivity measures assigned to each main modes of variations can be limited. An alternative should then rely on a unique global sensitivity indicator assigned to the whole range of variation of the functional input (following the spirit of [Lamboni et al., 2011] for functional output). In particular, what is of interest from a risk

management perspective is to distinguish whether the uncertainty is driven by the spatial uncertainty (e.g., heterogeneous permeability) or by a property, which can be considered to be scalar (homogeneous), because the characterization studies are different in both cases. In this view, different approaches can be identified in the literature:

- The "trigger parameter" method of [Crosetto and Tarantola, 2001]: this consists of defining an additional scalar parameter which governs the presence or absence of stochastic error and which is added to the spatial input (defined at its nominal value);
- The "macro-parameter" method of [Jacques et al., 2006]: this consists of replacing the spatial (functional) inputs by a set of correlated input parameters;
- The "categorical parameter" method of [Lilburne and Tarantola, 2009]: this consists of assigning a categorical indicator (i.e., a pointer variable which takes discrete values) to the set of spatial maps (possible realizations) of the spatial input.

Although these approaches deal with the high-dimensionality of the input/output variables when conducting VBSA, applying these approaches may still become too time-consuming when they are directly applied to long-running simulators. Additional studies should be carried out to combine these approaches with appropriate meta-modelling techniques.

5 A more flexible tool to represent epistemic uncertainties

The major pillar of the variance-based methods for sensitivity analysis, as described in the previous chapters, is the representation of epistemic uncertainty within the probabilistic setting. Yet, in situations where the data are very scarce, imprecise, vague, diffuse, fluctuating, incomplete, fragmentary, ambiguous, dubious, or linguistic (i.e. an environment of imperfect knowledge), this framework can be questionable: this is further discussed in Sect. 5.1. Alternative uncertain theories have been proposed in the literature (a brief introduction is available in Appendix D). In this chapter, the applicability of Fuzzy Sets [Zadeh, 1965] is discussed for representing different forms of epistemic uncertainties: representation and reasoning with vague qualitative statements (Sect. 5.2 and 5.3), imprecision (Sect. 5.4), and probabilistic model with imprecise parameters (Sect. 5.5). This chapter is mainly based on the work described in [Rohmer and Baudrit, 2011, Nachbaur and Rohmer, 2011].

5.1 On the limitations of the systematic use of probabilities

The probabilistic setting has successfully been used in a broad range of different domains and in particular for geo-technical applications [Einstein and Baecher, 1983, Nilsen, 2000, El-Ramly et al., 2002, Park et al., 2005]. Despite this success, the use of probabilities as a tool to represent epistemic uncertainties has been criticized by several authors [Ferson, 1996, Oberguggenberger and Fellin, 2002, Helton and Oberkampf, 2004, Baudrit and Dubois, 2006, Dubois and Guyonnet, 2011, Beer et al., 2013] in situations where the resources (time and budget) for hazard and risk assessments are limited, and where the available date are imprecise, scarce, incomplete, vague, qualitative, etc.

For instance, consider the situation where a geo-technical engineer has been asked for the stability analysis of an a priori unstable rocky cliff in the vicinity of buildings. The expertise

should be carried out "as quick as possible" taking into account the current data/information on the rock cliff characteristics, using for instance rock-mass classification like GSI index. The available data and information may only consist of:

- an estimate of the geometrical properties (slope angle, cliff height), which can only be done from the cliff toe;
- a qualitative estimate of the fracture distribution, e.g. "rock cliff is moderately to highly fractured";
- a rough idea of the presence of water and nature of the infilling materials of fractures; rock properties' values bounded by plausible physical values (e.g., "friction angle of a stiff dolomite may vary between 40° and 45°"). Such an estimate can possibly be refined by geotechnical surveys, which have been carried out in neighboring regions or in similar contexts.

Other examples can be given: the early design stage of a engineered structure [Beer et al., 2013]; the susceptibility assessment of the presence of abandoned cavities at regional scale like in the French context described by [Nachbaur and Rohmer, 2011]; the screening and selection phase of underground storage sites [Bouc et al., 2009]; the development of Risks Prevention Plans, which impose to work "in the state of knowledge" and "according to expert opinion" as outlined by [Cauvin et al., 2008], i.e. with no other resources that those available at the time of the study, etc.. These often correspond to situations where the resources (time and budget) for hazard and risk assessments are limited.

In such highly uncertain situations, the challenge is to develop appropriate mathematical tools and procedures for "accounting for all data and pieces of information, but without introducing unwarranted assumptions" [Beer et al., 2013]. The crucial concern of the uncertainty representation step can be formulated following [Aven and Zio, 2011]: "the representation of the knowledge available as input to the risk assessment in support of the decision making must be faithful and transparent: the methods and models used should not add information that is not there, nor ignore information that is there".

When a large number of observations are available, the most appropriate representation tool is the probability distribution, which can be inferred from data/observations. An example is the fit of power-law to rockfall volumes [Dewez et al., 2013]. Otherwise, test statistics might fail when the number of samples is low (typically below 20) so that the probability distribution should then be assumed. In some situations, such a law is theoretically known such as the power-law distribution relating the earthquake frequency to the magnitude (Gutenberg Richter law).

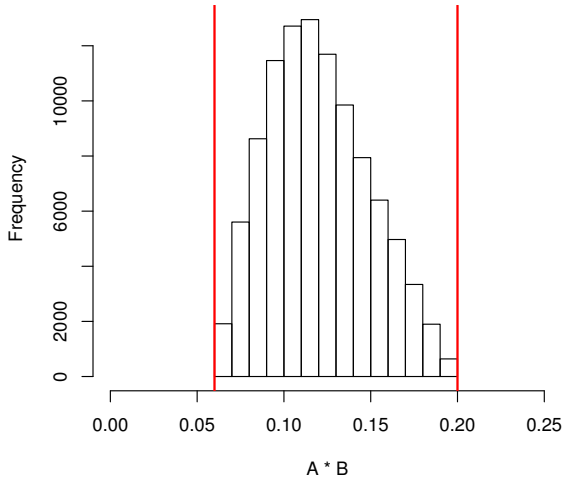

Figure 5.1: Histogram associated with $R = A \times B$ resulting from the Monte-Carlo-like sampling of two uniformly distributed random variables A and B (see text for details).

But in other situations, the distribution should be selected. This can rely on the principle of maximum entropy [Gzyl, 1995, Mishra, 2002]. For instance, when only bounds on the uncertain parameter are available, the application of this method leads to the selection of a uniform distribution. By using the example of an uncertain load acting on a floor beam with a strict requirement for the maximum deflection, [Beer et al., 2013] criticized this approach (stated as "disconnected from the engineering context"), because this would lead to an averaged result for the deflection, whereas the interest is on the maximum value. Besides, making the assumption of a uniform distribution introduces additional information: the original data only suggest that the uncertain load lies within bounds, whereas the probabilistic setting leads to assume that all possible loads have the same weight. This results in a change in the character of the information, which might impact the risk assessment's outcome.

More formally, this was discussed by [Ferson and Ginzburg, 1996] by using the formal example of the uncertain variable $R = A \times B$, with A and B two random independent variables respectively represented by an uniform distribution with support $[0.20; 0.40]$ and $[0.30; 0.50]$. Under the assumption of independence, the Monte-Carlo-based propagation procedure (using 1,000 random samples) results in a clear concentration of probability distribution of R as shown in Fig. 5.1.

Starting with inputs associated with very limited knowledge (here only the bounds), the choice of uncertainty representation (here the uniform distribution) virtually led to "more precisely

constrained" output information as shown by the concentration of the histogram of the output in Fig. 5.1. On the other hand, if the treatment was conducted with a strict consistency with the available data (here the bounds), the only information on the output would have been the bounds: 0.06 and 0.20 (outlined in red in Fig. 5.1).

The pure-probabilistic representations have been criticized for inducing an appearance of more refined knowledge with respect to the existing uncertainty than is really present [Klir, 1989, Klir, 1994]. When experimental data on material properties are insufficient to make a distinction between several probability distributions (e.g., gamma and lognormal), this choice may sound "arbitrary" (see. e.g., [Ditlevsen, 1994]). In the domain of geotechnical risk assessment, [Oberguggenberger and Fellin, 2002] showed that the estimates of failure probabilities are highly sensitive on this choice. On the other hand, [O'Hagan and Oakley, 2004] argued that the problem does not stem from the use of probability as such, but from the elicitation process (i.e. the procedure consisting in expressing the experts' knowledge and beliefs in probabilistic form). Nevertheless, in situations of data/information scarcity, expressing information in terms of mean, variance, or any statistical quantities may appear tedious, if not debatable. More specifically, this is supported by the results of some cognitive studies [Raufaste et al., 2003]. As outlined by [Dubois and Prade, 1994], the probability setting may be often too "rich" to be currently supplied by individuals: the identification of the probability distribution requires more information than what an expert is able to supply, which is often restricted to the 0.50 and 0.95 fractiles or a prescribed mode. Given these pieces of information, many mathematical probabilistic laws may exist [Dubois and Prade, 1994]. Hence, when the knowledge is restricted to the bounds (like in the afore-described cases), there is an infinity of probabilistic laws defined by such a support.

Alternative approaches in the probabilistic setting rely on the use of Bayesian methods (see examples in structural safety assessment provided by [Igusa et al., 2002] and brief introduction in Appendix D): this allows mixing subjective and objective information, i.e. perception regarding a probabilistic model (subjective probabilities consistent with the axioms of probability), and observations/data for model update. In this approach, a unique probability distribution represents the state of knowledge of experts. However, this may appear debatable in the phase of information collection, i.e. in the early phase of uncertainty treatment (see Fig. 1.3). In this approach, subjectivity is introduced at "the very beginning of the risk analysis chain, whereas it would make more sense to appear at the very end to support decision-making" [Dubois and Guyonnet, 2011]. In geo-technical risk assessments, [Oberguggenberger and Fellin, 2002] conveyed a similar reasoning by underlying the shortcomings of the probabilistic settings to interpret a failure probability as a frequency of failure. On the other hand, they also clearly outlined the usefulness of probabilities as means for

decision-making under uncertainty (i.e. at the final phase of the risk analysis chain), in particular for comparative studies of scenarios. Besides, this subjectivity influence only decays with a growing amount of data [Beer et al., 2013], which may not be always available in the situations considered here.

Alternatives to the probabilistic setting for representing epistemic uncertainties when data are scarce, imprecise or vague have then been proposed in the literature (see a brief introduction in Appendix D). Yet, it should be recognized that a consensus has not yet been reached on how to standardize such practices. The subsequent sections focus on the use of Fuzzy Sets (originally developed by [Zadeh, 1965]) as a tool for processing uncertain data. It is demonstrated how flexible and useful this tool can be in different situations of epistemic uncertainty encountered for geo-hazard assessments, namely:

- Representing and reasoning with vague concepts (Sect. 5.1 and 5.2);
- Handling imprecision (Sect. 5.3);
- Dealing with a probabilistic model whose parameters (like the mean, the variance, etc.) are imprecise (Sect. 5.4).

5.2 Handling vagueness

5.2.1 A motivating real-case: hazard related to abandoned underground structures

It is estimated that France contains more than 500,000 of abandoned underground structures whose partial or total ruin can have considerable socio-economic consequences for the community, e.g. [Nachbaur and Rohmer, 2011]. For example, in the French region of Picardy alone, more than 300 constructions were damaged through cavity collapse following the winter rains of 2000-2001 [Bouchut and Vincent, 2002]. These structures can correspond to quarries and marl pits, but also to various other types as diverse as war saps (covered frontline trenches), underground shelters, troglodyte dwellings, etc. [Tritsch et al., 2002]. In addition to these anthropogenic structures are the "natural" cavities such as the karsts in limestone environments.

In France, cartographic tools known as Risk Prevention Plans have been developed at municipal scale for determining cavity-associated risks [MATE, 1999]. However, faced with both the number and the diversity of such abandoned cavities, the authorities require decision-aid tools to be able to rank the risks at spatially larger scales (such as grouped-municipality, if not regional scale) and manage the resultant uncertainties [Waltham et al., 2005]. Contrary to

other natural phenomena, no single random variable can be identified for the overall measurement of the dreaded event, i.e. cavity collapse. In most cases, the studied underground structures are not accessible, which eliminates the use of a systematic deterministic approach. The particularity of mining cavities is the precision often provided by the existence and knowledge of mine plans and geometric parameters, even when incomplete. However, in the case of abandoned underground structures, limited and non-exhaustive input data are available, and they are seldom supported by geometric and mechanical parameters. At large spatial scales, the only predictive models that can be established consist in expressing qualitatively (or semi-quantitatively) the spatial probability of a surface instabilities appearance, known as "susceptibility".

Hence, assessing the presence/collapse susceptibility of abandoned cavities at a regional scale is associated with large uncertainties that are mainly related to the very nature of the phenomena, but also to the difficulty in collecting exhaustive information on often "forgotten" structures at such a spatial scale. In this context, the expert's role is essential, because he/she is able to synthesize the information resulting from the inventory and from the commonly imprecise, if not vague, criteria on the basis of his/her experience and his/her knowledge of the regional geological and historical and economic context. Fuzzy sets can provide a useful and flexible tool for processing such pieces of information.

5.2.2 Membership function

Let us consider the concept of membership function, which defines how each element x of the input space X (also named "universe of discourse") is mapped to a degree of membership (denoted μ). Under the classical theory of Boolean logic, the membership function of a set A is simply defined as a binary function that takes the value $\mu(x) = 1$ if the element belongs to A and the value $\mu(x) = 0$, otherwise. In Fig. 5.2A, the set A is graphically represented by a clearly defined boundary.
[Zadeh, 1965] originally observed that: "more often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership". This emphasizes the gap, which exists between the power of expressivity of real numbers and the limited level of precision found in mental representations of reality accounted for via natural language terms as outlined by [Dubois et al., 2001]. Examples in geo-hazard assessments are: "the density of fracture is moderate"; "the fault is highly conductive"; it is plausible to find karsts at a given depth", etc. The Fuzzy set theory introduces the concept of a set without a crisp (i.e. clearly defined) boundary. Such a set $\tilde{A}$ can contain elements with only a gradual

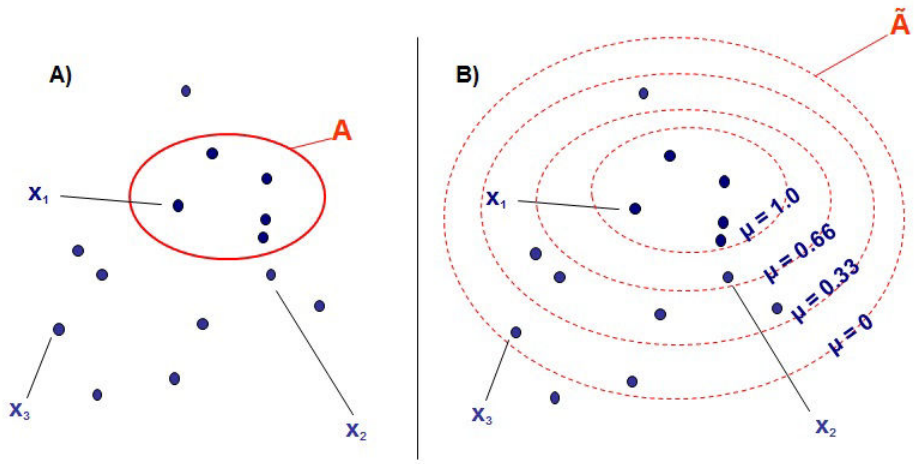

Figure 5.2: Graphical representation of the set A: A) under the classical Boolean theory; B) under the fuzzy set theory.

(partial) degree of membership (scaled between 0 and 1) as defined by Eq. 5.1.

$$\begin{aligned} \mu_{\tilde{A}} &\rightarrow [0;1] \\ \tilde{A} &= \{x;\mu_{\tilde{A}}(x)|x \in \tilde{A};\tilde{A} \subset X\} \end{aligned} \tag{5.1}$$

Graphically, the boundary of the fuzzy set $\tilde{A}$ is a progressive boundary (Fig 5.2B), so that the element x_2 depicted in Fig 5.2B is located in the "fuzzy" zone between the so-called "certain" boundaries (boundaries respectively assigned to $\mu = 0$ and $\mu = 1$) and is associated with a 66 % degree of membership, whereas under the classic Boolean theory, it is entirely excluded from the set A (Fig 5.2A).

Let us consider a usual example of vague concept: evaluating the range of ages assigned to the group of A="old people". In this case, the input space (i.e. universe of discourse) is all potential ages. By defining two thresholds, say 65 and 75 years, it is possible to assign A a gradual membership function so that people with age: i. superior to 75 years have $\mu = 1$; ii. below to 65 years have $\mu = 0$; and linearity decreasing grade μ from 1 to 0, otherwise.

5.2.3 Application

In the context of abandoned underground structures, the vague concept of "high susceptibility of presence" could be handled using the frequently-used criterion "depth of the worked geological formation" [Nachbaur and Rohmer, 2011]. In this case, the universe of discourse is

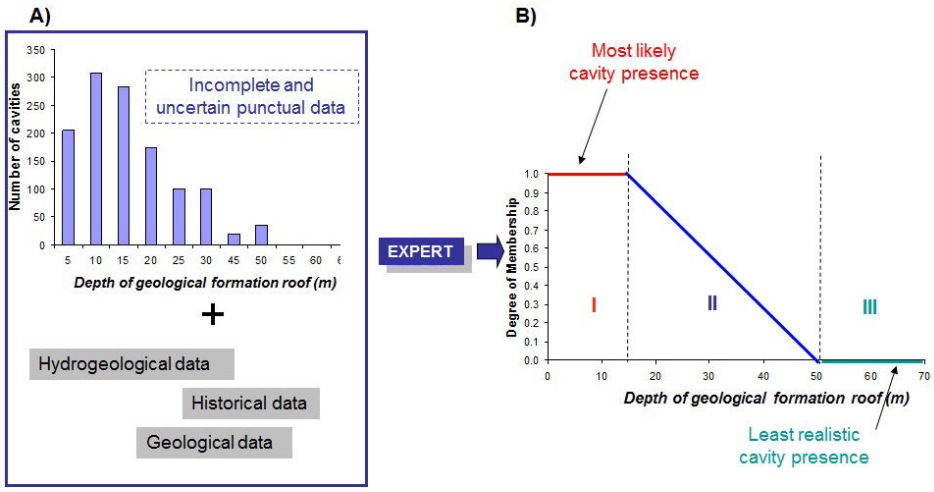

Figure 5.3: Construction of the membership function associated to the criterion for cavity susceptibility assessment using the depth of the exploited geological formation [Nachbaur and Rohmer, 2011].

all potential depths of the worked geological formation (equivalent to the ages in the previous example). Based on the collected data (Fig. 5.3A), as well as on regional knowledge of the context (both geological and historical), the expert could, for instance, state that "in this type of geological formation, I am certain that there are quarries at less than 15m depth and that, for technical mining conditions, no quarry could be present below 50m". Under the traditional (Boolean) approach, if the threshold is fixed at 15m, a cavity located at 16m is automatically and unfavourably placed in the same class as that located at 50m. The entire sense of the "gradual membership function" is thus to free one from data insufficiency [Ercanoglu and Gokceoglu, 2004]. Fuzzy sets can here be used to alleviate these so-called "threshold effects": a fuzzy set $\tilde{A}$ can be defined based on the depths to which one may encounter quarries, accompanied by a degree of membership measuring their influence on the event A="susceptibility of cavity presence is high". Taking the above example, Fig. 5.3B shows that the depths within the 0-15m interval (Zone I) are characterized by the maximum degree of membership. Below this depth, in the 15-50m interval (Zone II of Fig. 5.3B), the degree of membership decreases until below 50m depth (Zone III, Fig. 5.3B), where the expert estimates that the presence of a quarry is not possible, which corresponds to a null degree of membership. Here, the proposed transition is linear (Zone II), but the expert could assume a nonlinear curve. For instance, if the expert has knowledge (or observations) indicating that, while values located within zone II are possible, they are nevertheless very unlikely: convex functions can then be used [Baudrit et al., 2007b].

Thus, the gradual membership function introduced by [Zadeh, 1965] allows handling the

vagueness caused by the gradual nature of words [Dubois et al., 2001], here the "susceptibility is high". Note that vagueness corresponds to a broad class of uncertainty, which also includes ambiguity (see an extensive discussion by [Dubois et al., 2001]). This kind of vague human knowledge can thus be represented in a human-friendly, yet rigorous way. Interestingly, this concept joins the type of continuous function intuitively defined by [Thierry et al., 2009] (but without explicitly referring to the Fuzzy Set Theory) to assess gypsum dissolution collapse susceptibility from observations.

5.3 Reasoning with vagueness

The afore-described notion of gradual membership is the basis for Fuzzy Logic [Zadeh, 1975], which formalizes empirical and intuitive reasoning which the experts may establish from a few vague data. In particular, the application of this fuzzy logic was brought to rock mechanics by [Brown, 1979], and appears particularly well suited for geotechnical problems (see e.g. [El-Shayeb, 1999] and references therein).

5.3.1 A motivating real-case: the inventory of assets at risk

Let us illustrate the application of this approach considering the inventory of assets-at-risk in the context of the Lourdes earthquake scenario (see the setting in Fig. 5.4-left). In this case, the studied urbanized area is subdivided into districts, which define sets of buildings. Each building is classified into vulnerability typologies, which are groups of different types of buildings that may perform in a similar way during earthquakes. Within each defined district, the ratio of vulnerability typologies is evaluated by compiling an inventory based on either aerial photos or, site surveys (i.e. through visual inspection of a limited number of buildings chosen within the district).

In a classical statistical approach, this inventory would be developed from a representative sample i.e. a random sampling of the considered building set (see full description in [Bertil et al., 2009]). In such a context, the uncertainties would depend on the size of the sample. However, compilation of the buildings' inventory in each district faces several constraints in practice:

- financial/time constraints prevents the inspection of each building in detail since it is often carried out from the roadside;
- due to the spatial extension of the district, the expert is only able to inspect a few buildings, where it is possible to do so;
- both site surveys and analysis of aerial photos necessarily involves human judgements

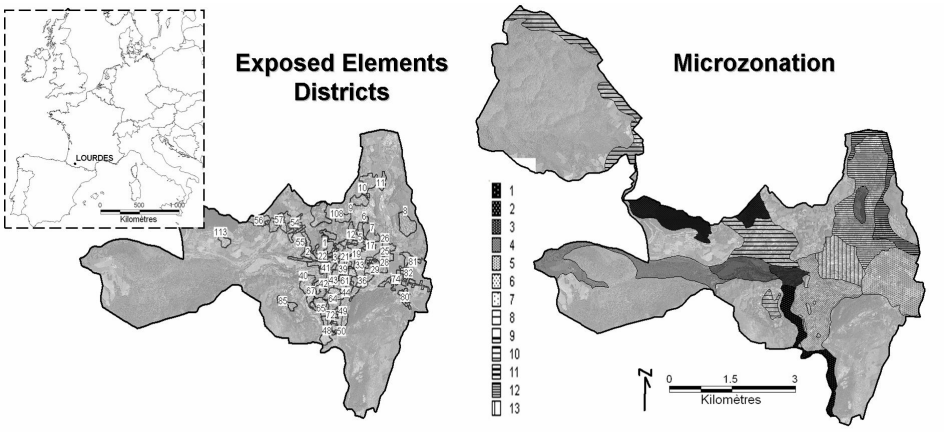

Figure 5.4: Situation of the French city of Lourdes: vulnerability analysis with the defined sets of buildings (on the left) and the defined geotechnical zones (on the right)

(in particular during the site surveys) though the recent use of advanced satellite- and ground-based remote sensing techniques have shown very promising results as demonstrated by the study of [Wieland et al., 2012];

- the heterogeneity (number) of typologies within a given district affects the representativeness of the sample as well.

In such a context, the probabilistic approach shows limitations and experts may only be able to qualify the degree of imprecision in the estimates of the ratio of vulnerability typologies, i.e. to state whether the imprecision in the inventory is “low” or “medium” or “high”. At this level, two sources of uncertainty should be dealt with: i) how to represent the vagueness linked with the qualified statement of imprecision (see Section 5.2)?; and ii) how to decide which class of imprecision to choose (i.e. decision under vagueness)?

5.3.2 Application of Fuzzy Logic

Experts may feel more comfortable in asserting a range of values from their specific or generic knowledge about each qualified imprecision statement than a single crisp value (e.g., 15% of imprecision). To handle this type of vagueness, each class of imprecision “low”, “medium” and “high” can be assigned a trapezoidal fuzzy set, which can be constructed as follows. By aggregating several experts’ opinions, the imprecision is preferentially judged as “low“ for a percentage ranging from 0 to 1%, but the value 5% is judged “still relevant” to be taken into account. It is interesting to note that this type of transformation of a vague concept (here “inventory with low imprecision”) into a fuzzy set allows a high degree of “precisiation”: this latter operation consists in transforming an object into another one, which in some specified sense is defined more precisely than the original one: this is a powerful feature of fuzzy logic

Table 5.1: Numerical choices for imprecision assessment considering the inventory of assets at risk at Lourdes. Density is expressed in numbers of building per km². Heterogeneity is expressed in numbers of different vulnerability typologies. Imprecision of the proportion of vulnerability classes in a given district is expressed in %.

Qualified statement	Core	Support
Low density	[0;750]	[0;1500]
Medium density	[1500;2000]	[750;3500]
High density	[3500;5000]	[2000;5000]
Low heterogeneity	[0;2]	[0;3]
Medium heterogeneity	[3;4]	[2;5]
High heterogeneity	[5;6]	[4;6]
Low imprecision	[0;1]	[0;5]
Medium imprecision	[5;10]	[1;15]
High imprecision	[10;30]	[15;30]

Table 5.2: Logical rules for imprecision assessment considering the inventory of assets at risk at Lourdes.

	Low heterogeneity	Medium heterogeneity	High heterogeneity
Low density	Low Imprecision	Low Imprecision	Medium Imprecision
Medium density	Low Imprecision	Medium Imprecision	High Imprecision
High density	Medium Imprecision	High Imprecision	High Imprecision

[Zadeh, 2008].

In the Lourdes case, logical rules (see Table 5.2) exist between two decision criteria, namely the district density (in terms of numbers of buildings per km²), the heterogeneity of the typologies (measured in terms of numbers of vulnerability classes), and the resulting imprecision of the inventory. Table 5.2 should be read as follows: IF “Low density” AND “Low heterogeneity” THEN “Low imprecision”. Such statements are named “Fuzzy rules”. Each decision criterion is also qualified using the following statement “low”, “medium” and “high” and to each qualified statement, a fuzzy set is assigned. The assumptions for the fuzzy set definition are given in Table 5.1.

Figure 5.5 illustrates the methodology, which is divided into three main steps.

- Step A) “Characterization”. a fuzzy set is assigned to each qualified decision criterion. The membership values in each decision criterion class of the considered district are then estimated;
- Step B) “Combination”. A district may be a member of a fuzzy imprecision class “to some degree” depending on its density value and on its heterogeneity (in terms of number

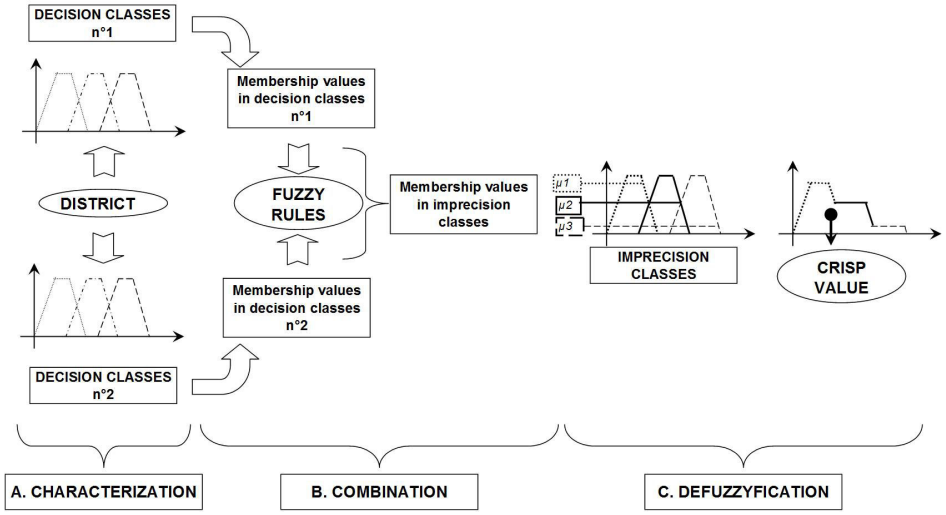

Figure 5.5: Methodology for inventory imprecision assessment adapted from the approximate reasoning [Zadeh, 1975].

of vulnerability classes). Given step A), the membership values μ of the corresponding district in each imprecision class are determined from the min/max combination approach (e.g. [Cox, 1994]). The membership function associated with the uncertainty on the imprecision evaluation is then constructed based on the combination of the fuzzy sets associated with each qualified statement ("low imprecision", "medium imprecision" and "high imprecision") and weighed by the corresponding membership values μ;

- Step C) "Defuzzyfication". The membership function associated with the uncertainty on the imprecision evaluation is then converted to a crisp value synthesizing the vagueness of the qualified imprecision. Here, the chosen "defuzzyfication" method is the "centroid" method (e.g. [Cox, 1994]). Graphically, this method consists in calculating the centre of gravity of the area under the curve of the membership function (i.e. "centroid"). The x-coordinate of this "centroid" represents the "defuzzified" value, which provides an indicator for the vagueness of the qualified statement related to the imprecision in the inventory of the vulnerability classes in each city district.

To illustrate, the Lourdes district n°18 is characterized by two vulnerability classes and 2,542 buildings per km². The corresponding membership (following step A) values are: μ(low heterogeneity) = 0.5; μ(medium heterogeneity) = 0.5; μ(high heterogeneity) = 0; μ(low density) = 0; μ(medium density) =0.64 and μ(high density) = 0.36. The combination under the min/max approach (step B) gives the results in Table 5.3. The resulting membership values in the

imprecision classes are: μ(low imprecision) = Max(0.5, 0, 0) = 0.5; μ(medium imprecision) = Max(0.36, .5, 0) = 0.5 and μ(high imprecision) = Max(0.36, 0, 0.36) = 0.36. For the Lourdes district n°18, the "defuzzyfication" process (step C) gives a value of 15.5%: this corresponds to the estimate of the imprecision on the proportion of each vulnerability classes in this district. This value synthesizes the whole vagueness associated with each qualified statement ("low imprecision", "medium imprecision" and "high imprecision"), each of them being weighted by the membership values estimated from the qualified statements for the district heterogeneity and density.

Table 5.3: Imprecision assessment in the district n°18 (2 vulnerability classes ; 2,542 buildings / km²).

	Low heterogeneity	Medium heterogeneity	High heterogeneity
Low density	Min(0;0.5)=0	Min(0;0.5)=0	Min(0;0)=0
Medium density	Min(0.64;0.5)=0.5	Min(0.64;0.5)=0.5	Min(0.64;0)=0
High density	Min(0.36;0.5)=0.36	Min(0.36;0.5)=0.36	Min(0.36;0)=0.36

The afore-described procedure can be improved by accounting for imprecise values of heterogeneity and density, i.e. by accounting inputs represented by fuzzy sets instead of crisp values. This can be achieved, for instance, by using the procedures termed as β-cut of [El-Shayeb, 1999].

5.4 Handling imprecision

Hereafter, a third situation of epistemic uncertainty is addressed, namely when the (numerical) value of a given model parameter cannot be precisely estimated owing to sparse data sets, i.e. the parameter is imprecise. This can be handled within the theory of Possibility, which is strongly related to fuzzy sets [Dubois and Prade, 1988, Zadeh, 1978].

5.4.1 Possibility theory

Let us first define a possibility distribution of some imprecise quantity x, which can takes on alternative values in the set U. Such a distribution assigns to each element u in U a degree of possibility $\pi(u) \in [0;1]$ of "being the correct description of a state of affairs" [Dubois, 2006]. This possibility distribution is a representation of what the expert knows about the value of x ranging on U (not necessarily a random quantity): it reflects the more or less plausible values of the unknown quantity x. These values are assumed to be mutually exclusive given that x takes on only one value, i.e. its true value. When $\pi_x(u) = 0$ for some u, it means that $x = u$ is considered an impossible situation. When $\pi_x(u) = 1$, it means that $x = u$ is just "unsurprising,

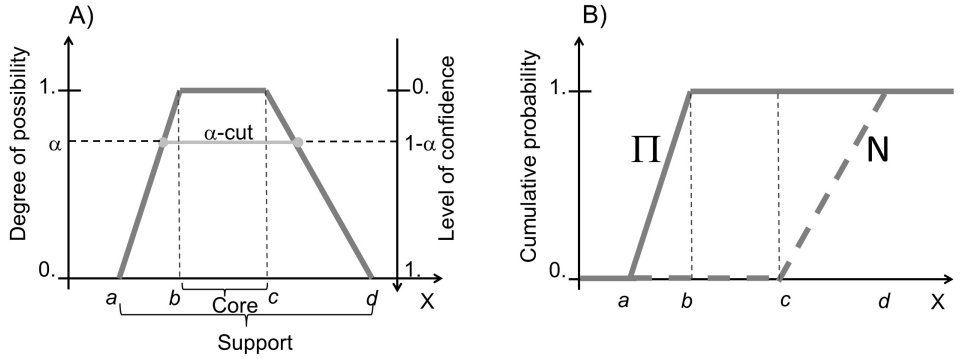

Figure 5.6: Illustration of a possibility distribution associated with an imprecise parameter (A) and definition of the measure of possibility Π and necessity N (B).

normal, usual". Yet it should be underlined that it is "a much weaker statement than when probability is 1.0" [Dubois, 2006]. A normalization condition is further assumed so that at least one value u is viewed as totally possible $\pi_x(u) = 1$.

From a fuzzy set's perspective, the possibility distribution π can be viewed as determined by the membership function μ of a fuzzy set F. In this vision, $\pi_x(u) = \pi(x = u|F)$ estimates the possibility that the variable x is equal to u, knowing the incomplete state of knowledge "x is F". Then, $\mu(u)$ estimates the degree of compatibility of the precise information $x = u$ with the statement to evaluate "x is F" [Dubois et al., 2000].

5.4.2 A practical definition

In practice, a possibility distribution can be defined as follows. The simplest approach to represent imprecision is the interval, which is defined by a lower and an upper bound. But in most cases, experts may provide more information by expressing preferences inside this interval, i.e. the interval can be nuanced [Beer et al., 2013]. For example, "expert is certain that the value for the model parameter is located within the interval $[a;d]$". However, according to a few measurements and his/her own experience, expert may be able to judge that "the value for the model parameter is most likely to be within a narrower interval $[b;c]$". The most likely interval $[b;c]$ designates the "core" of the possibility distribution π and is assigned a degree of possibility equal to one, whereas the "certain" interval $[a;d]$ (referred to as the "support" of π) is assigned a degree of possibility zero, such that values located outside this interval are considered impossible. A transition between the core and the support can be assumed to be linear, but the expert could also assume a nonlinear curve [Baudrit et al., 2007b]. Furthermore, some procedures exist to directly derive a possibility distribution from a histogram of sparse observations [Masson and Denœux, 2006].

The possibility distribution can also be viewed from the probabilistic point of view. The intervals defined as $\pi_\alpha = \{e; \pi(e) \geq \alpha\}$ are called α-cuts. They contain all the values that have a degree of possibility of at least α (lying between 0 and 1) (see Fig. 5.6). They formally correspond to the confidence intervals 1-α as traditionally defined in the probability theory, i.e $P(e \in \pi_\alpha) \geq 1 - \alpha$. Figure 5.6A depicts a trapezoidal possibility distribution associated with the core $[b; c]$ and the support $[a; d]$. Furthermore, possibility distribution encodes a probability family [De Cooman and Aeyels, 1999, Dubois and Prade, 1992] limited by an upper probability bound called the possibility measure $\Pi(e \in E) = sup_{e \in E} \pi(e)$ (see for instance the upper cumulative probability bound on Fig. 5.6B) and a lower probability bound $N(e \in E) = inf_{e \notin E}(1 - \pi(e))$ called the necessity measure, where E represents a specific interval on the real line. The Possibility measure refers to the idea of plausibility, while the Necessity measure is related to the idea of certainty (see further details in [Dubois, 2006] and references therein).

5.4.3 Illustrative real-case application

Hereafter, the afore-described method is applied to represent the imprecision of the amplification factor A_{LITHO} associated to lithological site effects in the domain of seismic risk analysis. At local scale, site effects phenomena exist, which might amplify the ground-motion. Such effects are taken into account by a multiplicative scalar coefficient representing the amplification, which depends on the geo-technical and geological properties of the soil. The amplification is assumed to be spatially homogeneous in the so-called geo-technical zones (see illustration in the case of Lourdes, Fig. 5.4-right). The geo-technical zones are defined from the processing of different sources of data, namely from geological, geophysical (spectral analysis of surface waves and analysis of ambient vibrations) and geo-technical information.

In the Lourdes case, 13 so-called geo-technical zones were defined (Fig. 5.4-right) and each of which are associated with response spectra based on 1D site response analyses [Modaressi et al., 1997]. The amplification factor A_{LITHO} is derived from the comparison between the four response spectra (two from synthetic time-histories and two from natural accelerograms) computed for the zones located on bedrock outcrops and the numerically calculated spectra for the various zones. A sensitivity analysis was carried out for each bedrock spectrum by [Bernardie et al., 2006]. It results in the definition of the interval within which Bernardie and co-authors (i.e. the panel of experts, who are in charge of the site effect analysis), are certain to find the real value. The latter defines the support of the possibility distribution associated to A_{LITHO}. Besides, the mean value between all the possible outcomes given by each bedrock spectrum plus one standard deviation is considered by Bernardie and co-authors

the most likely value for A_{LITHO}. The latter gives the core of the possibility distribution, here assumed to be triangular.

5.5 Handling probabilistic laws with imprecise parameters

In this section, a method is described to deal with a specific source of uncertainty combining aleatory and imprecision, namely a probabilistic model with ill-known (imprecise) parameters. A common example is a Gaussian law with unknown mean and variance. A traditional method is to rely on 2D Monte-Carlo approach, which consists in: 1. assigning a probability law to the considered random variable and to the parameters of this probability law; 2. nesting two ordinary Monte-Carlo simulations (see e.g., [Helton, 1994]). However, as pointed out by [Baudrit et al., 2008], it would make more sense to use intervals or confidence intervals to represent the imprecision on these parameters, i.e. the non-stochastic nature of parameter distributions. An alternative, originally proposed by [Baudrit et al., 2008], is to present the imprecision on the probabilistic model's parameters by means of possibility distributions (Sect. 5.4). The applicability of such an approach is here investigated using a probabilistic damage assessment model used for seismic risk analysis, namely the RISK-UE, level 1 model (Sect. 5.5.1). The application of such an approach results in very rich pieces of information (Sect. 5.5.2), which may be tedious to manage for decision-making: a method to summarize it is proposed in Sect. 5.5.3 under the formalism of fuzzy random variables (see a brief introduction in Appendix E).

5.5.1 A motivating example: the Risk-UE (level 1) model

To model the physical damage of earthquakes to buildings, the EMS-98 damage grades [Grünthal, 1998] can be used. Six damage grades D_k (with k from 0 to 5) are considered: $D_0 = 0$ corresponds to "no damage", $D_1 = 1$ to "slight damage", $D_2 = 2$ to "moderate damage", $D_3 = 3$ to "heavy damage", $D_4 = 4$ to "very heavy damage" and $D_5 = 5$ to "maximal damage". The Risk-UE methodology as described by [Lagomarsino and Giovinazzi, 2006] proposes a probabilistic approach for damage assessment. The probabilistic damage curve is defined as the cumulative probability distribution F of the event "$d \leq D_k$".

$$F(d \leq D_k) = \int_0^{D_k} p(\theta) d\theta \tag{5.2}$$

where p is the density probability of the Beta law, such that:

$$p(d) = \frac{d^{q-1}(6-d)^{7-q}}{6^7} \frac{\Gamma(8)}{\Gamma(q)\Gamma(8-q)} \tag{5.3}$$

where Γ is the gamma function. The curve determined by Eq. 5.2 corresponds to the probabilistic decision curve to support risk management. The form of the Beta law is determined by q, which depends on the mean damage value r_D such that [Lagomarsino and Giovinazzi, 2006]:

$$\frac{q}{8} = 0.007 \cdot r_D^3 - 0.0525 \cdot r_D^2 + 0.2875 \cdot r_D \tag{5.4}$$

The mean damage value r_D is the key parameter of the earthquake risk model. The correlation is defined as follows [Lagomarsino and Giovinazzi, 2006]:

$$r_D = 2.5(1 + \tanh(\frac{I + 6.25 \cdot V_i - 13.1}{2.3})) \tag{5.5}$$

For a set of buildings in a given city district, r_D directly correlates the seismic ground motion parameter, namely the macroseismic intensity I and the vulnerability index of the set of buildings V_i.

5.5.2 Problem definition

In a seismic analysis with no epistemic uncertainty, the result of the hazard and of the vulnerability assessment are crisp values. In this case, the output is a unique decision curve. The large impact of epistemic uncertainty has been underlined at each stage of the earthquake risk assessment (hazard, vulnerability assessment, etc.) in various studies, in particular [Steimen, 2004, Crowley et al., 2005]. In this context, the choice of the appropriate decision curve is uncertain.

If the epistemic uncertainties on the model parameters are simply represented by intervals, the risk model output would be a set of damage curves, which would be determined by its pair of lower and upper curves. If more information is available, the model parameters can be represented by a possibility distribution, i.e. by a set of intervals associated with a degree

of possibility corresponding to the α-cuts (see Sect. 5.4). For each degree α (ranging from 0 to 1), a pair of damage lower and upper bounds can be constructed, thus defining a family of damage curves associated with each degree α. Figure 5.7 illustrates the methodology to be used for the construction of the family of probabilistic damage curves on the basis of the α-cuts of the imprecise mean damage value r_D depicted in Fig. 5.7-left.

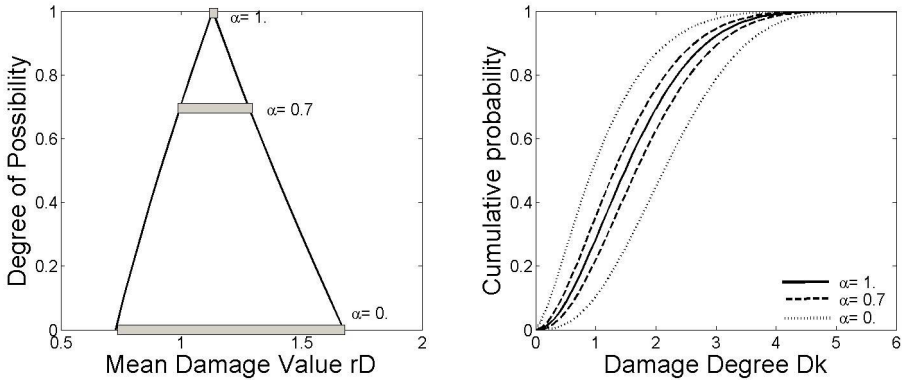

Figure 5.7: Illustration of the family of probabilistic damage curves (right figure) associated to the α-cuts of the imprecise parameter r_D (left figure)

The output of the propagation procedure gives the decision maker all the possible alternatives for the probabilistic decision curves. This set of decision curves should be summarized for an efficient use in risk management. The objective is to provide a simple measure of the whole epistemic uncertainty, while preserving a probabilistic format, which is familiar to the decision-makers. In this view, we use methodological tools for uncertainty processing developed by [Baudrit, 2005, Baudrit and Dubois, 2006, Baudrit et al., 2007b] in the formal framework of fuzzy random variables (a brief introduction to fuzzy random variable is provided in Appendix E; a more extensive description is provided by [Gil, 2001]).

5.5.3 Use for an informed decision

Under a situation of "partial ignorance", the damage grade D_k (k=0 to 5) can be seen as a fuzzy random variable. The results depicted in Fig. 5.7-right can be viewed as the second order possibility distribution induced by its α-cuts $[\underline{F}_\alpha; \overline{F}_\alpha]$. The strong relationship with the order-one model of fuzzy random variables (see Appendix E) allows summarizing the uncertainty on the damage grade in a pair of indicators $[\underline{P}; \overline{P}]$ associated to the event: "$d \leq D_k$; for $D_0 = 0, ..., D_5 = 5$", which bound the "true" probability. Fig. 5.8 gives the output of the synthesis methodology for the family of probabilistic damages of curves described in Fig. 5.7. The gap between the two indicators exactly reflects the incomplete nature of our knowledge, thus explicitly displaying "what is unknown", whereas the value of the probability bounds is

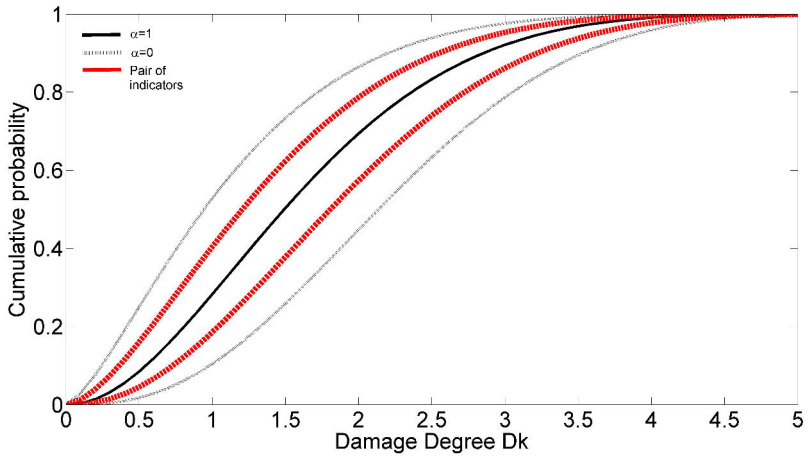

Figure 5.8: Synthesis in a pair of probabilistic indicators of all the possible alternatives for the probabilistic damage curves associated with the α-cuts of the imprecise parameter r_D (red lines). The pairs of probabilistic curves for α=0 and α=1 are also indicated.

related to the aleatoric uncertainty.

Therefore, this indicator can be used to underline the zones in the studied area, where the epistemic uncertainty is the highest i.e. where efforts should be made in terms of additional campaigns (e.g. vulnerability assessment).

Figure 5.9 gives the final output of the uncertainty propagation step considering the earthquake scenario of the city of Lourdes (South of France). Different types of epistemic uncertainties were dealt with: 1. imprecision on the amplification factor expressing lithological site effects (Sect. 5.4) and on the vulnerability measurement; 2. the vague information associated with the exposed building inventory (Sect. 5.3). The final result consists of a pair of maps, which respectably represent the lower and upper probability bounds of the event: "$d \geq D_4$". Prioritisation of the districts for further characterisation studies could then be based on the two indicators as shown by the following examples.

- The probability of exceeding the grade of damage D_4 is between 1.7 % and 11 % in the district n°10. When considering the district n°81, the lower indicator is 2.63 % and the upper one is 13.5 %. The ranking is straightforward, as both indicators show that the probability of the event "$d \geq D_4$" is the higher in the district n°81;
- When considering the lower and upper indicators in the district n°40 of respectably 0.9 % and 18 %, no conclusion can be drawn from the comparison of the indicators for districts n°81 and n°10, because the level of epistemic uncertainty is very large (around 17.9 %) in this district, whereas it reaches around 10 % in the other districts. Thus, such an analysis points out that additional investigations should be undertaken in the district

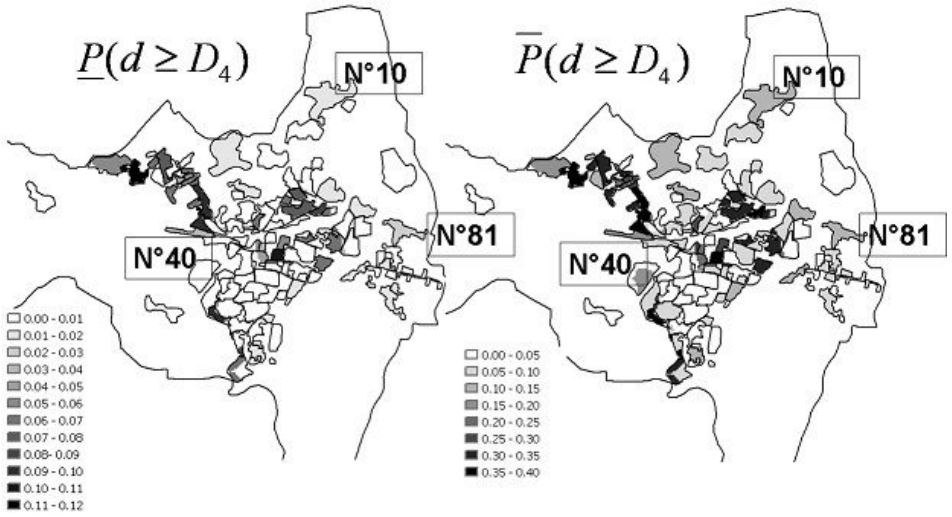

Figure 5.9: Mapping of lower (left) and upper (right) probabilistic indicator of the event: "exceeding damage grade D4", locations of district n°10, n°81 and n°40 are indicated (see Sect. 5.5.3 for details).

n°40.

5.6 Concluding remarks of Chapter 5

In the present chapter, the different criticisms available in the literature against the systematic use of probabilities in situations where the available data/information are scarce, incomplete, imprecise or vague were reviewed. On this basis, the use of a flexible uncertainty representation tool was investigated, namely Fuzzy Sets. Different situations were considered and examples of real cases in the context of geo-hazard assessments were used:

- Vagueness due to the gradual nature of words. The application of Fuzzy sets is illustrated in the context of susceptibility assessment of abandoned underground structures [Nachbaur and Rohmer, 2011]. In particular, it is shown how the so-called "threshold effect" can be handled when the expert defines classes of hazard / susceptibility;
- Reasoning with vague concepts. This is handled using Fuzzy Logic. This is illustrated with the treatment of imprecision associated to the inventory of assets at risk in the context of seismic risk analysis;
- Imprecision. This is handled by defining possibility distributions, which have a strong link with fuzzy sets. This is illustrated with the representation of uncertainty on the amplification factor of lithological site effects in the context of seismic risk analysis;
- Imprecision on the parameters of a probabilistic model. This is handled in the setting of fuzzy random variables. This is illustrated using a probabilistic damage assessment

model in the context of seismic risk analysis, namely the RISK-UE, level 1 model.

On this basis, the next chapter investigates how this “new tool” for uncertainty representation can be integrated in a sensitivity analysis procedure together with traditional probabilistic information.

6 Sensitivity analysis adapted to a mixture of epistemic and aleatory uncertainty

In the previous chapter, fuzzy set was described as a tool for processing data and handling different situations of epistemic uncertainty: its flexibility was demonstrated through real-case examples. In the present chapter, the question of how to conduct sensitivity analysis when the uncertain parameters are represented by different mathematical tools (i.e. a mixture of aleatory-epistemic uncertainty representations) is addressed. First, a state of the art of the approaches and procedures is provided (Sect. 6.1). On this basis, a major limitation is outlined, namely the computational cost of the existing procedures. In Sect. 6.2, a graphical tool is developed to alleviate this burden by using only the simulations required for uncertainty propagation in a post- processing manner. Among all the situations of sensitivity analysis, the study is focused here on a common situation for natural hazards, namely assessing the stability of a system. The applicability of the tool is demonstrated in a third section using real-case applications (Sect. 6.3).

6.1 State of the art of sensitivity analysis accounting for hybrid uncertainty representations

To the author's best knowledge, only a few procedures have been proposed to account for mixed aleatory-epistemic uncertainty representations in sensitivity analysis. The most popular technique was proposed by [Ferson and Troy Tucker, 2006], namely the "pinching" strategy. This consists of assessing how the imprecision on the output would reduce if additional information on the input parameters was available, i.e. if imprecision in the inputs was reduced. For instance, this can be performed by transforming an triangular possibility distribution to a constant value, e.g. to the core, or by transforming a Gaussian-type probability-box to a single Gaussian probability distribution (see Appendix D for an introduction to this un-

certainty representation approach). This strategy was further improved by [Alvarez, 2009] using the Hartley-like measure of non-specificity. A series of case studies was provided by [Oberguggenberger et al., 2009], who both used the classical (probabilistic) sensitivity analysis and the pinching methods applied using different concepts including random sets, fuzzy sets and pure interval bounding. They evaluated each approach regarding performance criteria (cost, accuracy, interpretability, ability to incorporate dependencies, and applicability to large scale problems), which showed the attractive features of imprecise probabilities. Following a strategy with a similar spirit than pinching (i.e. "reduce epistemic uncertainty in the input and estimate the effect on the output"), [Guo and Du, 2007] proposed an OAT-like analysis consisting in keeping one epistemic variable uncertain while the other ones are fixed at their average values and in measuring the effect through the difference between the upper and the lower probability distribution (i.e., belief and plausibility measures, see further details in Appendix D) using the Kolmogorov-Smirnov distance.

Let us illustrate the pinching procedure using the example of the probabilistic damage model for seismic risk assessment (see description in Sect. 5.5). Consider the fuzzy set depicted in black in Fig. 6.1 representing the imprecision in the vulnerability index V_i assigned to the class of vulnerability termed as M4 "masonry stone" as defined by [Lagomarsino and Giovinazzi, 2006]. [Giovinazzi and Lagomarsino, 2004] defined a pinching procedure using a filter function f as follows:

$$\begin{aligned} &f(V_i, V_i^*, V_f) = 1 && if && |V_i - V_i^* - V_f| \leq V_f/2 \\ &f(V_i, V_i^*, V_f) = 0 && if && |V_i - V_i^* - V_f| > 3V_f/2 \\ &f(V_i, V_i^*, V_f) = 1.5 - |V_i - V_i^* - V_f|/V_f && otherwise \end{aligned} \tag{6.1}$$

where V_f is the filter band-width, and V_i^* is the "most probable" value defined at the centroid of the fuzzy set. The smaller the band-width, the more precise the fuzzy set (i.e. centered around V_i^*). The filter function is then multiplied to the fuzzy set and normalized between 0 and 1.

Figure 6.1 respectively depicts the transformation of the M4 fuzzy set given different values of the filter bandwidth V_f varying from 0.1 to 0.3. The upper and lower bounds of the probability of exceeding the damage degree D_4 is calculated considering a constant macroseismic intensity of VII (as described in Sect. 5.1). Figure 6.2 provides the width of the probability interval (assimilated to an epistemic indicator) as a function of V_f. This type of curve can be useful to decide the level of imprecision reduction in the estimate of the vulnerability index of M4, which should be achieved so that a "reasonable" degree of epistemic uncertainty in the

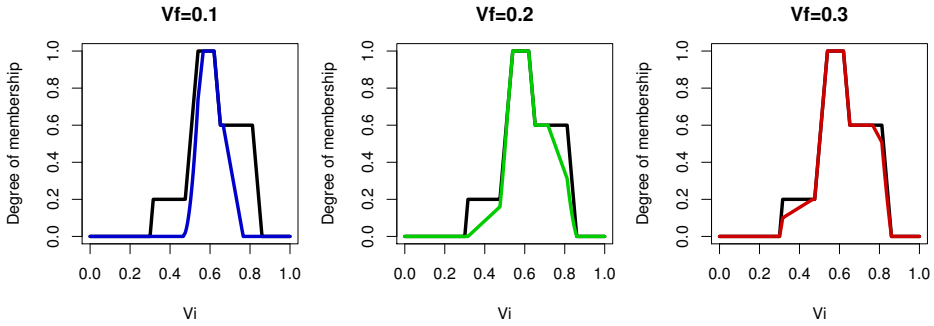

Figure 6.1: Example of pinching transformation of the fuzzy set assigned to the vulnerability index Vi of the M4 class of vulnerability (in black) considering a bandwidth of the filter: A) V_f=0.1; B) V_f=0.2; C) V_f=0.3

probability estimate can be reached.

For the three cases considered in Fig. 6.1, the epistemic indicator decreases from ≈ 2% to less than 0.5%. [Giovinazzi and Lagomarsino, 2004] proposed to link the filter width to the quantity, quality and origin of the data available from additional vulnerability surveys: V_f=0.08 for non specified existing data base; V_f =0.04 for data specifically surveyed for vulnerability purposes. As shown by Fig. 6.2, the addition of any data, whatever their origin, can lead to a large reduction in the epistemic indicator.

Though the implementation of the pinching-like procedure remains efficient in practice, it presents the drawback of being strongly dependent on the type of the pinching method used (as pointed out by [Ferson and Troy Tucker, 2006] and still recently by [Alvarez, 2009]). Take the example of the pinching of an interval. The most common case would consist in transforming it to a constant value. But what is the “most reasonable” constant value: the interval median? The lower or the upper bound? Besides, assuming that the whole imprecision can be reduced by additional data can sound "too optimistic". In practice, the addition of new data usually reduces the magnitude of epistemic uncertainty to a given level. As discussed by [Allaire and Willcox, 2012] in the probabilistic setting, imprecision reduction is rarely achievable to the true value of any uncertain parameter . A “less extreme” pinching option would consist in transforming it to an interval of smaller width or to a triangular possibility distribution, whose support is embedded within the interval bounds. The filter-based approach, as afore-described, appears efficient, but it should be recognized that the choice of V_f can be tedious, if not too arbitrary/subjective. To overcome the arbitrariness of choosing a specific pinching strategy, a systematic exploration of all the possible pinching methods through Monte-Carlo simulations could be developed as originally suggested by [Alvarez, 2009]. Yet, this might lead to an increase of the computation time cost.

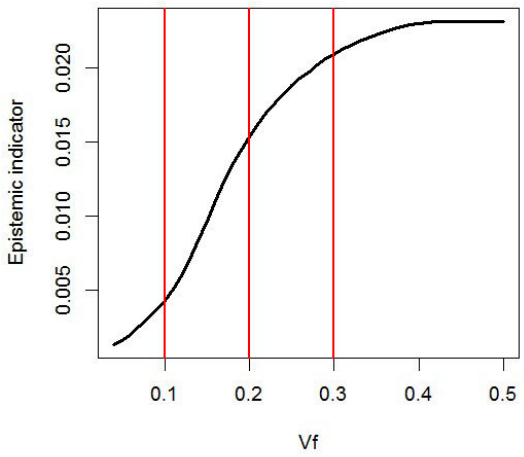

Figure 6.2: Epistemic indicator (width of the interval bounding the probability of exceeding the damage degree D4) as function of the filter width V_f applied to the fuzzy set of M4 vulnerability class. The three cases represented in Fig. 6.1 are outlined by a red vertical line.

Alternatives to pinching-like techniques have been proposed. [Hall, 2006] examined three approaches for deriving an uncertainty-based sensitivity either by extending variance-based-sensitivity analysis, or partial expected value of perfect information or relative entropy to imprecise probabilities. This yields interval-valued sensitivity indices. A first limitation of the approach is the need to search a sufficiently general bounded set of probability distributions that includes the distributions that maximize sensitivity: [Hall, 2006] proposed to use a set of beta-type probabilistic distributions. The second limitation highlighted by [Hall, 2006] is the computational expense of calculating imprecise sensitivity indices. Finally, [Hall, 2006] proposed a unique scalar-valued sensitivity indicator, i.e. aggregate entropy-based sensitivity measure based on entropy maximization as proposed by [Klir and Wierman, 1999]. This is both applicable to precise and imprecise probability distributions and is less computationally intensive than the approach based on imprecise sensitivity indices. Yet, [Hall, 2006] highlighted the limited sensitivity of the approach to changes of the definition of the input information, which is a feature of high interest in practice.

In conjunction with evidence theory (see description in Appendix D), [Helton et al., 2006a] proposed a three-step sampling-based sensitivity analysis:

- Initial exploratory analysis: the relationship between the variable of interest y and the input parameters can be explored through classical sampling-based approaches and examination of the scatterplots. This first screening step is used to identity the most

influential input parameters;

- Stepwise analysis: based on the ordering of the most important parameters in the first step, the lower and upper cumulative probability distribution, i.e. Plausibility Pl and Belief Bel (see further details in Appendix D), of y is constructed in turn by first only accounted for the Dempster-Shafer structure of the first most important parameter, then of the two first most important parameters, and so forth. In this manner, the incremental effect of the input parameters on Pl and Bel can be estimated;
- Summary analysis: compute sensitivity measures. Instead of computing upper and lower bounds of such sensitivity measures as proposed by [Hall, 2006], the following strategy is proposed: i. select precise probability distributions consistent with the Dempster-Shafer structure assigned to each input parameter; ii. numerically sample the selected probability distributions and compute the corresponding sensitivity measure (e.g. variance-based measure); iii. Repeat the procedure for other probability distributions. This results in a spectrum of sensitivity indices related to the spectrum of probability distributions used in the analysis.

Finally, it is worth noting the recent study by [Song et al., 2014], who proposed to use the original version of Borgonovo's measure of sensitivity and its extension to fuzzy sets to both handle random variables or fuzzy-valued variables for structual reliability assessment.

6.2 A graphical-based approach

6.2.1 Motivation

One major drawback of the afore-described procedures (except for the pinching OAT-like procedure) is their computational cost, especially when the model has a high computation time duration. To overcome such a limitation, robust meta-modelling strategies have been proposed, see e.g. [Eldred et al., 2011], via the combination of meta-modelling techniques (as the one described in Chapter 3) and optimization-based interval estimation. With good reasons, such types of sensitivity techniques focus on epistemic uncertainty, since it is this type of uncertainty, which is expected to be reduced through additional characterisation data. Yet, it would also be valuable to highlight the contribution of aleatoric uncertainty versus epistemic uncertainty as the actions in terms of risk management differ (see Sect. 1.3 and [Dubois, 2010]): a joint exploration of both types of uncertainty would be of interest from a risk management perspective.

In the following, another approach is proposed following the statement of [Helton et al., 2006a]: "An initial exploratory sensitivity analysis plays an important role in helping to guide any study

that involves uncertain inputs [...] The examination of scatterplots is a natural initial procedure". This motivated the development of an easy-to-use graphical exploration tool for sensitivity analysis with different types of uncertainty representations. Randomness is represented by means of probability distributions, whereas imprecision is represented by means of numerical possibility distributions (see Chapter 5). Contrary to previous chapters, the sensitivity analysis is focused on the estimate of failure probability instead of the variance: this is of high interest for stability analysis, often used for hazard assessments. This type of analysis usually relies on the use of a safety factor SF, which is defined as the ratio of the resisting to the driving forces acting on the mechanical system. If SF is below 1.0 (or some specified threshold), the failure is possible. The objective of the stability analysis is then to estimate the failure probability P for SF to be below the specified threshold. The proposed approach relies on the adaptation of a tool recently proposed by [Li and Lu, 2013] in the probabilistic setting, namely the Contribution to probability of Failure sample Plot CFP.

First the joint propagation of randomness and imprecision is recalled (Sect. 6.2.2). Then, the construction of CFP in the pure probabilistic setting is introduced (Sect. 6.2.3). On this basis, CFP is adapted to the case of mixture of both types of uncertainty (Sect. 6.2.4).

6.2.2 Joint propagation of randomness and imprecision

Consider that SF is related to n uncertain inputs X_i (whether random or imprecise) by a model G such that $SF = G(X_i)$ with $i = 1,...,n$. In situations where the only type of uncertainty is randomness, the most commonly used propagation approach relies on Monte-Carlo random sampling. In situations where both randomness and imprecision are present, the uncertainty propagation task should rely on more advanced procedures. Monte-Carlo-based procedures have recently been developed in order to both generate random numbers from cumulative probability distributions and intervals either from probability boxes [Zhang et al., 2010] or from possibility distributions [Baudrit et al., 2007b]. Here, we focus on the latter approach referred to as the independence random set method IRS [Baudrit, 2005].

This joint propagation method is based on the assumption of independence among all sources of uncertainty, whether of aleatoric or of epistemic nature. Figure 6.3 schematically depicts the main steps of the propagation procedure considering a random and an imprecise variable.

Consider k random input variables $X_i, (i = 1,...,k)$, each of them associated with a cumulative probability distribution F, and $n-k$ imprecise input variables X_i (i=k+1,...,n), each of them associated with a possibility distribution π. In this situation, the IRS procedure holds as follows ([Baudrit, 2005]: page 76):

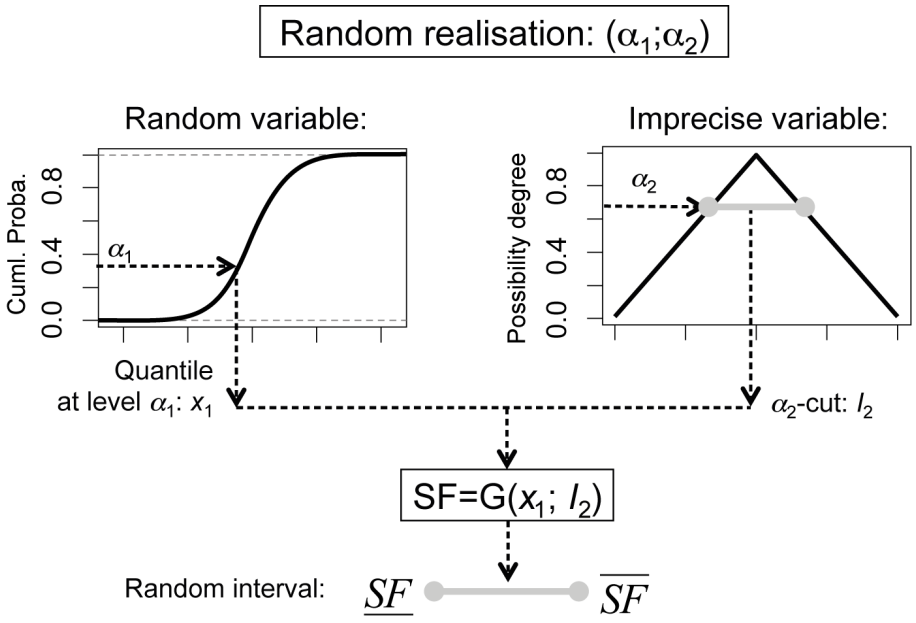

Figure 6.3: Main steps of the joint propagation of a random variable (left, represented by a cumulative probability distribution) and an imprecise variable (right, represented by a possibility distribution) using the independence Random Set procedure of [Baudrit et al., 2007b].

As a preliminary step, randomly generate from uniform probability distributions, m vectors of size n: α_i, i=1,...,n, such that $0 \leq \alpha_i \leq 1$. For each realisation:

- Step 1: generate k values for the k random input variables X_i by using the inverse function of F_i: , $i = 1,...,k$;
- Step 2: sample $n-k$ intervals I_i corresponding to the cuts of the possibility distributions with level of confidence $1-\alpha_i$, $i = k+1,\ldots,n$;
- Step 3: evaluate the interval $[\underline{SF};\overline{SF}]$ defined by the lower and upper bounds associated with SF as follows:

$$\begin{aligned} \underline{SF} &= inf_I(G(x_1,...,x_k,I_{k+1},...,I_n)) \\ \overline{SF} &= sup_I(G(x_1,...,x_k,I_{k+1},...,I_n)) \end{aligned} \tag{6.2}$$

Note that step 3 of the method should rely on global optimization techniques if the function G is not monotone, which can be very computationally intensive.

The output of the whole procedure then takes the form of m random intervals of the form $[\underline{SF}_k;\overline{SF}_k]$, with $k = 1,...,m$. This information can be summarised in the form of a pair of indicators, namely the Plausibility (Pl) and Belief (Bel) functions bounding some probability

of being below a given threshold T, namely $P(A)$ with the event "$A = SF \leq T$" in the formal framework of the evidence theory [Shafer et al., 1976] as proposed by [Baudrit et al., 2007b, Baudrit et al., 2007a]. Considering [Baudrit, 2005], the indicators bounding P can be calculated as follows:

$$\begin{array}{lcl} Bel(A) & = & \frac{1}{m} Card(k, \overline{SF}_k \leq T) \\ Pl(A) & = & \frac{1}{m} Card(k, \underline{SF}_k \leq T) \end{array} \tag{6.3}$$

where Card denotes the cardinality operator of the considered set.

An example of post-processing is provided in Section 6.3.2 on Fig. 6.6: the "real" failure probability is bounded by the optimistic Bel (i.e. the failure probability is the lowest) and the pessimistic bound Pl. Yet, this type of format for presenting the final result can be difficult to use in practice. Our own return of experiences with the French Environment Ministry showed us that decision-makers are more comfortable in using a single value for the failure probability. In particular, [Oberguggenberger and Fellin, 2002] outlined the usefulness of probabilities as means for decision-making under uncertainty (i.e. at the final phase of the risk analysis chain) for comparative studies of scenarios. Thus, a solution may consist in restricting the decision using the pessimistic bound. However, as pointed out by [Dubois and Guyonnet, 2011], in some situations for risk analysis, restricting the analysis on Pl may be too conservative, because this neglects information leading to less pessimistic outcomes. In order to support decision-making with a more nuanced indicator, [Dubois and Guyonnet, 2011] proposed to weight the bounds by an index w, which reflects the attitude of the decision-maker to risk (i.e. the degree of risk aversion) such that the randomly generated intervals are replaced by a unique value: $w \cdot \underline{SF}_k + (1-w) \cdot \overline{SF}_k$. If w equals 1.0, more weight is given to the pessimistic bound and the aversion to risk of the decision-maker is the highest. This results in a single probability value P_w, which both reflects the imprecise character of P and the attitude (aversion) of the decision-maker to risk. In this sense, this can only be chosen by the decision-maker himself. Though this sounds "artificial", this has minor consequences compared to situations, where the systematic use of probabilities is not questioned, i.e. from the very beginning of the uncertainty treatment chain when choosing the mathematical model representing the uncertainty as discussed by [Dubois and Guyonnet, 2011]. In the following, the sensitivity analysis is focused on this unique indicator P_w.

6.2.3 Contribution to probability of failure sample plot

Graphical tools have been proposed by several authors to provide valuable information on the relationship between the model inputs and the output (see a review by [Tarantola et al., 2012]). Among the existing methods, plots based on the contribution of the input to a given statistical characteristic of the output have shown to be very efficient to investigate the regional importance of inputs. This allows localizing the regions of the input's domain of variation, which contribute the most to the evolution of the considered statistical characteristic, like the mean [Bolado-Lavin et al., 2009], the variance [Tarantola et al., 2012] or the probability of being below a specified threshold, i.e. the probability of failure [Li and Lu, 2013]. The latter graphical tool, denoted in the following CFP, is the main ingredient of the proposed approach.

This plot relates any given fraction q of the smallest values of the considered uncertain variable $x_i = 1,..,n$ (i.e. the quantile of x_i) to the fraction of the probability $P(SF \leq T)$. In the pure probabilistic setting, the CFP curve is then constructed as follows:

- Step 1: randomly generate m samples of vectors of size n: $\boldsymbol{x}^{(j)} = (x_i, i = 1,...,n)$, with j=1,...,m; then, estimate the corresponding safety factor $SF^{(j)}$ (j=1,...,m) by running the model G;
- Step 2: compute the indicator function H associated with each realisation $SF^{(j)}$ (j=1,..,m) such that $H(SF^{(j)}) = 1.0$ if $SF^{(j)} \leq T$, and 0.0 otherwise;
- Step 3: considering each input variable x_i, sort in ascending order the samples $(x_i^{(1)};...;x_i^{(m)})$ and the corresponding set of indicator values $(H^{(i,1)};...;H^{(i,m)})$, with $i = 1,\ldots,n$;
- Step 4: at each quantile q, the contribution CFP to probability of failure sample plot is computed for each input variable x_i as follows:

$$CFP_{x_i}(q) = \frac{\sum_{j=1}^{|q\times m|} H^{(i,j)}}{\sum_{j=1}^{m} H^{(i,j)}} \tag{6.4}$$

where $|q \times m|$ is the largest integer not greater than $q \times m$, $q \in [0;1]$

See an example of CFP in Fig. 6.4. By investigating the position of the CFP curve relatively to the diagonal (i.e. the first bisector), several specific input-output relationships can be outlined. If the CFP curve is close to the diagonal, it indicates that the corresponding variable contributes to the same manner to P throughout its range of variation such that it can be fixed at any nominal value with little influence on P. On the other hand, where the CFP curve is very steep, the contribution to P is large (larger than on average). Thus, the location of the

Table 6.1: Description of the main steps and methods of the joint exploration procedure of possibilistic and probabilistic uncertainty.

Step	Description	Methods
1	Depending on the data/information available, select the most appropriate mathematical framework (either probabilistic or possibilistic) for representing the uncertainty on the parameters of the stability analysis;	Probability distribution and possibility distribution (Sect. 5.4)
2	Propagate uncertainty to estimate the bounds of failure probability P;	Monte-Carlo sampling and IRS procedure (Sect. 6.2.2)
3	Select a value for degree of risk aversion w and estimate P_w;	This reflects the attitude of the decision-maker regarding risk (Section 6.2.2)
4	Apply the procedure for computing the CFP by sorting in ascending order the levels of confidence, which have served to sample the random parameters and the ones, which have served to sample the cuts associated with the imprecise parameters.	Contribution to probability of failure sample plot in the probabilistic framework described in Sect. 6.2.3 and adaptation to the possibilistic framework described in Sect. 6.2.4.

maximum departure from the diagonal indicates the region of the parameter's domain where it has the greatest sensitivity [Bolado-Lavin et al., 2009]. If the CFP curve does not cross the diagonal, it indicates a monotonic input-output relationship (increasing if below the diagonal and decreasing if above it). Furthermore, [Saint-Geours et al., 2015] clarified the link between first-order sensitivity indices (main effects) and Contribution to the Sample Mean plots.

6.2.4 Adaptation to possibilistic information

Hereafter, CFP is adapted to both handle probabilistic and possibilistic information to investigate the influence on the unique indicator P_w, which both reflects the imprecision in the probability for SF to be inferior to 1.0 and the risk aversion of the decision-maker.

First, recall the Monte-Carlo sampling strategy as afore-described. Considering only probabilistic information, let us notice that constructing the CFP curve by directly using the samples of X_i or of the quantiles q_i (i.e. levels of confidence) is equivalent due to the positive monotonic behaviour of the cumulative probability distribution: $X_i = F^{-1}(q_i)$.

Considering the possibilistic information, the random sampling is conducted on the degrees of possibility α_i, which is have a link with the levels of confidence: $q_i = 1 - \alpha_i$. For each

imprecise variable, the result is a set of intervals (cuts), which are made nested by ordering the levels of confidence such that $q_i = 0$ corresponds to the core (the smallest interval) and $q_i = 1$ corresponds to the support (the largest interval). An example of possibility distribution seen as a set of nested cuts is provided in Fig. 6.3 (see also Fig. 5.6). Hence sorting the set of q_i in an ascending order is equivalent to sorting the corresponding intervals with increasing width, i.e. with increasing imprecision. Thus, we propose to construct the CFP curve by using the samples of levels of confidence q_i associated with each cut and/or random number. Regarding the possibilistic information, the analysis of the CFP curve then allows identifying the ranges of q_i for which the cuts most greatly influence the probability P_w. For instance, it can be identified whether the intervals close to the core (i.e. close to the most likely value) or close to the support (i.e. related to the spread) contribute the most. The summary of the whole procedure is provided in Table 6.1.

Note that a second option would have consisted in constructing the CFP curve by using the lower and upper bounds of the sampled intervals. Yet, this would have implied handling twice more inputs, which have rapidly become intractable when considering a high number of inputs.

6.3 Case studies

In this section, it is shown how the proposed graphical tool can be useful to explore the influence of both randomness and imprecision using several test cases. First, a simple model is used (Sect. 6.3.1). Then, application is conducted for stability analysis in real cases: in the domain of steep slope stability (Sect. 6.3.2) and in the domain of mine pillar stability (Sect. 6.3.3). Finally, a more complex case is investigated using a numerical model for pillar stability analysis (Sect. 6.3.4): the feasibility of the tool is further explored by combining the approach with meta-modelling techniques as described in Chapter 3.

6.3.1 Simple example

To better illustrate the types of results, which can be derived, let us first consider a simple model: $SF(X;\epsilon) = X^2 + r.\epsilon + 0.75$ where X is an imprecise variable represented by a triangular possibility distribution with core 0.0 and support $[-1.0; 1.0]$; ϵ is a random variable following the uniform distribution between 0 and 1; r is a scalar value, which dictates the "strength" of the randomness in the model. Here, the analysis is conducted with $P_{w=0} = Bel(SF \leq 1.0)$ i.e. we use the most optimistic probability bound such that the degree of risk aversion equals 0 ($w = 0$). By construction, the intervals with levels of confidence ranging from 0 (core) to 0.5

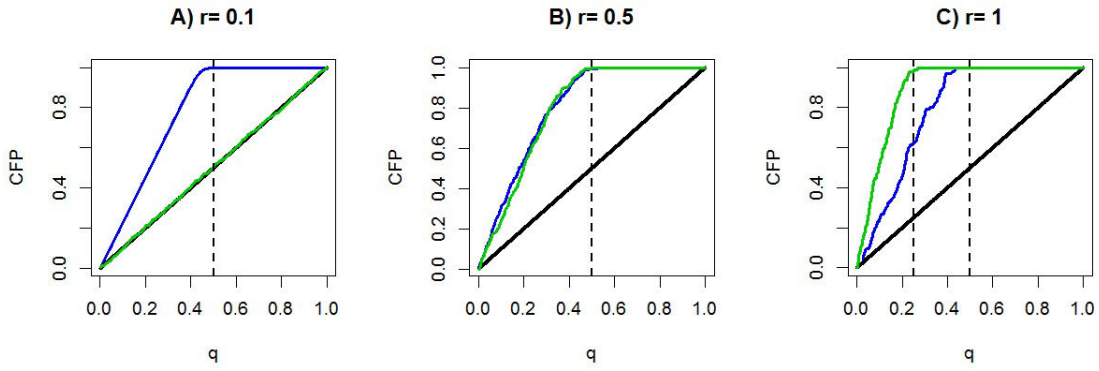

Figure 6.4: CFP curves associated with the random variable ϵ (green) and the imprecise variable X (blue) considering the simple model described in Sect. 6.3.1 and different cases of "randomness strength r".

contribute to P_w. Consider three cases: A) $r = 0.1$; B) $r = 0.5$; C) $r = 1.0$. In Fig. 6.4, the CFP curves are depicted for both variables considering each of these cases.

Considering case A), the strength of randomness r is very low such that the CFP curve associated with ϵ is close to the diagonal. As expected the CFP curve associated with X increases from 0 to 0.5 and then reaches a horizontal plateau at 1.0: above the level of confidence of 0.5 (i.e. close to the support), the randomly generated intervals have their lower and upper bounds both exceeding the threshold at 1.0, hence the probability P_w is not influenced by them. Considering case B), both CFP curves present the same trend, i.e. they increase up to 0.5. Here, the contribution of randomness and imprecision to P_w is the same. Considering case C), r is high such that the contribution of ϵ is the highest for levels of confidence above 0.25. This is indicated by a CFP curve increasing from 0 to 0.25.

6.3.2 Case study n°1: stability analysis of steep slopes

The first application case is based on the stability analysis described by [Collins and Sitar, 2010] for steep slopes in cemented sands. The safety factor SF is estimated using infinite slope assumptions accounting for the toe height H_t and the possible presence of a tension crack of height H_{tc} in the upper part of the cliff (of total height H and of slope inclination β, Fig. 6.5).

The expression for SF reads as follows:

$$SF = \frac{2c(\frac{H_S+H_t}{H^2-(H_S+H_{tc})^2}) + \gamma(\cos\beta)^2\tan(\phi)}{\gamma\cos(\beta)\cos(\beta)} \tag{6.5}$$

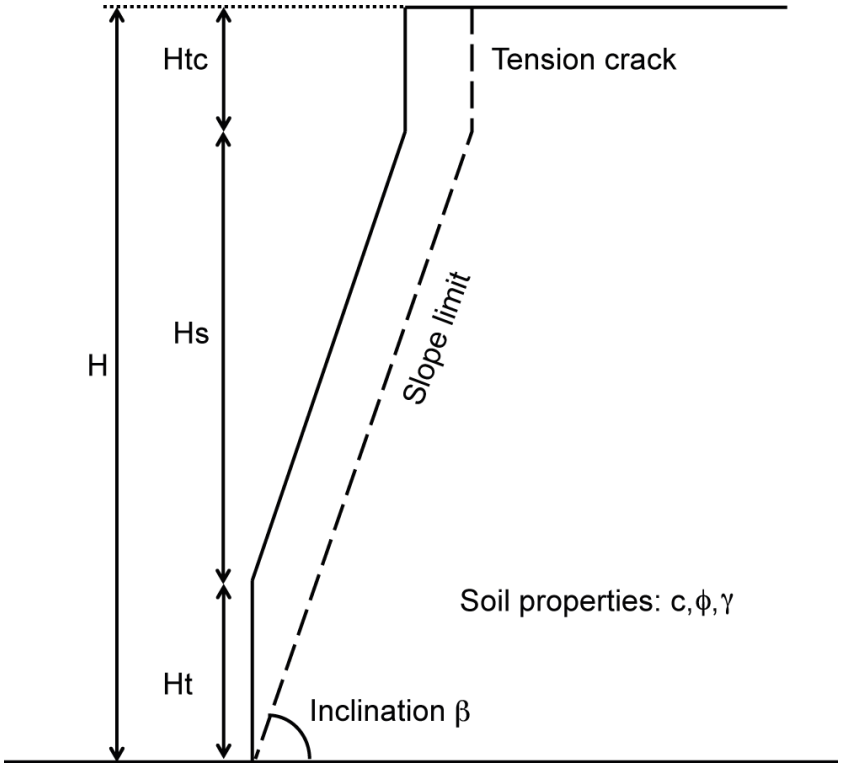

Figure 6.5: Schematic description of the failure geometry applied in [Collins and Sitar, 2010] to assess stability of steep slopes in cemented sands.

where c is the soil cohesion, ϕ is the soil internal friction angle, γ is the unit weight, and H_s is the slope height defined as $H-(H_{tc}+H_t)$.

Based on the case study in northern California described in [Collins and Sitar, 2010], the slope total height and inclination are considered at their constant values: H =27 m and $\beta = 57°$. The imprecision is associated with two parameters: the tension crack's height H_{tc}, and the toe height H_t. These imprecise parameters are respectively modelled by a trapezoidal possibilistic distribution with core $[2;3m]$ and support $[1;5m]$, and by a triangular possibilistic distribution with core $2m$ and support $[1;3m]$. Note that the imprecision on H_{tc} is the largest, because its estimate is made difficult on site due to the practical difficulties to access to this part of the cliff. The soil properties c, ϕ are assumed to be random and modelled by a Gaussian probability distribution with mean value of respectively 8 kPa (1 kPa = 10^3 Pa) and of 39°, and standard deviation defined as 10 % of the mean value. The uncertainty in γ is assumed to be negligible and fixed at its constant value of 17 kN/m^3.

Based on the afore-described assumptions, variability and imprecision are jointly propagated in the model using the IRS propagation method (Sect. 6.2.2) with 10,000 Monte-Carlo samples. The result of the propagation is post-processed using the procedure described in Sect. 6.2.2.

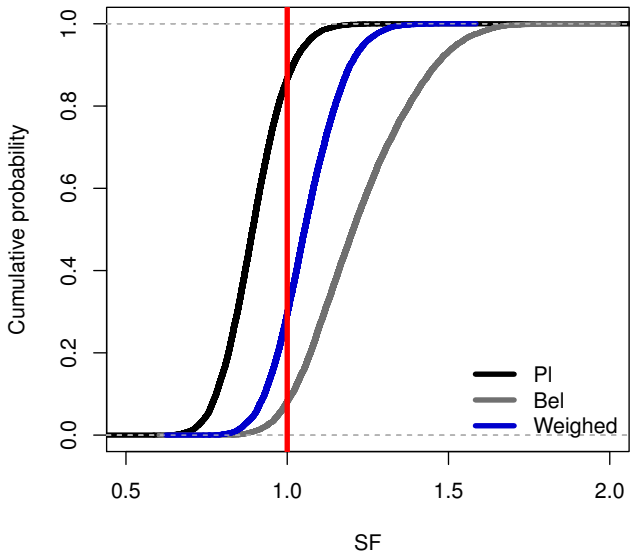

Figure 6.6: Plausibility Pl and Belief Bel functions resulting from the joint propagation of variability and imprecision in the slope stability analysis. The vertical line represents the threshold used to support decision-making. The blue line (denoted "Weighed") represents the distribution resulting from the weighting procedure of Pl and Bel by a degree of risk aversion $w = 0.5$ using the approach by [Dubois and Guyonnet, 2011].

The result is depicted on Fig. 6.6.

Focusing on the probability that SF remains below 1.0, we conduct the procedure of Table 6.1. Figure 6.7 depicts the CFP curves associated with P_w assuming different degrees of risk aversion: $w = 0.25$ (larger weight given to the most optimistic probability bound); $w = 0.50$; $w = 0.75$ (larger weight given to the most pessimistic probability bound). Several observations can be made:

- the CFP curve associated with the random parameter ϕ presents the largest deviation from the diagonal whatever the degree of risk aversion. The quantiles below $\approx 60\%$ ($w = 0.25$); below $\approx 70\%$ ($w = 0.50$) and below $\approx 80\%$ ($w = 0.75$) contribute the most to P_w (as indicated by the steep increase of the CFP curve);
- the imprecise parameter H_{tc}, though presenting a large imprecision (large width of the core and of the support), appears to have a minor influence whatever the degree of aversion w since its CFP curve remains very close to the diagonal;
- the imprecise parameter H_t, though presenting a relatively small imprecision (a spread of 1 m), appears to play a significant role. In particular, when a large weight is given to

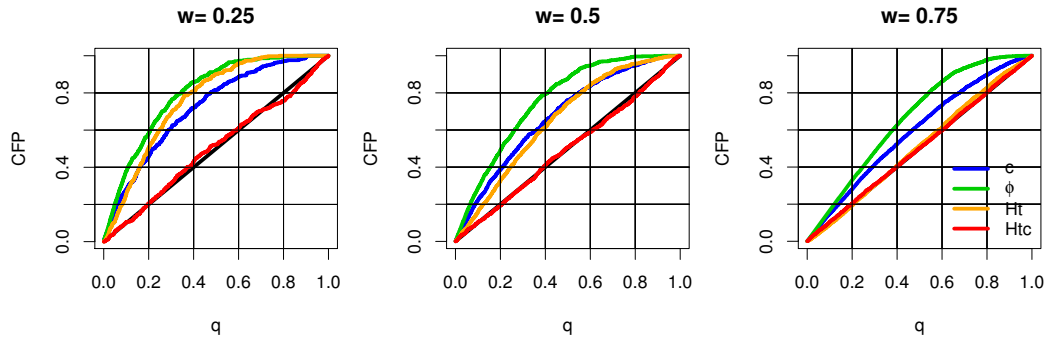

Figure 6.7: CFP curves associated with the random and imprecise variables of the slope stability analysis (case n°1).

the most optimistic probability bound ($w < 0.50$), it corresponds to the second most important parameters. The increase of CFP is steep for q from 0 to $\approx 70\%$ (over this domain, the CFP increases from 0 to > 90%), hence indicating that for this imprecise parameter, the cuts close to the support contribute the least.

From a risk management perspective, possible actions regarding the imprecise parameters could be focused on data acquisition. Based on this sensitivity analysis, it can be concluded that no extra characterisation effort is necessary considering H_{tc}: this is of interest in practice, because the tension crack is hardly accessible as being located in the upper part of the cliff. On the other hand, a special attention should be paid to estimate more precisely the toe geometry (using for instance terrestrial laser surveys as described by [Dewez et al., 2013]), which is "easier" to characterise in practice than the tension crack. Regarding the treatment of randomness associated with c and ϕ, their contribution cannot be reduced, because randomness is part of the system under study, i.e. it is its property. In this case, practical actions could be of preventive nature, e.g. by applying a safety margin, for instance by increasing the safety threshold from 1.0 to, say, 1.2.

6.3.3 Case study n°2: stability analysis in post-mining

The second application case is based on the stability analysis of [Piedra-Morales, 1991] in the “Beauregard” abandoned underground quarry located in the Pont-Eveque city near Lyon in the South-East of France. The “room-and-pillar” method was used for the extraction of Molasse rock of the Miocene geological epoch. The stability of the two pillars (Fig. 6.8) was assessed through a deterministic approach using the safety factor SF defined as the ratio of the pillar compressive strength R_c over the mean vertical stress S acting on the pillar. This indicator of mine pillar stability can be calculated using the Tributary Area Theory, which considers the

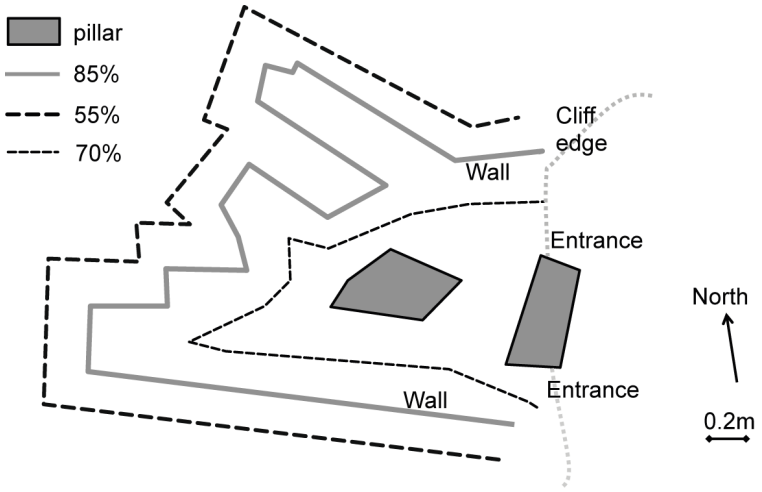

Figure 6.8: Schematic representation of the underground quarry of Beauregard (adapted from [Piedra-Morales, 1991]). Three zonation methods were used to estimate the extraction ratio f (ranging from 0.70 to 0.85).

total overburden load directly over the pillar and the extraction ratio f (defined as the ratio of the mined rock volume to the total rock volume):

$$SF = \gamma H \frac{1}{1-f} \tag{6.6}$$

where γ the unit weight of the overburden rock formation (expressed in kN/m^3) and H, the mining depth.

The measurement of the pillar compressive strength R_c was carried out using a digital rock strength index apparatus (Franklin Press), which yielded the following results: average value of 4.593 MPa (1 MPa = 10^6 Pa) with standard deviation of 1.412 MPa. This randomness was represented by a Gaussian (normal) probabilistic distribution whose parameters (mean and standard deviation) were constrained by the measurement results. The second random parameter is γ with an average value of 23 kN/m^3. We represent its uncertainty by a Gaussian probabilistic distribution assuming a standard deviation of 1 kN/m^3.

During the study, the estimate of f was made difficult, because of the geometry of the quarry (small dimensions with only two pillars, Fig. 6.8). Three methods were used, each of them using a different zonation assumption of the mined and total rock volume. This respectively yielded f = 55%, 70% and 85%, but the value of 70% was considered the more realistic. Based on this "expert" information, we modelled the imprecision on this parameter using a triangular

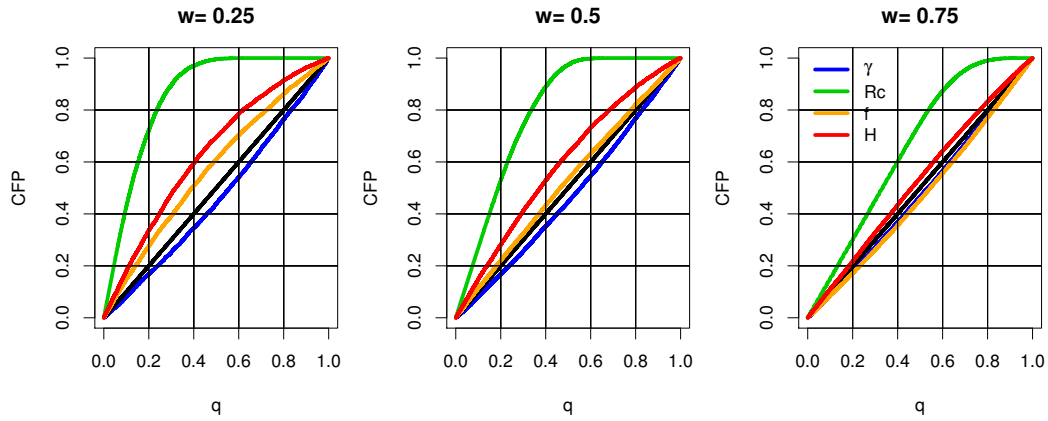

Figure 6.9: CFP curves associated with the random and imprecise variables of the mine pillar stability analysis.

possibilistic distribution with core 70% and support [55;85%]. Finally, the mining depth H appears to be variable over the quarry area and is modelled here by a triangular possibilistic distribution with core $12.5m$ and support $[6.5;16m]$.

The original study concluded to an absence of pillar failure, because the minimum calculated value for SF reached 5.0. Here, we investigate to which extent the assumptions on the extraction ratio could have influenced these conclusions, i.e. whether the estimate of the extraction ratio should have been more accurate regarding the failure probability $P(SF \leq 5.0)$. Based on the assumptions made for uncertainty representation, variability and imprecision are jointly propagated in the model using the IRS propagation method with 25,000 Monte-Carlo samples.

Figure 6.9 depicts the CFP curves associated with P_w assuming different degrees of risk aversion: $w = 0.25$ (the most optimistic); $w = 0.50$; $w = 0.75$ (the most pessimistic). Several observations can be made:

- the CFP curve associated with the random parameter R_c presents the largest deviation from the diagonal whatever the degree of risk aversion. The quantiles below $\approx 40\%$ ($w = 0.25$); below $\approx 50\%$($w = 0.50$) and below $\approx 60\%$($w = 0.75$) contribute the most to P_w (as indicated by the steep increase of the CFP curve);
- the other random parameter γ appears to have little influence, since it has the least deviation from the diagonal whatever the degree of aversion;
- the second most important source of uncertainty was the imprecision on H so that the cuts with α above 0.5 (q ranging from 0. to 0.5, i.e. close the core) have the greatest influence as indicated by the steep increase of the CFP;
- the imprecise extraction ratio f has low-to-moderate influence (the greatest influence

is for $w = 0.25$, but remains below the one of H).

Based on this sensitivity analysis, it can be concluded that the variability in R_c has the greatest contribution, which may be related to the heterogeneous nature of the Molasse rock extracted in this quarry. This analysis further provides ground for confidence on the original stability analysis, because it highlights that the influence of f remains limited though it presents a large imprecision due to the difficulty to estimate it in practice.

6.3.4 Case study n°3: numerical simulation for stability analysis in post-mining

In this section, we address the question of the influence of weak bands in the stability of mine pillars. This issue has been described and discussed regarding field data collection in limestone mines in the USA by [Esterhuizen and Ellenberger, 2007]. By weak bands, we refer here to very thin (with thickness from a few millimeters to centimeters, i.e. small compared to the total pillar thickness) sub-horizontal discontinuities, which typically layer limestone pillars and are generally composed of indurated clays or seat earths, i.e. materials closer to soils than rocks. In the case described by [Esterhuizen and Ellenberger, 2007], the presence of weak bands may have played a major role in the reduction of the average pillar strength, which appeared to be only 10-15% of the uniaxial compressive strength of limestone beds.

Model set-up and parameters

We rely on 2D finite-element simulations for evaluating the strength of a limestone pillar. The geometry is assumed known: we consider a 1m-high rectangular pillar with width of 0.8m and larger dimension in the normal direction so that we assume 2D plain strain conditions (the pillar considered here can be assimilated to a mine wall), see Fig. 6.10A). Horizontal weak bands are assumed uniformly distributed along the pillar thickness as schematically depicted in Fig. 6.10A). A Drucker-Prager plastic law [Drucker and Prager, 1952] is assigned to both rock materials (limestone and clay), with plastic yield surface F defined as follows:

$$F = \sigma_d + T \times \sigma_m - t \leq 0 \tag{6.7}$$

where σ_m is the mean stress defined by trace($\underline{\underline{\sigma}}$)/3 ; σ_d is the equivalent deviatoric stress defined as $\sqrt{\underline{\underline{\sigma_d}} : \underline{\underline{\sigma_d}}}$ with $\underline{\underline{\sigma}} = \underline{\underline{\sigma}} - \sigma_m \underline{\underline{1}}$ the deviatoric stress tensor; the shear strength of the material is related to T (linked to the internal friction coefficient) and to t (linked to the internal friction coefficient and to the internal cohesion). It should be underlined that

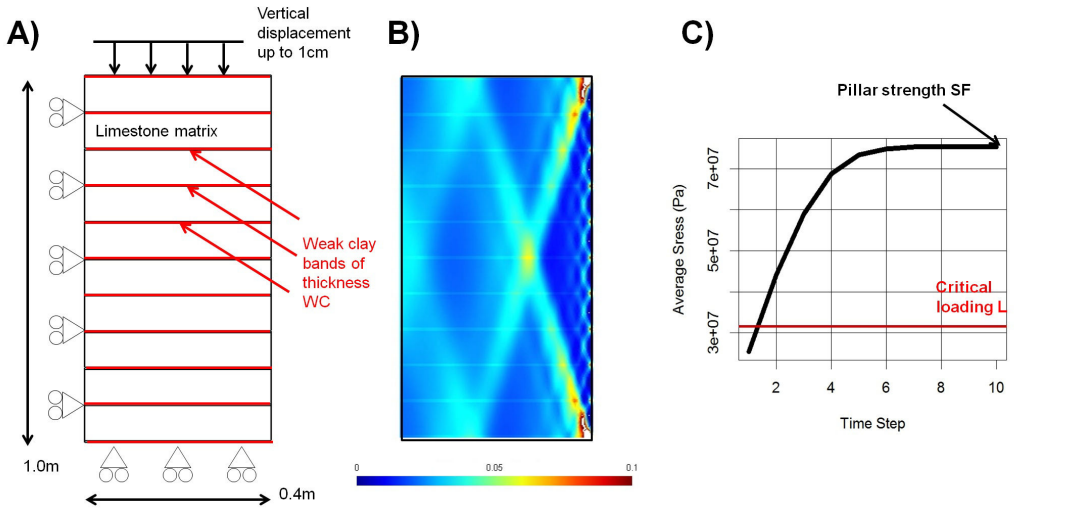

Figure 6.10: A) Model geometry and boundary conditions for evaluating the stress evolution at mid height of pillar during loading (imposed vertical displacement up to -1cm); B) Map of plastic shear strain at the end of the loading T_L=0.51; t_L=4.9 MPa; T_C=0.27; t_C=0.55 MPa; W_C=3.9mm (see definition of the parameter in Table 6.2); C) Typical stress evolution during loading.

more sophisticated rheological law can be assumed (e.g., Hoek and Brown law) and that the following procedure for uncertainty treatment can be applied whatever this assumption.

Vertical displacements are fixed at the bottom of the pillar and a monotonically increasing vertical displacement loading of -1 cm is imposed at the top (during 10 time steps). The problem was solved using the finite-element simulator Code Aster[1]. The average (spatial) stress at mid-height of the pillar is computed during loading following the procedure by [Esterhuizen and Ellenberger, 2007]. The spatial distribution of plastic strain over the whole pillar face is illustrated in Fig. 6.10B) after 6 time steps for the following parameters: T_L=0.51; t_L=4.9 MPa; T_C=0.27; t_C=0.55 MPa; W_C=3.9mm (see definition of the parameters in Table 6.2). A typical stress evolution is depicted in Fig. 6.10C). Two parts can be distinguished: 1. in a first part, the stress evolution is linear with loading indicating a elastic behaviour of the system (here before time step N°5); 2. the stress is blocked at a constant value indicating irreversible behaviour (i.e. plastic strain). We define the strength SF of the system as the maximum average stress reached during loading. The objective of the study is to estimate the probability that SF is inferior to a critical loading applied on the pillar at depth, here assumed to be 31 MPa.

[1] http://www.code-aster.org/

Table 6.2: Assumptions on the uncertainty representation for the properties of the rock materials composing the mine pillar.

Rock material	Property	Representation
Limestone	Young's modulus	Fixed at 35 GPa
Limestone	Poisson's coefficient	Fixed at 0.30
Limestone	Shear strength T_L	Normal law with mean of 0.47 and standard deviation of 0.05
Limestone	Cohesion t_L	Normal law with mean of 5.00 MPa and standard deviation of 0.25 MPa.
Weak band	Young's modulus	Fixed at 5 GPa
Weak band	Poisson's coefficient	Fixed at 0.25
Weak band	Shear strength T_C	Triangular Possibility distribution with core at 0.30 and support $[0.20, 0.40]$
Weak band	Cohesion t_C	Triangular Possibility distribution with core at 1.00 MPa and support $[0.50, 1.50]$ MPa.
Weak band	Thickness W_C	Triangular Possibility distribution with core at 5 mm and support $[2.5, 10.0]$ mm.

Uncertainty representation

Knowledge on both materials (limestone and weak band) are dissimilar. Elastic and strength parameters of the limestone rock materials are assumed to be well-constrained by a large collection of in-site observations, or at least from literature data. Probabilistic laws are assumed to capture the variability of such properties. On the other hand, the material composing the weak band can be heterogeneous and data can be very scarce due to practical difficulties to characterise them on site. This is mainly due to the small thickness of the band (of the order of a few mm), which prevents any proper collection of rock samples in site. An other source of information can be experts' opinions, and to process such qualitative data in flexible manner, we rely on Fuzzy Sets.

In the following, we assume the elastic properties to be known for both rock materials (see Table 6.2) and uncertainty treatment is restricted to the parameters of the yield surface and to the thickness W_C of the weak bands. This latter parameter is known to be difficult to characterize in site and can vary a lot from one pillar to another. Assumptions on the uncertainty representation are summarised in Table 6.2.

Uncertainty propagation

The objective is to evaluate the probability that SF is below a critical loading, here assumed of circa 30 MPa. Due to imprecision on the properties of the weak bands, this probability

is imprecise itself and calculated using procedures described in Sect. 6.2.2. Contrary to the afore-described case studies, the difficulty of the present case stems from the optimization procedure to find the upper and lower bounds of SF given intervals on the imprecise weak bands' parameters: this can be very computationally extensive, because SF is not necessarily monotonic with respect to these parameters. In our case, we relied on numerical gradient estimates, namely the limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm "L-BFGS-B" of [Byrd et al., 1995] for solving the bound constrained global optimisation problem. Finding the lower bound given the intervals on the band thickness W_C, and on the shear strength parameters T_C and t_C typically requires about one hour of calculation using a single computer unit. Note that an additional computationally intensive task resides in the update of the mesh model given each new thickness value.

To alleviate this computation problem, we rely on meta-modelling techniques as described in Chapter 3. Considering a mixture of aleatoric and epistemic uncertainties, [Eldred et al., 2011] proposed to rely on kriging-type meta-models to ease the global optimisation (using approaches described by [Jones et al., 1998]) regarding the propagation of interval-valued variables and on spectral methods (polynomial chaos expansions) regarding the propagation of random variables. Here, we considered Fuzzy Sets, i.e. nested intervals assigned to different levels of confidence. Keeping in mind that the hybrid propagation can be formulated using the levels of confidence of both random variables and Fuzzy Sets (see Sect. 6.2.4), we propose to extent the strategy originally developed by [Lockwood et al., 2012] for jointly propagating intervals and probabilities, to the case of fuzzy sets and probabilities. The basic idea is to construct two mappings (approximations) between the levels of confidence and the lower and upper bounds of the safety factor SF. This reads as follows:

- Step 1: randomly generate n_0 vectors of size n (the number of uncertain parameters) assuming uniform probability distributions: α_i (between 0 and 1), i=1,..,n. On this basis, generate k values of aleatoric parameters using their inverse cumulative distributions and the $n-k$ cuts, denoted I, assigned to the levels of confidence 1-α;
- Step 2: fix the aleatoric parameters at their sampled values; the optimisation problem is then solved to find the lower and upper bounds, $\underline{SF}$ and $\overline{SF}$, given each randomly generated interval I. This step can be eased using kriging-based optimisation method [Jones et al., 1998]. Re-conduct the procedure n_0 times;
- Step 3: at the end of step 2, we have two training data sets of the form $(\alpha_i;\overline{SF})_j$ and $(\alpha_i;\underline{SF})_j$, where $i=1,...,n$ and $j=1,...,n_0$. For each training data set, approximate the relationship between the bounds of SF and the levels of confidence α using a kriging-type meta-model;
- Step 4: once the quality of the approximation has been validated, use the meta-model to

evaluate the bounds on the probability of interest (using techniques described in Sect. 6.2.2) and apply the graphical tool to explore the contribution of each parameter.

The afore-described procedure is applied to the pillar case using $n_0 = 25$ different long-running simulations. We use the logarithm (base 10) of SF to improve the quality of the kriging meta-model (with linear trend, Matérn-type covariance and maximum-likelihood-based evaluation of the kriging parameters). Figure 6.11 shows the results of the leave-one-out cross validation procedure for both bounds: they appear to be approximated with a high level of quality (here indicated by a coefficient of determination R^2_{CV} superior to 0.98 in both cases).

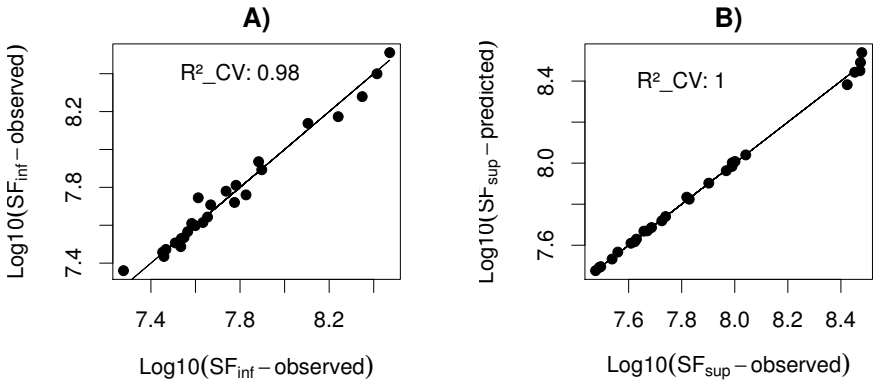

Figure 6.11: Leave-One-Out Cross-Validation procedure applied to both kriging-type meta-models for the lower (A) and upper (B) bounds of the average stress SF.

On this basis, both meta-models are used to evaluate the probability envelope as depicted in Figure 6.12 using 250,000 random samples of the levels of confidence. Obviously, this would not have been feasible using directly the "true" simulator. In the following, we focus on the weighed probability derived from both bounds using w=0.50.

The samples used for the uncertainty propagation are used as inputs of the graphical tools for sensitivity analysis. Figure 6.13A) shows the CFP curves for all parameters: the shear strength T_L of the limestone rock formation clearly drives the value of the probability of interest with the largest contribution of the quantiles from 0 to $\approx 25\%$. To ease the analysis of the other uncertain parameters, we slightly change the presentation of the results: Fig. 6.13B) shows the corresponding CFP curve by representing the evolution of $CFP(q) - q$ instead of CFP alone. The most influential parameters are the ones deviating the most from the horizontal line at 0. This clearly outlines the weak band's thickness W_C as the second most important parameter with the largest contribution.

From this analysis, actions in terms of risk management should focus on the application of a additional safety margin to "protect" from the randomness of the limestone shear strength.

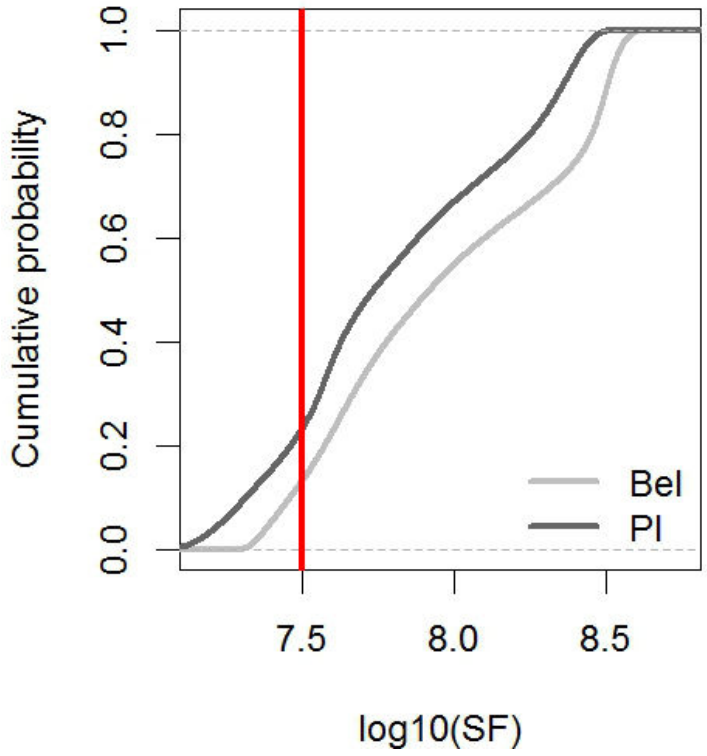

Figure 6.12: Upper (Pl) and Lower (Bel) probability distribution bounding the true probability assigned to the average stress at the end of loading. The vertical line outlines the critical threshold at $log10(SF)$=7.5 ($SF \approx 31$ MPa).

Regarding the influence of the weak bands' thickness, additional characterisation studies to better constrain this parameter could be undertaken. Interestingly, strength parameters of the weak band only slightly influence the results (for the considered assumptions), hence indicating that a precise characterisation is not necessary here: this is of high interest, because these parameters can be hard to characterize in site in practice due to the small thickness of the bands.

6.4 Concluding remarks for Chapter 6

In the present chapter, the different methods / procedures for handling a mixture of randomness and imprecision in sensitivity analysis have been reviewed. On this basis, a major limitation has been outlined, namely their computation time cost, especially when using numerical simulators. These new theories for uncertainty representation basically rely on interval-valued tools, so that uncertainty propagation usually involves optimisation procedure,

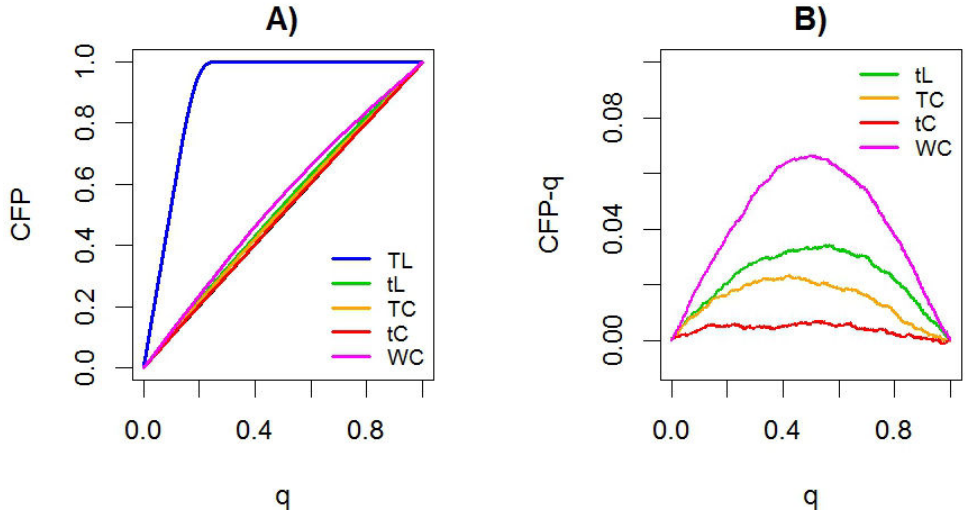

Figure 6.13: CFP derived from the joint propagation of imprecision and randomness for the pillar case using 250,000 random samples for the levels of confidence. A) CFP for all parameters; B) Representation of CFP-q versus q for all parameters except for the shear strength of the limestone.

which can be highly computationally intensive.

In this context, an easy-to-use graphical tool was proposed. It is based on an adaptation of the contribution to failure probability plot of [Li and Lu, 2013] to both handle probabilistic and possibilistic information for stability analysis. It should be acknowledged that the proposed modification of the graphical tool remains minor in practice, because it is merely based on the interpretation of the possibilistic distribution as sets of intervals associated with different levels of confidence. But to the authors' best knowledge, this "simple" modification has rarely been done in the statistical literature, whereas it can be of interest from a "knowledge" perspective. The advantages are the following ones:

- It allows placing on the same level random and imprecise parameters, i.e. it allows the comparison of their contribution in the probability of failure. Concrete actions from a risk management perspective can then be decided according to the contribution of both types of uncertainty;
- It allows highlighting the regions of the quantiles and of the nested intervals which contribute the most to the bounds of the failure probability, e.g. close to the core or close to the support. It allows investigating whether future investigations should be concentrated on the evaluation of the most likely value of the imprecise variable or on the spread;
- Such a graphical procedure has the great advantage to be conducted in a post-processing

step, i.e. by relying only on the model runs required for uncertainty propagation, i.e. at no extra computational cost.

It should be underlined that the proposed procedure can be easily extended to imprecise probabilities (represented for instance by a probability box) via sampling-based joint propagation approaches like the one of [Zhang et al., 2010] in the domain of structural safety.

To demonstrate the applicability of this tool, three real-cases were studied. These application examples can provide an additional interest for the geotechnical community. The easy-to-use tool allows handling situations, where it is questionable to use probabilities for uncertainty treatment:

- The first application case corresponds to stability analysis of steep slopes. The main imprecise parameters in this case are the tension crack's height located in the upper part of the cliff and the toe height;
- The second one corresponds to a stability analysis of abandoned underground quarry (based on the study by [Piedra-Morales, 1991]), where the extraction ratio was imprecise because it could only be estimated with great difficulties (due to the particular geometry of the quarry);
- The third one corresponds to the stability analysis of a pillar using numerical simulations. The imprecise parameters are the characteristics of the weak bands layering the pillar. Due to the high computation time cost of the joint propagation (imposing to both solve constrained optimisation problems and to apply Monte-Carlo-based sampling), we described a strategy relying on meta-modelling techniques similar to the ones detailed in Chapter 3.

Though the primary motivation of the graphical tool was to alleviate the computation burden of alternative approaches, some limitations may still exist when dealing with the estimates of low probabilities. In those situations, computational efficiency should be further improved by using, for instance, the adaptive radial-based importance sampling strategy proposed by [Li and Lu, 2013]. Finally, a very recent study by [Li et al., 2014] extended the CFP for handling to epistemic uncertainties in the framework of evidence theory: it would be of high interest to compare our approach with theirs and to set up a common framework to handle epistemic uncertainties whatever the type of uncertainty representation tools used, in particular by taking advantages of the bridges between the new uncertainty theories [Dubois, 2007].

7 Conclusions

7.1 Achieved results

The central topic of the present work is the treatment of epistemic uncertainty of type "parameter" in geo-hazard assessments. Contrary to aleatory uncertainty, this type of uncertainty can be reduced through additional measurements (additional lab tests, in site experiments, etc.) or modelling (e.g., through numerical simulations) or R&D efforts. Therefore, a possible option to manage this facet of uncertainty is to: 1. identify the contribution of the different input parameters in the uncertainty on the final hazard outcome; 2. rank them in terms of importance; 3. Decide accordingly the allocation of additional resources. This is the purpose of sensitivity analysis.

In the first chapter, the advanced variance-based global sensitivity analysis VBSA was applied on an analytical model for slope failure assessment. On this basis, the richness of the information of VBSA was demonstrated. By providing quantitative measures of sensitivity, the most important sources of parameter uncertainty can be identified (using the main effects) as well as the parameters of negligible influence (using the total indices). Besides, some key attributes of the model behaviour can be identified (using the sum of the main effects). Yet, to the author's best knowledge, this kind of analysis has rarely been conducted in the domain of geo-hazard assessments (except for the study by [Hamm et al., 2006]). This can be explained by the specificities of this domain, which impose to consider several constraints / limitations, which were are at the core the present PhD thesis.

The first limitation is related to the computation CPU time cost of most numerical models supporting geo-hazard assessments. The time required for a single simulation can reach several minutes or even hours, either because they are large-scale (for instance the application of simple slope stability analysis at the spatial scale of a valley), or because the underlying processes

are difficult to be numerically solved. This latter issue is illustrated by the La Frasse landslide model, as described in the introduction, which has a CPU time of about 4 days, because it involves a complex elastoplastic model describing the behaviour of the slip surface material. Despite the extensive research work on the optimization of the computation algorithms, VBSA remains computationally intensive, as it imposes to run a large number of simulations (> 1,000). In this context, VBSA can be made possible via the combination with meta-modelling techniques, for instance using kriging: this was demonstrated using the modelling study carried out by [Laloui et al., 2004] for the La Frasse landslide in [Rohmer and Foerster, 2011]. To the author's best knowledge, the application of such kinds of technique is original in the domain of landslide risk assessment. The computation burden was largely alleviated and the sensitivity of the landslide-induced displacements to the values of the slip surface properties was investigated: the first order sensitivity indices associated to these properties were derived using only 30 different simulations. In addition to proving the applicability of this strategy, the impact of meta-model error (i.e. the additional uncertainty introduced because the true simulator was replaced by an approximation) was discussed by treating the problem under the Bayesian formalism. This allowed assigning confidence intervals to the derived sensitivity measures: the importance ranking could then be done accounting for the limited knowledge on the "true" simulator (i.e. the limited number of long-running simulations), hence increasing the confidence in the importance ranking.

The second limitation is related to the nature of the parameters (input or output): VBSA deals with scalar parameters. Yet, in the domain of geo-hazard, parameters are often functional, i.e. they are complex functions of time or space (or both). This means that parameters can be vectors with possible high dimension (typically 100-1,000). For instance, the outputs of the La Frasse model correspond to temporal curves of the displacements (discretized in 300 steps) at any nodes of the mesh, i.e. the outputs are vectors of size 300 at any spatial location. Another example is the spatial distribution of hydraulic conductivities of a soil formation. Focusing on the functional output case, a methodology to carry out dynamic (global) sensitivity analysis of landslide models was described combining: 1. basis set expansion to reduce the dimensionality of the functional model output; 2. extraction of the dominant modes of variation in the overall structure of the temporal evolution; 3. meta-modelling techniques to achieve the computation, using a limited number of simulations, of the Sobol' indices associated to each of the modes of variation. Using the La Frasse case, it was shown how to extract useful information on dynamic sensitivity using a limited number (a few tens) of long running simulations [Rohmer, 2013]. The analysis of the sensitivity measures assigned to the dominant modes of variation is interpreted by adopting the perspective of the risk practitioner in the following fashion: "identifying the properties, which influence the most the possible

occurrence of a destabilization phase (acceleration) over the whole time duration or on a particular time interval". However, a limitation should be underlined, namely the physical interpretation of the dominant modes of variation, especially compared to the traditional time-varying VBSA (more easily interpretable, but also intractable for very long time series). Based on the study on the functional output, the applicability of the proposed methodology was also investigated for the case of functional inputs. Nevertheless, three limitations were outlined: 1. The number of necessary components in the basis expansion of the functional input can be very large (over several tens), which might impose to largely increase the number of training samples in order to reach a satisfactory level of approximation quality of the meta-model; 2. the level of truncation of the basis set expansion should be decided before carrying out any long-running simulations, i.e. with limited possibility to assess the impact of leaving a given level of information on the quality of the approximation; 3. as afore-mentioned, the physical interpretation of the sensitivity measures can be tedious. On this basis, a few lines of improvement were highlighted and constitute perspective for future research works.

Finally, a third limitation is related to the representation of uncertainty. By construction, VBSA is based on the assumption that the variance can capture the main features of the uncertainty. This not always holds true (e.g., [Auder and Iooss, 2008] in the probabilistic setting). Yet, in the domain of geo-hazard assessments, this validity is even more questionable, because data are often scarce, incomplete or imprecise. The major criticisms available in the literature against the systematic use of probability in such situations were reviewed. On this basis, the use of a flexible uncertainty representation tool was investigated, namely Fuzzy Sets to handle different situations of epistemic uncertainty. For each situation, examples of real cases in the context of geo-hazard assessments were used: hazard assessment related to underground abandoned cavities at regional scale [Nachbaur and Rohmer, 2011] and the earthquake risk scenario at Lourdes [Rohmer and Baudrit, 2011]. The situations are: i. Vagueness due to the gradual nature of words; ii. Reasoning with vague concepts; iii. Imprecision; iv. Imprecision on the parameters of a probabilistic model.

On this basis, the issue of sensitivity analysis considering a mixture of randomness and imprecision was addressed. Based on a literature review, a major limitation was outlined, namely the computation time cost: new approaches for uncertainty representation basically rely on interval-valued tools, the uncertainty propagation then involves optimisation procedure, which can be highly computationally intensive. In this context, an adaptation of the contribution to failure probability plot of [Li and Lu, 2013] to both handle probabilistic and possibilistic information was proposed [Rohmer and Verdel, 2014]: it allows placing on the same level random and imprecise parameters, i.e. it allows the comparison of their contribution in the probability of failure so that concrete actions from a risk management perspective

can be decided accordingly. Besides, the analysis is conducted in a post-processing step, i.e. at no extra computational cost and using only the samples of random intervals and of random numbers necessary for the propagation phase. The applicability of this easy-to-use tool was demonstrated using three cases, where it is questionable to use probabilities to treat uncertainty.

In summary, the present work should be seen as an effort to handle epistemic parameter uncertainties in geo-hazard assessments. First, the achievement is of methodological nature (methodology for conducting VBSA using long running simulators, methodology for conducting VBSA adapted to functional outputs, methodology for conducting sensitivity analysis adapted to both imprecision and randomness). This methodological work takes advantages of recent advances in the statistical community (VBSA, basis set expansion, Fuzzy Sets, Fuzzy random variables, hybrid propagation, etc.) to answer practical questions (what drives the uncertainty on the results of the hazard assessment? How to conduct multiple simulations when the simulation code takes one hour to be run? How should the uncertainty be treated when the only pieces of information available restrict to vague statements and a few quantitative estimates?). This was done either through the combination (e.g., meta-models, basis set expansion, VBSA) or the adaptation of these techniques (e.g., contribution to probability of failure sample plot). A great attention has been paid to investigate the applicability of each proposed technique / procedure i.e. by highlighting the pros and cons through the confrontation to real cases. This constitutes the second achievement of the present work.

7.2 Open questions and Future developments

7.2.1 Model uncertainty

In the present work, the analysis was restricted to one single type of epistemic uncertainty, namely parameter uncertainty. Other types exist (see Sect. 1.3 in the introduction). Among them, model uncertainty has been recognised to have a significant influence on the geo-hazard assessment's results, if not larger than parameter uncertainty. Some examples of studies are provided by [Saeidi et al., 2013] considering the impact of building damage assessment methods for risk analysis in mining subsidence regions and by [Eidsvig et al., 2014] for quantifying the model uncertainty in debris flow vulnerability assessment.

Following [Pappenberger and Beven, 2006], a model can be defined (in hydrology, but the definition can be considered valid for other domains of application): "a model is an abstract construct to represent a system for the purposes of reproducing, simplifying, analyzing, or understanding it. Any model is based on a perceptual model (summary of our (personal)

perceptions on how a system responds), which gets translated into a conceptual model (mathematical description and implementation as a procedural model (computer code))". Based on this definition, two forms of model uncertainty can be identified.

1. Uncertainty can appear in the structure/form of the model, which depends on the choice of variables, dependencies, processes and so forth regarded as relevant and prominent for their purpose in the model;

2. Uncertainty can stem from the unambiguous choice of the "best" model to be used: in some cases, a set of different models (e.g. differing in their structure and input variables) can be considered equally adequate regarding two criteria: i. fitting past observations: they can accurately reproduce the previously observed behaviour of the considered natural system and ii. predictive capability: they can be used with confidence for predicting "yet-unseen" behaviour. This is exemplified by the extensively debated issue of selecting appropriate ground motion prediction equations (e.g., [Bommer et al., 2010]). Another example is the influence of the selected mechanical model (planar shear model; Bishop's method with circular shear surface, wedge analysis, Finite Element method) and the definition of safety within it for assessing a slope stability as discussed by [Fellin, 2005].

The criteria defining the degree to which model is considered "adequate" can be appreciated through the validation of model-based predictions with respect to real observations. Regarding the first criterion (fitting the past observations), this can be achieved by calibrating the uncertain input parameters, so that the model outputs can accurately reproduce the observations. The calibration exercise (also known as history matching in the domain of petroleum reservoir engineering or back analysis in the domain of geotechnical engineering) can be conducted either from a Bayesian updating perspective (e.g., [Kennedy and O'Hagan, 2001]), from a data assimilation perspective using for instance Ensemble Kalman Filter or from an optimization-based perspective (see a review by [Oliver and Chen, 2011] in the domain of petroleum reservoir engineering). The common difficulty to all methods relates to the computational hurdle posed by the huge computation time cost of numerical models: the use of meta-models turns here to be a key element. In particular, the Gaussian-Process-type meta-model defined in a Bayesian setting (as described in Chapter 3) offers a flexible framework to incorporate a discrepancy term (defined as the differences between the model outputs and the measurements) in the learning phase of the meta-model [Kennedy and O'Hagan, 2001]. Though this procedure can overcome the computation problem posed by the use of the numerical model, it remains limited by the number of model parameters that should be calibrated (of the order of 10). In this context, a preliminary sensitivity analysis can be very useful via, for

instance, the variable of interest defined as the differences with the observations. An example of advanced procedures is the combination of GSA and the GLUE approach in the domain of hydrology [Ratto et al., 2001].

Regarding the second criterion (predictive capability), it is often assumed that if the numerical model has been carefully calibrated then it will be able to forecast the system's behaviour. Though this assumption turns to be valid in many studies (see for instance good achievements for shallow, rainfall-triggered landslides forecasting by [Schmidt et al., 2008]), this not always holds true as shown by [Carter et al., 2006] with an example from the petroleum industry and by [Konikow and Bredehoeft, 1992] for groundwater models. Though the confrontation to the real measurements is an efficient approach for reducing model uncertainty, models remain "simplified representations of the phenomena and compliance between the model assumptions and the properties of the system being analysed never exist in an absolute sense" as underlined for instance by [Nilsen and Aven, 2003]. Besides, the modelling task cannot be considered a stand-alone exercise in the sense that model construction has always a purpose and in the case considered here, it is meant to support decision-making for risk management. In this view, it is based on a trade-off between accuracy and simplicity of the model [Aven and Zio, 2011] together with an explicit and transparent description of the model assumptions. Developing robust and efficient methods (and accepted by the community) for dealing with such type of epistemic uncertainty still constitutes a matter of ongoing research (e.g. [Parry and Drouin, 2009] in the field of nuclear safety).

From a sensitivity analysis's perspective, future research studies should focus on the following question: given both sources of epistemic uncertainty (model and parameter), should the effort be primarily focused on the characterisation task to reduce uncertainty in the model parameters' value or on the modelling task. For instance, should the effort be done on the in-site characterization studies or on the sophistication of the soil material's behaviour? To illustrate this latter issue, Laloui and co-authors showed the benefit of using the Hujeux constitutive law compared to the commonly-used Mohr-Coulomb one in the La Frasse case. To the author's best knowledge, few studies exist allowing weighing the importance of model uncertainty against parameter uncertainty. Most existing studies restrict to a "one-factor-at-a-time" (OAT) approach applied to a set of plausible models (for instance [Rabinowitz and Steinberg, 1991]). Recently, [Jacques et al., 2006] proposed a methodology to handle two situations related to model uncertainty using the tools of VBSA, namely when an uncertain input becomes known (determinist) and when the goal is to estimate the sensitivity indices of a model corresponding to the sum of two models (either without shared common inputs, or a mixture of shared and not-shared ones) for which sensitivity indices are available. Further work should be done to generalise VBSA to this problem.

7.2.2 Use of new uncertainty theories for practical decision-making

In chapter 5, Sect. 5.1, the main arguments against the systematic use of probabilities were outlined in situations where the available data are scarce, imprecise or vague. New theories for uncertainty treatment have been proposed in the literature (see a recent review by [Dubois and Guyonnet, 2011]) and in particular the Possibility Theory, which relies on the use of Fuzzy Sets. By representing both types of uncertainty with the most appropriate mathematical format, it allows being "inter-subjective" as discussed by [Aven and Zio, 2011], i.e. in the sense that it ensures that the "representation corresponds to documented and approved information and knowledge" (contrary to the Bayesian setting). This hybrid treatment of uncertainty leads to a communication of uncertainty in the form of ill-known (imprecise) probabilities, which can be summarized, for instance, using an interval where the true probability should lie. Such a format has the advantage of providing a clear picture of what is not known (see Chapter 5). Therefore, the width of the probability interval can be used as a sensitivity index to indicate that additional effort should be done to reduce epistemic uncertainty as exemplified by the pair of maps for the seismic risk scenario at Lourdes, France (Fig. 5.7).

To my mind, adopting such an approach goes in the right direction because "if our science is to be meaningful, we should aim to communicate the limitations of the predictions" [Pappenberger and Beven, 2006]. Yet, it should be recognized that this transparency can turn to be a disadvantage depending on the perception of the audience to which the message is conveyed. For practical decision-making, the format used to communicate uncertainty is of primary importance. The concept of probability itself can pose problem as its interpretation can be twofold, either a frequency value or a degree of belief (within the Bayesian setting). Depending on the adopted frameworks, people perceive and act differently in situation of uncertainty (see the surveys reported by [Patt and Dessai, 2005]).

Hence, using interval-valued probabilities (within new formal settings as described in Appendix D) might add difficulties for practical decision making. The complicated representation of uncertainty has often been a major criticism to these new theories (see e.g., [Lindley, 2000]), which can appear to be less transparent than those of probability as outlined by [Rougier and Beven, 2013]. Put in other words, the danger is to add more confusion than insights [Aven and Zio, 2011]. Recall that using an interval-valued probability is not a "new" format for risk analysis: it has even been recommended for situations where the probability values are low like in the nuclear setting as reported by [Ellingwood and Kinali, 2009] within a Bayesian setting. Yet, it should be acknowledged that decision-makers may not feel comfortable in using such a format. Should the most pessimistic value, say the lower bound, be used? If so, the more optimistic values are neglected. Otherwise, should the average value be used?

[Dubois and Guyonnet, 2011] extensively discussed such a problem and defined an index which reflects the attitude of the decision-maker to risk (i.e. the degree of risk aversion), see Chapter 6. This allows introducing subjectivity only at the very end of the risk analysis chain instead of introducing it from the very beginning, in the phase dedicated to the mathematical representation of uncertainty.

Despite this formal clarification, familiarisation of the decision maker to these new settings remains difficult and deserves future research. Efforts like the recent studies by [Pedroni et al., 2013, Loschetter et al., 2015], aiming at comparing the results obtained by the standard probabilistic (Bayesian or frequentist) approach and the ones derived from the use of new uncertainty theories, should also be intensified. Interval-valued probabilities may be a appropriate tool for "fulfilling the transparency requirement of any risk assessment, but this may not be sufficient to achieve the level of confidence necessary for assisting the deliberation process and decision making" [Aven and Zio, 2011]. Put in other words, the width of the probability interval provides a clear picture of the "flaws in the assessment process", but if it is not transferred cautiously from scientists to end-users, this might undermine the confidence in the risk analysis, potentially leading to a loss of credibility in the results. Take a situation of large degree of epistemic uncertainty when the lower and the upper bounds can respectively reach 0 and 1: this can be hard to communicate. A possible line of research could be based on lessons drawn from the Intergovernmental Panel on Climate Change (IPCC). They propose to use words to describe probabilities (see the background paper by [Schneider and Moss, 1999]). Though this format can present some limitations in practices [Patt and Dessai, 2005], a code of practice shared by all practitioners and a comprehensive terminology to convey the uncertain results as represented by these new theories is desirable in the future.

Finally, an additional barrier to the systematic use of these techniques is their practical implementation, which goes in pair with the afore-described criticism of sophistication. Manipulating hybrid mathematical tools goes at the expense of additional computational efforts: in a broad sense, the hybrid uncertainty propagation imposes not only to randomly sample probability distributions, through Monte-Carlo-like techniques, but also to solve constrained optimization problems (see e.g., [Baudrit et al., 2006]). The combination with meta-modelling techniques as proposed by [Eldred et al., 2011, Lockwood et al., 2012] is promising: this is exemplified with the third case tackled in Chapter 6, Sect. 6.3.4. Yet, additional effort should be done in the future to accurately account for the meta-model error, as discussed in Chapter 3, in the final results of the analysis to efficiently support decision-making.

A Functional decomposition of the variance: the Sobol´ indices

The present appendix describes the basic concepts underlying the computation of the Sobol´ indices using a numerical model f. The appendix is based on [Iooss, 2011] and on chapter 4 of [Saltelli et al., 2008].

Consider a square integrable function f over the n-dimensional unit hypercube $\Omega^n = [0;1]^n$. Define $\boldsymbol{X}$ and Y respectively the n-dimensional vector of scalar (here assumed to be independent) inputs and the scalar output of f so that $Y = f(X_1, X_2, ..., X_n)$. [Sobol', 1990] proposed the following decomposition of f into summands of increasing dimension:

$$f(x) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{i<j} f_{ij}(x_i, x_j) + ... + f_{1,...,n}(x_1, ..., x_n) \tag{A.1}$$

This expansion is called the high-dimensional model representation HDMR of f. Assume that each term has zero mean $\int_0^1 f(x_i)dx_i = 0$, then all terms are orthogonal such that $\int_0^1 f(x_i)f(x_j)dx_i dx_j = 0$. In this manner, the different terms of Eq. A.1 can be expressed in a unique manner using the conditional expectations of Y:

$$\begin{aligned} f_0 &= E(Y) \\ f_i &= E(Y|X_i) - f_0 \\ f_{ij} &= E(Y|X_i, X_j) - f_i - f_j - f_0 \end{aligned} \tag{A.2}$$

A direct relationship with the variance exists in the following manner:

$$\begin{array}{lcccl} Var(f_i(X_i)) & = & V(E(X|X_i)) & = & V_i \\ Var(f_{ij}(X_i,X_j)) & = & V(E(X|X_i,X_j)) - V(E(X|X_i)) - V(E(X|X_j)) & = & V_{ij} - V_i - V_j \end{array} \tag{A.3}$$

An indicator of importance of X_i could be the reduction of uncertainty in Y (here represented by the variance of Y) if we were able to fix X_i to its true value x_i^*, i.e. $Var(Y|X_i = x_i^*)$. Yet, in most cases, this true value is unknown and to overcome this problem, we can use the average of this conditional variance over all possible values for x_i^*, i.e. $E(Var(Y|X_i = x_i^*))$. This is the expected variance of the output provided that X_i is fixed. Recall now the following property:

$$Var(Y) = Var(E(Y|X_i = x_i^*)) + E(Var(Y|X_i = x_i^*)) \tag{A.4}$$

If X_i is very influential, this means that fixing it implies the expected variance (right term of Eq.A.4) to be almost zero. Therefore, an importance measure can be defined using the left term of Eq.A.4, which is called the variance of the conditional expectation. The greater the influence of X_i, the larger this quantity. On this basis the first order sensitivity index, ranging from 0 to 1, can be written by dividing V_i by the unconditional variance $V(Y)$ as follows:

$$S_i = \frac{V(E(Y|X_i))}{V(Y)} \tag{A.5}$$

In a similar manner, the second-order index measuring the joint effects of X_i and X_j can be defined

$$S_{ij} = \frac{V_{ij}}{V(Y)} \tag{A.6}$$

The higher order terms can be expressed similarly.

On this basis, the ANOVA-HDMR decomposition can be written:

$$\begin{array}{lcl} V_Y & = & \sum_{i=1}^{n} V_i + \sum_{i<j} V_{ij} + ... + V_{1,...,n} \\ 1 & = & \sum_{i=1}^{n} S_i + \sum_{i<j} S_{ij} + ... + S_{1,...,n} \end{array} \tag{A.7}$$

To ease the practical use of the $2^n - 1$ sensitivity indices, the total index has been introduced [Homma and Saltelli, 1996] as the sum of all indices related to the X_i (first and higher order), hence measures not only the contribution of X_i alone, but also all the possible interactions with the other inputs. For instance for $n = 3$, S_{T_1} corresponds to $S_1 + S_{12} + S_{13} + S_{123}$. Recalling Eq.A.7, it is equivalent to write: $1 - S_{23} - S_2 - S_3$. Thus, if the total index is null, this means that the uncertainty on Y is mostly driven by the second and the third input parameter. Therefore, fixing the first parameter to any value has negligible influence. Formally, the total index can then be defined as follows :

$$S_{T_i} = 1 - \frac{V(E(Y|X_{-i}))}{V(Y)} \tag{A.8}$$

where X_{-i} corresponds to $(X_1, ..., X_{i-1}, X_{i+1}, ..., X_n)$.

In practices, Monte-Carlo-based numerical procedures can be used to compute the first and total indices. Here, we present the one proposed by [Saltelli, 2002], which furthers extent the one of [Sobol', 1990] and [Homma and Saltelli, 1996].

Generate a $(N \times 2 \cdot n)$ matrix of random numbers (N is the number of samples and n is the number of uncertain parameters). Then, define three matrices: $\boldsymbol{A}$ and $\boldsymbol{B}$ by simply splitting the original matrix into matrices of equal size and a third one $\boldsymbol{C}_i$ formed by all columns of $\boldsymbol{B}$ except for the i^{th} column taken from $\boldsymbol{A}$. Now, run the simulator f and compute $y_A = f(\boldsymbol{A})$, $y_B = f(\boldsymbol{B})$ and $y_{C_i} = f(\boldsymbol{C}_i)$. The first order sensitivity index can be defined as follows:

$$S_i = \frac{y_A \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} = \frac{1/N \cdot \sum_{j=1}^{N} y_A^{(j)} \cdot y_{C_i}^{(j)} - f_0^2}{1/N \cdot \sum_{j=1}^{N} y_A^{(j)} \cdot y_A^{(j)} - f_0^2} \tag{A.9}$$

where f_0 is the mean $1/N \cdot \sum_{j=1}^{N} y_A^{(j)}$.

Similarly, the total index holds as follows:

$$S_{T_i} = 1 - \frac{y_B \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} = \frac{1/N \cdot \sum_{j=1}^{N} y_B^{(j)} \cdot y_{C_i}^{(j)} - f_0^2}{1/N \cdot \sum_{j=1}^{N} y_A^{(j)} \cdot y_A^{(j)} - f_0^2} \tag{A.10}$$

The total cost in terms of number of necessary model runs is $N \times (n+2)$ corresponding to

$N+N$ for the computation of $\boldsymbol{A}$ and $\boldsymbol{B}$, plus n times N for matrix $\boldsymbol{C}_i$ (defined for each of the n input parameters).

B Universal kriging equations

We introduce here the equations for universal kriging modelling. For a more complete introduction to kriging meta-modelling and full derivation of equations, the interested reader can refer to [Sacks et al., 1989, Jones et al., 1998, Forrester et al., 2008, Santner et al., 2003].

Let us consider the deterministic response of the simulator $y = f(\boldsymbol{x})$ as a realization of a Gaussian stochastic process F so that $f(\boldsymbol{x}) = F(\boldsymbol{x}, \omega)$ where ω belongs to the underlying probability space Ω. In the following, we use the notation $F(\boldsymbol{x})$ for the process and $F(\boldsymbol{x}, \omega)$ for one realization. The process F results from the summation of two terms:

- $f_0(\boldsymbol{x})$ the deterministic mean function, which takes the general form $\sum_{i=1}^{p} \beta_i \cdot h_i(\boldsymbol{x})$ where $h(\boldsymbol{x}) = (h_1(\boldsymbol{x}), ..., h_p(\boldsymbol{x}))$ is a vector of p linearly independent known functions (named basis functions), and $\boldsymbol{\beta}$ is a vector of unknown coefficients;
- $Z(\boldsymbol{x})$ the Gaussian centred stationary stochastic process characterized by a zero mean and the covariance matrix $\boldsymbol{C}$, which depends on the variance σ_Z^2 and on the correlation function R, which governs the degree of correlation through the use of the vector of length-scale parameters $\boldsymbol{\theta}$ between any input vectors.

The covariance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is then expressed as $C(\boldsymbol{u}, \boldsymbol{v}) = \sigma_Z^2 \cdot R(\boldsymbol{u}, \boldsymbol{v})$, where $\boldsymbol{u} = (u_1, u_2, ..., u_n)$ and $\boldsymbol{v} = (v_1, v_2, ..., v_n)$ are two input vectors of dimension n.

Let us define $\boldsymbol{X}_D$ the design matrix composed of n_0 vectors of n-dimensional input parameters $\boldsymbol{x}$ (i.e. the training samples) to be simulated so that $\boldsymbol{X}_D = (\boldsymbol{x}^{(1)}; ...; \boldsymbol{x}^{(n_0)})$ and $\boldsymbol{y}_D$ the vector of model output associated with each selected training samples so that $\boldsymbol{y}_D = (y^{(1)} = f(\boldsymbol{x}^{(1)}), ..., y^{(n_0)} = f(\boldsymbol{x}^{(n_0)}))$.

Under the afore-described assumptions, the distribution of the model output for a new input vector of input conditions x^* follows a Gaussian distribution conditional on the design

matrix $\boldsymbol{X}_D$ and of the corresponding model outputs $\boldsymbol{y}_D$ with expected value $\hat{y}$ for the new configuration $\boldsymbol{x}^*$ and variance s^2 respectively defined by the universal kriging equations Eq.B.1 and B.2:

$$\hat{y}(\boldsymbol{x}^*) = h(\boldsymbol{x}^*)^T \cdot \hat{\boldsymbol{\beta}} + \boldsymbol{c}(\boldsymbol{x}^*)^T \cdot \boldsymbol{C}_D^{-1} \cdot (\boldsymbol{y}_D - \boldsymbol{H}_D \cdot \hat{\boldsymbol{\beta}}) \tag{B.1}$$

with $\hat{\boldsymbol{\beta}} = (\boldsymbol{H}_D^{-1} \cdot \boldsymbol{C}_D^{-1} \cdot \boldsymbol{H}_D)^{-1} \boldsymbol{H}_D^T \cdot \boldsymbol{C}_D^{-1} \cdot \boldsymbol{y}_D$ the vector of generalised least square estimates of $\boldsymbol{\beta}$.

$$s(\boldsymbol{x}^*)^2 = \sigma_Z^2 - \boldsymbol{c}(\boldsymbol{x}^*)^T \cdot \boldsymbol{C}_D^{-1} \cdot c(\boldsymbol{x}^*) + (\boldsymbol{h}(\boldsymbol{x}^*)^T - \boldsymbol{c}(\boldsymbol{x}^*)^T \cdot \boldsymbol{C}_D^{-1} \cdot \boldsymbol{H}_D) \cdot (\boldsymbol{H}_D^T \cdot \boldsymbol{C}_D^{-1} \cdot \boldsymbol{H}_D)^{-1} \cdot (\boldsymbol{h}(\boldsymbol{x}^*)^T - \boldsymbol{c}(\boldsymbol{x}^*)^T \cdot \boldsymbol{C}_D^{-1} \cdot \boldsymbol{H}_D)^T \tag{B.2}$$

where $\boldsymbol{H}_D = [\boldsymbol{h}(\boldsymbol{x}^{(1)}), \boldsymbol{h}(\boldsymbol{x}^{(2)}), ..., \boldsymbol{h}(\boldsymbol{x}^{(n_0)})]$ is the design $n_0 \times p$ matrix, and $\boldsymbol{c}(\boldsymbol{x}^*)$ is the covariance vector between the test candidate $\boldsymbol{x}^*$ and the training samples.

Note that the above equations can easily be written using the correlation matrix $\boldsymbol{R}$ and its relationship between $\boldsymbol{C}_D$ and σ_Z^2.

The specific case, where the basis functions reduce to a unique constant function, corresponds to the "ordinary kriging" as described in Sect. 3.3.

The most common technique for estimating the kriging parameters $(\boldsymbol{\beta}, \sigma_Z, \boldsymbol{\theta})$ relies on the likelihood function f defined as follows (see details e.g., in [Santner et al., 2003]):

$$f(\boldsymbol{y}_D | \boldsymbol{\beta}, \sigma_Z, \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma_Z^2)^{n_0/2} \sqrt{det \boldsymbol{R}}} \exp\left(-0.5 \frac{(\boldsymbol{y}_D - \boldsymbol{H}\boldsymbol{\beta})^T \boldsymbol{R}^{-1} (\boldsymbol{y}_D - \boldsymbol{H}\boldsymbol{\beta})}{\sigma_Z^2}\right) \tag{B.3}$$

The maximum likelihood estimate of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{y}_D \tag{B.4}$$

Substituting $\boldsymbol{\beta}$ by its estimate in Eq. B.3 and maximizing it with respect to σ_Z leads to its generalized least-squares estimate:

$$\hat{\sigma}_Z^2 = 1/n_0 (\boldsymbol{y}_D - \boldsymbol{H}\hat{\boldsymbol{\beta}})^T \boldsymbol{R}^{-1} (\boldsymbol{y}_D - \boldsymbol{H}\hat{\boldsymbol{\beta}}) \tag{B.5}$$

Substituting by their estimates in Eq. B.3, the likelihood f becomes:

$$f(\boldsymbol{y}_D | \boldsymbol{\theta}) = (2\pi\hat{\sigma}_Z^2)^{-n_0/2} (det \boldsymbol{R})^{0.5} \exp(-0.5 n_0) \tag{B.6}$$

The estimate of $\boldsymbol{\theta}$ is then found by maximizing the logarithm of Eq. B.6 using global optimisation procedures as developed by [Roustant et al., 2012]. Finally, it should be noted that the universal kriging equations correspond to the conditional mean and variance in the Bayesian framework [Helbert et al., 2009], in the case where an non-informative prior is given to the regression coefficients and the covariance parameters are taken at constant values (given by the maximisation of the likelihood function, see further details in Appendix C).

C Key ingredients of a bayesian treatment of kriging-based meta-modelling

In this appendix, we describe the key ingredients of Bayesian treatment of kriging-type meta-model for computing Sobol' indices together with a measure reflecting the meta-model error. The appendix is primarily based on [Gramacy, 2005, Le Gratiet, 2013].

C.1 Principles of Bayesian Model Averaging

Following a similar description as [Hoeting et al., 1999], consider the following problem. A geotechnical engineer has gathered data concerning pillar failures under different conditions (pillar height to width ratio, rock formation, presence of fractures, etc.). The engineer proposes a model M^* to link failure with the parameters describing those conditions by checking that M^* fits the data reasonably well and decides to use the model in a predictive manner, i.e. to use it for not-yet-seen conditions. Though this is a standard way of proceeding, this may not be fully satisfactory since it does not handle the possibility of an alternative model M^{**} that can also provide a good fit to the data but might lead to substantively different predictions. Bayesian model averaging (e.g., [Hoeting et al., 1999]) provides an approach for dealing with such an issue. Define y the quantity of interest. Bayesian techniques rely on the definition of prior probability distributions, which encode the prior knowledge (or ignorance) of the modeling scientist(s), and, for instance, can be based on past experiments. The posterior probability distribution assigned to y is the updated prior probability on y given newly observed data D: it is defined as the average of the posterior distributions under each of the models considered $M_1, ..., M_K$, weighed by their posterior model probability given as

$$p(y|D) = \sum_{k=1}^{K} p(y|M_k, D) \cdot p(M_k|D) \tag{C.1}$$

where the posterior probability for model M_k is given by

$$p(M_k|D) = \frac{p(D|M_k) \cdot p(M_k)}{\sum_{i=1}^{K} p(D|M_i)p(M_i)} \tag{C.2}$$

where the integrated likelihood of model M_k holds as

$$p(D|M_k) = \int p(D|\boldsymbol{\psi}_k, M_k) \cdot p(\boldsymbol{\psi}_k|M_k) d\boldsymbol{\psi}_k \tag{C.3}$$

with $\boldsymbol{\psi}_k$ corresponds to the vectors of parameters of model M_k, $p(\boldsymbol{\psi}_k|M_k)$ is the prior of $\boldsymbol{\psi}_k$ under model M_k, $p(D|\boldsymbol{\psi}_k, M_k)$ is the likelihood, and $p(M_k)$ is the prior probability that M_k is the true model given that one of the models considered is true.

C.2 Monte-Carlo-based procedures

Families of priors, which induce posterior distributions in the same family when combined with a likelihood are called "conjugate". These have the advantages of leading to analytically tractable posteriors. When conjugate priors cannot be found to adequately encode prior beliefs about models and parameters, posterior inference usually proceeds by simulation. Markov chain Monte Carlo (MCMC) can be used (e.g. [Gelman et al., 2013]) for posterior inference by simulation. A Markov chain is established so that its stationary distribution is the targeted posterior distribution. The transition probabilities from state θ_n to θ_{n+1} of the Markov chain, representing samples from the posterior of θ, can be set up either using the Metropolis-Hastings (MH) [Metropolis et al., 1953, Hastings, 1970], or Gibbs [Geman and Geman, 1984] algorithms.

The MH algorithm proceeds by proposing a new θ^* from a proposal distribution $q(\theta^*|\theta_n)$. The next (n + 1st) sample from the posterior for θ is chosen based on a ratio of posterior and proposal distributions, called the MH acceptance ratio defined as

$$\alpha = \min\{1, \frac{p(y|\theta^*)p(\theta^*)q(\theta_n|\theta^*)}{p(y|\theta_n)p(\theta_n)q(\theta^*|\theta_n)}\} \tag{C.4}$$

Here the estimates of $p(y)$ is not required, since the interest is only on a ratio of posteriors. The randomly proposed θ^* is accepted or rejected based on α:

$$\begin{aligned} \theta_{n+1} &= \theta^* \quad \text{with probability } \alpha \\ \theta_{n+1} &= \theta_n \quad \text{with probability } 1-\alpha \end{aligned} \tag{C.5}$$

Gibbs sampling is a special case of MH where $q(\theta^*|\theta_n) = p(\theta^*|y)$ so that $\alpha = 1$ and all proposals are accepted. Parameters with conditionally conjugates priors can usually be sampled using Gibbs algorithm.

C.3 Bayesian kriging

In this section, we consider the problem of bayesian model averaging considering the set of kriging models (details and definition of the parameters in Appendix B), each of them defined by parameters (aka hyper-parameters), corresponding to the regression coefficients, the variance of the stochastic process and the length-scales, namely $\boldsymbol{\psi} = \{\boldsymbol{\beta}, \sigma_Z, \boldsymbol{\theta}\}$.

Following [Rasmussen and Williams, 2006], we assume that the hyper-parameters are linked together, hence defining a hierarchical scheme. At the lowest level, the parameter $\boldsymbol{\beta}$ is considered. At the second level, the parameter σ_Z controls the distribution $\boldsymbol{\beta}$. At the top level, the knowledge of $\boldsymbol{\theta}$ controls the prior distribution of $\boldsymbol{\beta}$ and σ_Z^2. When little is known, an uninformative (improper) distribution on θ such that $p(\theta) \propto \frac{1}{\theta}$, can be defined. See [Helbert et al., 2009] for a discussion on more informative process-based priors. The process variance σ_Z^2 is commonly assigned an inverse gamma prior distribution, denoted $p(\sigma_Z^2|\boldsymbol{\theta})$. A normal distribution is then assigned to $\boldsymbol{\beta}$ whose variance is related to σ_Z^2, and denoted $p(\boldsymbol{\beta}|\sigma_Z, \boldsymbol{\theta})$.

Let us consider $\boldsymbol{y}_D$, the n_0 observations corresponding to the long-running simulations' results given different inputs' configurations $\boldsymbol{x}$. The posterior distribution for any new input configuration $\boldsymbol{x}^*$ is

$$p(y(\boldsymbol{x}^*)|\boldsymbol{y}_D) = \int p(y(\boldsymbol{x}^*)|\boldsymbol{y}_D, \boldsymbol{\theta}) p(\boldsymbol{\theta}|\boldsymbol{y}_D) d\boldsymbol{\theta} \tag{C.6}$$

The first term in the integral of Eq. C.6 is integrated with respect to σ_Z^2

$$p(y(\boldsymbol{x})|\boldsymbol{y}_D, \boldsymbol{\theta}) = \int p(y(\boldsymbol{x})|\boldsymbol{y}_D, \boldsymbol{\theta}, \sigma_Z^2) p(\sigma_Z^2|\boldsymbol{y}_D, \boldsymbol{\theta}) d\sigma_Z^2 \tag{C.7}$$

The first term in Eq. C.7 is integrated with respect to $\boldsymbol{\beta}$

$$p(y(\boldsymbol{x})|\boldsymbol{y}_D, \sigma_Z, \boldsymbol{\theta}) = \int p(y(\boldsymbol{x})|\boldsymbol{y}_D, \boldsymbol{\beta}, \boldsymbol{\theta}, \sigma_Z^2) p(\boldsymbol{\beta}|\boldsymbol{y}_D, \sigma_Z, \boldsymbol{\theta}) d\boldsymbol{\beta} \tag{C.8}$$

The first term in the integral of Eq. C.8 follows a normal distribution whose parameters (mean m^* and variance s^*) are given by the Simple Kriging equations at any new input $\boldsymbol{x}^*$ as follows

$$m^*(\boldsymbol{x}^*) = \boldsymbol{h}^T \cdot \boldsymbol{\beta} + \boldsymbol{r}(\boldsymbol{x}^*)^T \cdot \boldsymbol{R}_D^{-1} \cdot (\boldsymbol{y}_D - \boldsymbol{H}_D \cdot \boldsymbol{\beta}) \tag{C.9}$$

and

$$s^*(\boldsymbol{x}^*) = \sigma_Z \cdot (1 - \boldsymbol{r}^T \cdot \boldsymbol{R}_D^{-1} \cdot \boldsymbol{r}) \tag{C.10}$$

The second term of Eq. C.8 is evaluated using the Bayes rule from the prior on $\boldsymbol{\beta}$ as follows

$$p(\boldsymbol{\beta}|\boldsymbol{y}_D, \sigma_Z^2, \boldsymbol{\theta}) = \frac{f(\boldsymbol{y}_D|\boldsymbol{\beta}, \sigma_Z^2, \boldsymbol{\theta}) p(\boldsymbol{\beta}|\sigma_Z^2, \boldsymbol{\theta})}{p(\boldsymbol{y}_D|\sigma_Z^2, \boldsymbol{\theta})} \tag{C.11}$$

where the first term at the numerator of Eq. C.11 is the likelihood with respect to the hyper-parameters given as follows

$$f(\boldsymbol{y}_D|\boldsymbol{\beta}, \sigma_Z^2, \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma_Z^2)^{n_0/2}\sqrt{\det \boldsymbol{R}}} \exp\left(-\frac{(\boldsymbol{y}_D - \boldsymbol{H}\boldsymbol{\beta})^T \boldsymbol{R}^{-1} (\boldsymbol{y}_D - \boldsymbol{H}\boldsymbol{\beta})}{2\sigma_Z^2}\right) \tag{C.12}$$

The term at the denominator of Eq. C.11 holds as follows

$$p(\boldsymbol{y}_D|\sigma_Z^2, \boldsymbol{\theta}) = \frac{f(\boldsymbol{y}_D|\boldsymbol{\beta}, \sigma_Z^2, \boldsymbol{\theta}) p(\boldsymbol{\beta}|\sigma_Z^2, \boldsymbol{\theta})}{p(\boldsymbol{\beta}|\boldsymbol{y}_D, \sigma_Z^2, \boldsymbol{\theta})} \tag{C.13}$$

The conditional probability $p(\boldsymbol{\beta}|\boldsymbol{y}_D, \sigma_Z^2, \boldsymbol{\theta})$ at the denominator of Eq. C.13 is tractable and follows a normal distribution whose parameters depend on $\boldsymbol{H}$, $\boldsymbol{R}$, σ_Z (see Appendix B), and on the parameters of the normal prior distribution assigned to $\boldsymbol{\beta}$. Full expression can be found in [Santner et al., 2003].

The second term in the integral of Eq. C.6 is derived from the Bayes formulae

$$p(\boldsymbol{\theta}|\boldsymbol{y}_D) = \frac{p(\boldsymbol{y}_D|\boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\boldsymbol{y}_D)} \tag{C.14}$$

In practice, MCMC methods (see Sect. C.2) are used to evaluate $p(\boldsymbol{\theta}|\boldsymbol{y}_D)$ up to a multiplicative constant so that there is no need to evaluate $p(\boldsymbol{y}_D)$. The conditional probability at the numerator is evaluated using the prior on σ_Z as follows

$$p(\boldsymbol{y}_D|\boldsymbol{\theta}) = \frac{p(\boldsymbol{y}_D|\sigma_Z^2, \boldsymbol{\theta}) p(\sigma_Z^2|\boldsymbol{\theta})}{p(\sigma_Z^2|\boldsymbol{y}_D, \boldsymbol{\theta})} \tag{C.15}$$

where $p(\sigma_Z^2|\boldsymbol{y}_D, \boldsymbol{\theta})$ is tractable and follows a inverse gamma distribution whose parameters depend on $\boldsymbol{H}$, $\boldsymbol{R}$, and on the parameters of the normal prior distribution assigned to $\boldsymbol{\beta}$. Recall that the conditional probability at the numerator is given by Eq. C.13.

C.4 Deriving a full posterior distribution for the sensitivity indices

Recently an in-depth discussion on this problem was provided by [Le Gratiet et al., 2014]. Hereafter, we focus on the procedure developed by [Gramacy and Taddy, 2012]. The basic idea is to carry out an estimation of the Sobol' indices from random realisations of the conditional Gaussian Processes (including the probability on the kriging parameters).

At each iteration of the MCMC-based sampling chain of the kriging posterior distribution (whose parameters correspond to the ones at this MCMC iteration), $N \cdot (n+2)$ random samples, named locations, of the Monte-Carlo algorithm (here the one of [Saltelli, 2002], see also Appendix A) are drawn to compute the sensitivity indices (main and total effects). By allowing random draws of the input locations, the Monte Carlo error of the integral estimates can be included in the posterior variability of the indices and the posterior moments will not be dependent upon any single estimation input set. Using predicted output over this input set, a single realisation of the sensitivity indices is calculated through Saltelli's scheme. At the end of the MCMC, we have a representative sample from the posterior for the main and total indices. Therefore, by conditioning on the predicted output, a posterior sample of the indices can be derived incorporating variability from both the integral estimation (Monte-Carlo error) and uncertainty about the model output (incorporating uncertainty on the kriging parameters).

D Brief introduction to the main uncertainty theories

The present appendix is mainly based on [Dubois and Guyonnet, 2011, Baudrit et al., 2007a, Aven and Zio, 2011]. It aims at providing the basic concepts of the main existing theories for uncertainty treatment. In particular, refer to [Dubois, 2007] for a comprehensive review on how all these theories are linked under a unified vision.

D.1 Probability

Probability is a single-valued measure of uncertainty: uncertainty about the occurrence of a measurable event $A \in \Omega$ (Ω is the referential) is represented by a single number $P(A) \in [0;1]$. Two interpretations of probability exist in the literature.

In the relative frequency interpretation, probability is defined as the relative frequency of occurrence of A in a finite sample. The lack of knowledge about the true value of $P(A)$ is termed epistemic uncertainty. Whereas epistemic uncertainty can be reduced (by extending the size of the sample), the aleatory uncertainty cannot.

In the Bayesian (also named subjective) interpretation, probability is derived by the risk practitioner based on his/her background knowledge. In this vision, the probability of an event A represents the degree of belief of the risk practitioner regarding the occurrence of A: in this sense, the probabilities are named subjective. All subjective probabilities are seen as conditioned on the background knowledge: they are probabilities in the light of current knowledge as outlined by [Lindley, 2006]. The probability can be assigned with reference to either betting or some standard event [Aven and Zio, 2011]. If linked to betting, $P(A)$ is the price at which the risk practitioner is neutral between buying and selling a ticket that is worth one unit of payment (say 1 euro) if the event occurs, and is worthless otherwise [De Finetti and de Finetti, 1977]. If linked to a reference to a standard, the risk practitioner compares his/her uncertainty about

the occurrence of the event A with some standard events, e.g. drawing a favourable ball from an urn that contains $P(A) \cdot 100\%$ favourable balls [Lindley, 2000].

One common approach to risk analysis is to use probability (seen as the expression of epistemic uncertainty based on belief) on the true value of a "relative-frequency-interpreted" probability: this is the probability of the frequency approach [Kaplan and Garrick, 1981]. By taking the expected value of the relative frequency-based probability with respect to the epistemic-based probabilities, both aleatory and epistemic uncertainties can be accounted for following this approach.

D.2 Imprecise probability

This theory is described in [Walley, 1991]. A probabilistic representation is incomplete if a family of probability functions $\mathscr{F}_P$ is used in place of a single distribution P, because the available information is not sufficient to select a single one in $\mathscr{F}_P$. Bounds on the probability of events A can be computed as follows:

$$\begin{aligned} P^*(A) &= \sup(P(A), P \in \mathscr{F}_P) \\ P_*(A) &= \inf(P(A), P \in \mathscr{F}_P) \end{aligned} \tag{D.1}$$

The upper bound $P^*(A)$ can be used to measure the degree of plausibility of A, evaluating to what extent A is not impossible, i.e. there is no reason against the occurrence of A. The lower bound $P_*(A)$ can be used to measure the degree of certainty of A. This is similar to the standard probabilistic framework where the degree of belief in an event equals its frequency of occurrence if the latter is available. Each event A is then assigned an interval $[P_*(A); P^*(A)]$, the width of which is a measure of information lacking, i.e. of the epistemic uncertainty.

D.3 Evidence theory

The evidence theory [Shafer et al., 1976] allows imprecision and variability to be treated separately within a single framework. Contrary to probability theory which assigns probability weights to elements of Ω, this theory assigns weights to any subsets, called focal sets. A typical example, as provided by [Baudrit et al., 2007a], is a measurement device with a systematic error (imprecision) and a random error (variability). From repeated experiments, we can get N samples of intervals of the form $[a_i - \delta; a_i + \delta]_{(i=1,\ldots,N)}$, where a_i is the observed quantity and δ is the systematic error. Each interval can be assigned a probability m_i of observing a_i.

In this manner, we define a mass distribution m_i on intervals. Due to lack of knowledge, this probability can be assigned to any element within the interval $[a_i - \delta; a_i + \delta]$. Lower and upper probabilities can then be derived, respectively the belief and plausibility functions (denoted Bel and Pl) defined from the mass distribution m.

Define,

$$m : \mathscr{P}(\Omega) \longrightarrow [0;1], \sum_{E \in \mathscr{P}(\Omega)} m(E) = 1 \tag{D.2}$$

where Ω is the referential, $\mathscr{P}(\Omega)$ is the power set of Ω (the set of all subsets of Ω), E is a focal element of $\mathscr{P}(\Omega)$, when $m(E) > 0$.

Then,

$$\begin{array}{lcl} Bel(A) & = & \sum_{E, E \subseteq A} m(E) \\ Pl(A) & = & \sum_{E, E \cap A \neq \emptyset} m(E) = 1 - Bel(A^c) \end{array} \tag{D.3}$$

The imprecise evidence that asserts A is gathered by $Bel(A)$. After [Dempster, 1967], it is the minimum amount of probability that can be assigned to A by sharing the probability weights defined by the mass function among single values in the focal sets. On the other hand, the imprecise evidence that does not contradict A is gathered by $Pl(A)$: it is the maximum amount of probability that can be assigned to A in the same fashion.

Note that [Shafer et al., 1976] does not refer to an underlying probability space : $Bel(A)$ is assumed to quantify an agent's belief per se with no reference to a probability. The mathematical tool common to Dempster's upper and lower probabilities and to the Shafer's view is the notion of random disjunctive set [Dubois, 2007]. In this sense, the tools are generally referred to as Dempster-Shafer structures.

D.4 Probability bound analysis

[Ferson, 1996] suggested a combined probability-interval analysis, referred to as a probability bound analysis. A p-box [Ferson et al., 2002] is defined by a pair of cumulative distributions on the real line such that $F^* \leq F^*$, bounding the cumulative distribution of an imprecisely known probability function. A p-box can be seen as a parameterized probability model whose

parameters (like mean and variance) are only known to belong to an interval [Dubois, 2007]. Uncertainty propagation is carried out in the traditional probabilistic way for some parameters, and intervals are used for others.

D.5 Possibility theory

Possibility theory [Dubois and Prade, 1988] is convenient to represent consonant imprecise knowledge. A possibility distribution is a mapping from X to the unit interval such that $\pi(x) = 1$ for some $x \in X$. When some elements in X are considered to be more plausible than others for x, degrees of possibility $\pi(x) \in [0;1]$ can be assigned to $x \in X$, with condition that $\pi(x) = 0$ if $x \notin X$ and $\pi(x) = 0$ for at least one value $x \in X$. From a fuzzy set's perspective, the possibility distribution π can be viewed as determined by the membership function μ of a fuzzy set F. In this vision, $\pi_x(u) = \pi(x = u|F)$ estimates the possibility that the variable x is equal to u, knowing the incomplete state of knowledge "x is F". Then, $\mu(u)$ estimates the degree of compatibility of the precise information $x = u$ with the statement to evaluate "x is F" [Dubois et al., 2000]. Possibility theory provides two evaluations of the likelihood of an event A, respectively the possibility and the necessity measures, denoted Π and N:

$$\Pi(A) = sup_{x \in A}(\pi(x)),\ N(A) = 1 - \Pi(A^c) \tag{D.4}$$

The possibility measure of an event expresses the extent to which this event is plausible, i.e., consistent with a possible state of the world. Necessity measure expresses the certainty of events. Interestingly, the possibility measure can be viewed as an upper bound of a probability degree [Dubois and Prade, 1992]. In a coherent manner, we can define the set of probability measures encoded by a possibility distribution π with upper and lower probabilities induced by N and Π as follows:

$$\mathscr{F}_P = \{P, \forall A \subseteq X, N(A) \leq P(A) \leq \Pi(A)\} \tag{D.5}$$

E Fuzzy Random Variable

The present appendix is based on the description of Fuzzy Random Variable available in [Rohmer and Baudrit, 2011]. In the literature, fuzzy random variables can be interpreted in different ways depending on the context of the study (see [Gil, 2001] for an overview). Here, a fuzzy random variable is interpreted as a possibility distribution over classical random variables (named the second order model, see [Couso et al., 2002, Couso and Sánchez, 2008]). In the following, we briefly introduce the basic notions of fuzzy random variable inducing a second order possibility measure in order to have a better understanding of the fuzzy random variable post processing.

Let us consider the random variable $T = f(X,Y)$ and its associated probability measure, where $f : \mathscr{R}^2 \rightarrow \mathscr{R}$ is a known mapping, X a random variable and Y is another imprecisely known random variable described by a fuzzy set $\tilde{Y}$ associated with the possibility distribution: $\pi_{\tilde{Y}} : \mathscr{R} \rightarrow [0,1]$. These values represent the possibility grade that Y coincides with y. Then, $\tilde{T} = f(X,\tilde{Y})$ is a fuzzy random variable defined by the extension principle:

$$\pi_{\tilde{T}}(t) = sup_{t=f(x,y)} \pi_{\tilde{Y}}(Y) \tag{E.1}$$

Fig. E.1A) displays twenty samples $(\tilde{T}_i)_{i=1,\ldots,20}$ of $\tilde{T}$ derived from a Monte Carlo sampling combined with fuzzy interval analysis, where f corresponds to the addition so that X is assigned a uniform probability distribution on $[0,1]$, while Y is represented by a triangular possibility distribution of core 3 and support $[2,4]$.

The fuzzy random variable $\tilde{T}$ represents the imprecise information about T. Let us define the α-cuts of $\tilde{T}$ such that: $[\tilde{T}_\alpha] = (f(x,y)|y \in [\tilde{T}_\alpha])$. On this basis, $T \in [\tilde{T}_\alpha]$ with a confidence level

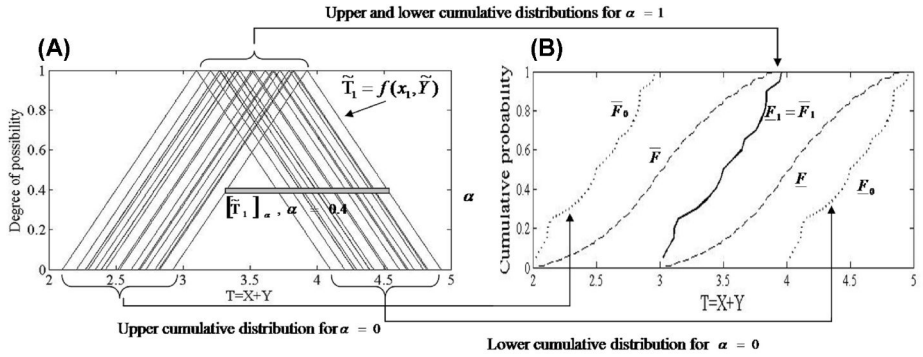

Figure E.1: a) A twenty sampling of the fuzzy random variable $T = X + Y$ where X is uniformly distributed on $[0,1]$ and Y is represented by a possibility distribution of core 3 and support $[2,4]$; b) First-order versus second-order model induced from a fuzzy random variable sampling (adapted from [Rohmer and Baudrit, 2011]).

$1-\alpha$.

For sake of clarity, let us assume that X takes a finite number of different values $x_1,...,x_m$ with respective probabilities $p_1,...,p_m$. In the discrete case, we can define a lower and upper probability bound $[\underline{P}_\alpha^T, \overline{P}_\alpha^T]$ of P^T with a confidence level $1-\alpha$ such that:

$$\underline{P}_\alpha^T(A) = \sum_{i,[\tilde{T}_i]_\alpha \subseteq A} p_i \leq P(T \in A) \leq \sum_{i,[\tilde{T}_i]_\alpha \cap A \neq \emptyset} p_i = \overline{P}_\alpha^T(A) \tag{E.2}$$

where $p_i = 1/20$ in the given example. For each α, $\underline{P}_\alpha^T$ gathers the imprecise evidence that asserts the statement "$T \in A$". On the other hand, $\overline{P}_\alpha^T$ gathers the imprecise evidence that does not contradict the same statement. For instance, Fig. E.1B) depicts the lower and upper cumulative probability bounds $[\underline{F}_\alpha, \overline{F}_\alpha]$ for $\alpha = 1$ and $\alpha = 0$ resulting from the 20 samples by means of the following expression:

$$\begin{aligned} \underline{F}_\alpha(x) &= 1/20 \cdot Card(i | [\tilde{T}_i]_\alpha \subseteq (-\infty, x])) \\ \overline{F}_\alpha(x) &= 1/20 \cdot Card(i | [\tilde{T}_i]_\alpha \cap (-\infty, x] \neq \emptyset)) \end{aligned} \tag{E.3}$$

where "Card" corresponds to the cardinality operator (i.e. size) of the considered set, ∞ is the infinite bound and $\emptyset$ is the null set.

In the continuous case, the same principle can be followed [Smets, 2005]. A second-order possibility distribution is then defined, over a set of probability measures by means of α-cuts

$[\underline{P}_\alpha^T, \overline{P}_\alpha^T]$. In this framework, the probability of a given event A is not precise meaning that it can not be represented by a crisp value, but rather by a range of possible values. The described formal concepts let us state for instance that "the probability that the true probability of the event A is 0.5 ranges between 0.4 and 0.7". [Couso et al., 2002] showed that the following interval (Eq. E.4) provides the smallest envelope of the "true" probability of A given the available information.

$$[\underline{P}^T(A), \overline{P}^T(A)] = [\int_0^1 \underline{P}_\alpha^T(A)d\alpha, \int_0^1 \overline{P}_\alpha^T(A)d\alpha] \tag{E.4}$$

This means that the measure $\underline{P}^T(A)$ (resp. $\overline{P}^T(A)$) corresponds to the greatest lower bound (respectively smallest upper bound) that we give to the probability of A.

Bibliography

[Abdallah, 2009] Abdallah, M. (2009). *Vulnérabilité des ouvrages en maçonnerie à des mouvements de terrain : méthodologie d'analyse par méthodes statistiques et par plans d'expériences numériques sur les données de la ville de Joeuf*. PhD thesis, Université de Lorraine.

[Abrahamson, 2000] Abrahamson, N. A. (2000). State of the practice of seismic hazard assessment. In *GeoEng 2000*, volume 1, pages 659–685, Melbourne, Australia.

[Allaire and Willcox, 2012] Allaire, D. L. and Willcox, K. E. (2012). A variance-based sensitivity index function for factor prioritization. *Reliability Engineering & System Safety*, 107:107–114.

[Alvarez, 2009] Alvarez, D. A. (2009). Reduction of uncertainty using sensitivity analysis methods for infinite random sets of indexable type. *International journal of approximate reasoning*, 50:750–762.

[Antoniadis et al., 2012] Antoniadis, A., Helbert, C., Prieur, C., and Viry, L. (2012). Spatio-temporal metamodeling for west african monsoon. *Environmetrics*, 23:24–36.

[Archer et al., 1997] Archer, G., Saltelli, A., and Sobol, I. (1997). Sensitivity measures, anova-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation*, 58:99–120.

[Ascough(II) et al., 2008] Ascough(II), J. C., Maier, H. R., Ravalico, J. K., and Strudley, M. W. (2008). Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecological Modelling*, 219:383–399.

[Aubry et al., 1986] Aubry, D., Chouvet, D., Modaressi, A., and Modaressi, H. (1986). Gefdyn software - logiciel d'analyse du comportement mécanique des sols par éléments finis avec prise en compte du couplage sol-eau-air (gefdyn software, finite element analysis of soil mechanical behaviour taking into account the soil-water-air coupling. Technical report, Ecole Centrale Paris.

[Aubry et al., 1982] Aubry, D., Hujeux, J.-C., Lassoudière, F., and Meimon, Y. (1982). A double memory model with multiple mechanisms for cyclic soil behaviour. In *International symposium on numerical models*, page 3–13, Balkema.

[Auder et al., 2012] Auder, B., Crecy, A. D., Iooss, B., and Marquès, M. (2012). Screening and metamodeling of computer experiments with functional outputs. application to thermal–hydraulic computations. *Reliab. Eng. Syst. Safety*, 107:122–131.

[Auder and Iooss, 2008] Auder, B. and Iooss, B. (2008). Global sensitivity analysis based on entropy. In *ESREL 2008 Conference, ESRA*, page 3–13, Valencia, Spain.

[Aven and Renn, 2009] Aven, T. and Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *Journal of Risk Research*, 12:1–11.

[Aven and Zio, 2011] Aven, T. and Zio, E. (2011). Some considerations on the treatment of uncertainties in risk assessment for practical decision-making. *Reliability Engineering and System Safety*, 96:64–74.

[Baecher and Christian, 2005] Baecher, G. B. and Christian, J. T. (2005). *Reliability and statistics in geotechnical engineering*. John Wiley & Sons.

[Bailles et al., 2013] Bailles, A., Vandromme, R., Desramaut, N., Sedan-Miegemolle, O., and Grandjean, G. (2013). Changing patterns in climate-driven landslide hazard: an alpine test site. In *Landslide Science and Practice*, pages 93–98. Springer.

[Baudrit, 2005] Baudrit, C. (2005). *Représentation et propagation de connaissances imprécises et incertaines: Application à l'évaluation des risques liés aux sites et sols pollués*. PhD thesis, Université Paul Sabatier-Toulouse III.

[Baudrit et al., 2007a] Baudrit, C., Couso, I., and Dubois, D. (2007a). Joint propagation of probability and possibility in risk analysis: Towards a formal framework. *International Journal of Approximate Reasoning*, 45(1):82–105.

[Baudrit and Dubois, 2006] Baudrit, C. and Dubois, D. (2006). Practical representations of incomplete probabilistic knowledge. *Computational statistics & data analysis*, 51(1):86–108.

[Baudrit et al., 2006] Baudrit, C., Dubois, D., and Guyonnet, D. (2006). Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment. *Fuzzy Systems, IEEE Transactions on*, 14(5):593–608.

[Baudrit et al., 2008] Baudrit, C., Dubois, D., and Perrot, N. (2008). Representing parametric probabilistic models tainted with imprecision. *Fuzzy sets and systems*, 159(15):1913–1928.

[Baudrit et al., 2007b] Baudrit, C., Guyonnet, D., and Dubois, D. (2007b). Joint propagation of variability and imprecision in assessing the risk of groundwater contamination. *Journal of contaminant hydrology*, 93(1):72–84.

[Beer et al., 2013] Beer, M., Ferson, S., and Kreinovich, V. (2013). Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing*, 37(1):4–29.

[Bellman et al., 1961] Bellman, R., Bellman, R. E., Bellman, R. E., and Bellman, R. E. (1961). *Adaptive control processes: a guided tour*, volume 4. Princeton University Press Princeton.

[Berger et al., 2001] Berger, J. O., De Oliveira, V., and Sansó, B. (2001). Objective bayesian analysis of spatially correlated data. *Journal of the American Statistical Association*, 96(456):1361–1374.

[Bernardie et al., 2006] Bernardie, S., Delpont, G., Nédellec, J., Negulescu, C., and Roullé, A. (2006). Microzonage de lourdes. Technical report, BRGM RP/53846-FR.

[Bertil et al., 2009] Bertil, D., Bengoubou-Valérius, M., Péricat, J., and Auclair, S. (2009). Scénarios départementaux de risque sismique en guadeloupe. Technical report, BRGM RP-57488-FR.

[Biarez et al., 1994] Biarez, J., Hicher, P.-Y., et al. (1994). *Elementary mechanics of soil behaviour: saturated remoulded soils*. AA Balkema.

[Bieri, 2006] Bieri, S. (2006). Disaster risk management and the systems approach. Technical report, World institute for disaster management.

[Bolado-Lavin et al., 2009] Bolado-Lavin, R., Castaings, W., and Tarantola, S. (2009). Contribution to the sample mean plot for graphical and numerical sensitivity analysis. *Reliability Engineering & System Safety*, 94(6):1041–1049.

[Bommer et al., 2006] Bommer, J., Spence, R., and Pinho, R. (2006). Earthquake loss estimation models: time to open the black boxes. In *Proceedings of the first European conference on earthquake engineering and seismology. Geneva, Switzerland*.

[Bommer et al., 2010] Bommer, J. J., Douglas, J., Scherbaum, F., Cotton, F., Bungum, H., and Fäh, D. (2010). On the selection of ground-motion prediction equations for seismic hazard analysis. *Seismological Research Letters*, 81(5):783–793.

[Borgonovo, 2007] Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784.

[Bouc et al., 2009] Bouc, O., Audigane, P., Bellenfant, G., Fabriol, H., Gastine, M., Rohmer, J., and Seyedi, D. (2009). Determining safety criteria for co< sub> 2</sub> geological storage. *Energy procedia*, 1(1):2439–2446.

[Bouchut and Vincent, 2002] Bouchut, J. and Vincent, M. (2002). Recensement et analyse des mouvements de terrain survenus en 2000–2001 dans les arrondissements de montdidier (somme) et de clermont (oise).

[Boulahya et al., 2007] Boulahya, F., Dubus, I., Dupros, F., and Lombard, P. (2007). Footprint@ work, a computing framework for large scale parametric simulations: application to pesticide risk assessment and management. In *Forum EGEE Enabling Grids for E-sciencE, Manchester, UK*, page 160.

[Brown, 1979] Brown, C. B. (1979). A fuzzy safety measure. *Journal of the Engineering Mechanics Division*, 105(5):855–872.

[Busby et al., 2007] Busby, D., Romary, T., Touzani, S., Feraille, M., Noetinger, B., and Hu, L. (2007). Reservoir forecasting under uncertainty: an integrated approach. In *International Meeting on Complexity in Oil Industry, 5-9 August 2007, Natal, Brazil*, page 160.

[Byrd et al., 1995] Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208.

[Cabantous et al., 2011] Cabantous, L., Hilton, D., Kunreuther, H., and Michel-Kerjan, E. (2011). Is imprecise knowledge better than conflicting expertise? evidence from insurers' decisions in the united states. *Journal of Risk and Uncertainty*, 42(3):211–232.

[Caers and Zhang, 2004] Caers, J. and Zhang, T. (2004). Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models.

[Campbell et al., 2006] Campbell, K., McKay, M. D., and Williams, B. J. (2006). Sensitivity analysis when model outputs are functions. *Reliability Engineering & System Safety*, 91(10):1468–1472.

[Campolongo et al., 2011] Campolongo, F., Saltelli, A., and Cariboni, J. (2011). From screening to quantitative sensitivity analysis. a unified approach. *Computer Physics Communications*, 182(4):978–988.

[Carter et al., 2006] Carter, J. N., Ballester, P. J., Tavassoli, Z., and King, P. R. (2006). Our calibrated model has poor predictive value: An example from the petroleum industry. *Reliability Engineering & System Safety*, 91(10):1373–1381.

[Cauvin et al., 2008] Cauvin, M., Salmon, R., Verdel, T., et al. (2008). Dealing with uncertainties in the context of post mining hazard evaluation. In *Post-Mining 2008. Symposium Proceedings*.

[Collins and Sitar, 2010] Collins, B. D. and Sitar, N. (2010). Stability of steep slopes in cemented sands. *Journal of geotechnical and geoenvironmental engineering*, 137(1):43–51.

[Couso et al., 2002] Couso, I., Montes, S., and Gil, P. (2002). Second order possibility measure induced by a fuzzy random variable. In *Statistical modeling, analysis and management of fuzzy data*, pages 127–144. Springer.

[Couso and Sánchez, 2008] Couso, I. and Sánchez, L. (2008). Higher order models for fuzzy random variables. *Fuzzy Sets and Systems*, 159(3):237–258.

[Cox, 1994] Cox, E. (1994). *The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems*. Academic Press Professional, Inc., San Diego, CA, USA.

[Crosetto and Tarantola, 2001] Crosetto, M. and Tarantola, S. (2001). Uncertainty and sensitivity analysis: tools for gis-based model implementation. *International Journal of Geographical Information Science*, 15(5):415–437.

[Crowley et al., 2005] Crowley, H., Bommer, J. J., Pinho, R., and Bird, J. (2005). The impact of epistemic uncertainty on an earthquake loss model. *Earthquake engineering & structural dynamics*, 34(14):1653–1685.

[De Cooman and Aeyels, 1999] De Cooman, G. and Aeyels, D. (1999). Supremum preserving upper probabilities. *Information Sciences*, 118(1):173–212.

[De Finetti and de Finetti, 1977] De Finetti, B. and de Finetti, B. (1977). Theory of probability, volume i. *Bull. Amer. Math. Soc. 83 (1977), 94-97 DOI: http://dx. doi. org/10.1090/S0002-9904-1977-14188-8 PII*, pages 0002–9904.

[de Rocquigny et al., 2008] de Rocquigny, E., Devictor, N., and Tarantola, S. (2008). *Uncertainty in industrial practice: a guide to quantitative uncertainty management*. John Wiley & Sons.

[Deck and Verdel, 2012] Deck, O. and Verdel, T. (2012). Uncertainties and risk analysis related to geohazards: From practical applications to research trends. *Risk management for the future—theory and cases. Intech—www. intechopen. com*.

[Dempster, 1967] Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pages 325–339.

[Dewez et al., 2013] Dewez, T. J., Rohmer, J., Regard, V., Cnudde, C., et al. (2013). Probabilistic coastal cliff collapse hazard from repeated terrestrial laser surveys: case study from mesnil val (normandy, northern france). *Journal of Coastal Research*, 65:702–707.

[Ditlevsen, 1994] Ditlevsen, O. (1994). Distribution arbitrariness in structural reliability. *Structural Safety and Reliability*, pages 1241–1247.

[Drucker and Prager, 1952] Drucker, D. C. and Prager, W. (1952). Soil mechanics and plastic analysis for limit design. *Quarterly of Applied Mathematics*, 10(2):157–165.

[Dubois, 2006] Dubois, D. (2006). Possibility theory and statistical reasoning. *Computational statistics & data analysis*, 51(1):47–69.

[Dubois, 2007] Dubois, D. (2007). Uncertainty theories: a unified view. In *IEEE Cybernetic Systems Conference, Dublin, Ireland, Invited Paper*, pages 4–9.

[Dubois, 2010] Dubois, D. (2010). The role of epistemic uncertainty in risk analysis. In *Scalable Uncertainty Management*, pages 11–15. Springer.

[Dubois et al., 2001] Dubois, D., Esteva, F., Godo, L., and Prade, H. (2001). An information-based discussion of vagueness. In *Fuzzy Systems, 2001. The 10th IEEE International Conference on*, volume 2, pages 781–784. IEEE.

[Dubois and Guyonnet, 2011] Dubois, D. and Guyonnet, D. (2011). Risk-informed decision-making in the presence of epistemic uncertainty. *International Journal of General Systems*, 40(02):145–167.

[Dubois et al., 2000] Dubois, D., Nguyen, H. T., and Prade, H. (2000). Possibility theory, probability and fuzzy sets misunderstandings, bridges and gaps. In *Fundamentals of fuzzy sets*, pages 343–438. Springer.

[Dubois and Prade, 1988] Dubois, D. and Prade, H. (1988). *Possibility theory*. Springer.

[Dubois and Prade, 1992] Dubois, D. and Prade, H. (1992). When upper probabilities are possibility measures. *Fuzzy sets and systems*, 49(1):65–74.

[Dubois and Prade, 1994] Dubois, D. and Prade, H. (1994). Possibility theory and data fusion in poorly informed environments. *Control Engineering Practice*, 2(5):811–823.

[Eidsvig et al., 2014] Eidsvig, U., Papathoma-Koehle, M., Du, J., Glade, T., and Vangelsten, B. (2014). Quantification of model uncertainty in debris flow vulnerability assessment. *Engineering Geology*.

[Einstein and Baecher, 1983] Einstein, H. H. and Baecher, G. B. (1983). Probabilistic and statistical methods in engineering geology. *Rock mechanics and rock engineering*, 16(1):39–72.

[El-Ramly et al., 2002] El-Ramly, H., Morgenstern, N., and Cruden, D. (2002). Probabilistic slope stability analysis for practice. *Canadian Geotechnical Journal*, 39(3):665–683.

[El-Shayeb, 1999] El-Shayeb, Y. (1999). Apport de la logique floue à l'évaluation de l'aléa mouvement de terrain des sites géotechniques: propositions pour une méthodologie générale. *Institut National Polytechnique de Lorraine. École des mines de Nancy*.

[Eldred et al., 2011] Eldred, M., Swiler, L. P., and Tang, G. (2011). Mixed aleatory-epistemic uncertainty quantification with stochastic expansions and optimization-based interval estimation. *Reliability Engineering & System Safety*, 96(9):1092–1113.

[Ellingwood and Kinali, 2009] Ellingwood, B. R. and Kinali, K. (2009). Quantifying and communicating uncertainty in seismic risk assessment. *Structural Safety*, 31(2):179–187.

[Ellsberg, 1961] Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669.

[Environmental Protection Agency, 2009] Environmental Protection Agency, E. (2009). Guidance on the development, evaluation, and application of environmental models. Technical Report EPA/100/K-09/003, Office of the Science Advisor, Council for Regulatory Environmental Modeling.

[Ercanoglu and Gokceoglu, 2004] Ercanoglu, M. and Gokceoglu, C. (2004). Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (west black sea region, turkey). *Engineering Geology*, 75(3):229–250.

[Esterhuizen and Ellenberger, 2007] Esterhuizen, G. S. and Ellenberger, J. L. (2007). Effects of weak bands on pillar stability in stone mines: field observations and numerical model assessment. In *Proceedings of the 26th International Conference on Ground Control in Mining, Morgantown, WV*, pages 336–342.

[European Commission, 2009] European Commission, E. (2009). Impact assessment guidelines. Technical report, European Commission.

[Fellin, 2005] Fellin, W. (2005). Ambiguity of safety definition in geotechnical models. In *Analyzing uncertainty in civil engineering*, pages 17–31. Springer.

[Ferson, 1996] Ferson, S. (1996). What monte carlo methods cannot do. *Human and Ecological Risk Assessment*, 2(4):990–1007.

[Ferson and Ginzburg, 1996] Ferson, S. and Ginzburg, L. R. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2):133–144.

[Ferson et al., 2002] Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D. S., and Sentz, K. (2002). *Constructing probability boxes and Dempster-Shafer structures*, volume 835. Sandia National Laboratories.

[Ferson and Troy Tucker, 2006] Ferson, S. and Troy Tucker, W. (2006). Sensitivity analysis using probability bounding. *Reliability Engineering & System Safety*, 91(10):1435–1442.

[Forrester et al., 2008] Forrester, A., Sobester, A., and Keane, A. (2008). *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons.

[Friedman and Stuetzle, 1981] Friedman, J. H. and Stuetzle, W. (1981). Projection pursuit regression. *Journal of the American statistical Association*, 76(376):817–823.

[Gelman et al., 2013] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.

[Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741.

[Ghanem and Spanos, 1991] Ghanem, R. G. and Spanos, P. D. (1991). *Stochastic finite elements: a spectral approach*, volume 387974563. Springer.

[Gil, 2001] Gil, M. A. (2001). Fuzzy random variables. *Information Sciences*, 133(1):1–2.

[Ginsbourger et al., 2009] Ginsbourger, D., Dupuy, D., Badea, A., Carraro, L., and Roustant, O. (2009). A note on the choice and the estimation of kriging models for the analysis of deterministic computer experiments. *Applied Stochastic Models in Business and Industry*, 25(2):115–131.

[Giovinazzi and Lagomarsino, 2004] Giovinazzi, S. and Lagomarsino, S. (2004). A macroseismic method for the vulnerability assessment of buildings. In *13th World Conference on Earthquake Engineering, Vancouver, BC, Canada*, pages 1–6.

[Gorsevski et al., 2006] Gorsevski, P. V., Gessler, P. E., Boll, J., Elliot, W. J., and Foltz, R. B. (2006). Spatially and temporally distributed modeling of landslide susceptibility. *Geomorphology*, 80(3):178–198.

[Gramacy, 2005] Gramacy, R. B. (2005). *Bayesian treed Gaussian process models*. PhD thesis, UNIVERSITY OF CALIFORNIA SANTA CRUZ.

[Gramacy, 2007] Gramacy, R. B. (2007). tgp: an r package for bayesian nonstationary, semiparametric nonlinear regression and design by treed gaussian process models. *Journal of Statistical Software*, 19(9):6.

[Gramacy and Lee, 2009] Gramacy, R. B. and Lee, H. K. (2009). Adaptive design and analysis of supercomputer experiments. *Technometrics*, 51(2):130–145.

[Gramacy and Lee, 2012] Gramacy, R. B. and Lee, H. K. (2012). Cases for the nugget in modeling computer experiments. *Statistics and Computing*, 22(3):713–722.

[Gramacy and Taddy, 2012] Gramacy, R. B. and Taddy, M. (2012). Categorical inputs, sensitivity analysis, optimization and importance tempering with tgp version 2, an r package for treed gaussian process models. *Journal of Statistical Software*, 33(i06).

[Grünthal, 1998] Grünthal, G. (1998). Cahiers du centre européen de géodynamique et de séismologie: Volume 15–european macroseismic scale 1998. *European Center for Geodynamics and Seismology, Luxembourg*.

[Guo and Du, 2007] Guo, J. and Du, X. (2007). Sensitivity analysis with mixture of epistemic and aleatory uncertainties. *AIAA journal*, 45(9):2337–2349.

[Gzyl, 1995] Gzyl, H. (1995). *The method of maximum entropy*. World Scientific Singapore.

[Hall, 2006] Hall, J. W. (2006). Uncertainty-based sensitivity indices for imprecise probability distributions. *Reliability Engineering & System Safety*, 91(10):1443–1451.

[Hamm et al., 2006] Hamm, N., Hall, J., and Anderson, M. (2006). Variance-based sensitivity analysis of the probability of hydrologically induced slope instability. *Computers & geosciences*, 32(6):803–817.

[Handcock and Stein, 1993] Handcock, M. S. and Stein, M. L. (1993). A bayesian analysis of kriging. *Technometrics*, 35(4):403–410.

[Hansen, 1984] Hansen, A. (1984). Landslide hazard analysis. In Brunsden, D., P. D., editor, *Slope Instability*, pages 523–602. Wiley and Sons, New York.

[Hastie et al., 2009] Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., and Tibshirani, R. (2009). *The elements of statistical learning*, volume 2. Springer.

[Hastings, 1970] Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109.

[Helbert et al., 2009] Helbert, C., Dupuy, D., and Carraro, L. (2009). Assessment of uncertainty in computer experiments from universal to bayesian kriging. *Applied Stochastic Models in Business and Industry*, 25(2):99–113.

[Helton, 1994] Helton, J. C. (1994). Treatment of uncertainty in performance assessments for complex systems. *Risk analysis*, 14(4):483–511.

[Helton et al., 2006a] Helton, J. C., Johnson, J. D., Oberkampf, W., and Sallaberry, C. J. (2006a). Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty. *Reliability Engineering & System Safety*, 91(10):1414–1434.

[Helton et al., 2006b] Helton, J. C., Johnson, J. D., Sallaberry, C. J., and Storlie, C. B. (2006b). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 91(10):1175–1209.

[Helton and Oberkampf, 2004] Helton, J. C. and Oberkampf, W. (2004). Alternative representations of epistemic uncertainty. *Reliability Engineering & System Safety*, 85(1):1–10.

[Hemez and Atamturktur, 2011] Hemez, F. M. and Atamturktur, S. (2011). The dangers of sparse sampling for the quantification of margin and uncertainty. *Reliability Engineering & System Safety*, 96(9):1220–1231.

[Higdon et al., 2008] Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008). Computer model calibration using high-dimensional output. *Journal of the American Statistical Association*, 103(482).

[Hill et al., 2013] Hill, L., Sparks, R., and Rougier, J. (2013). Risk assessment and uncertainty in natural hazards. *Risk and uncertainty assessment for natural hazards, edited by: Rougier, JC, Sparks, RS J., and Hill, LJ*, pages 1–18.

[Hoeting et al., 1999] Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical science*, pages 382–401.

[Holden, 2011] Holden, C. (2011). Kinematic source model of the 22 february 2011 mw 6.2 christchurch earthquake using strong motion data. *Seismological research letters*, 82(6):783–788.

[Homma and Saltelli, 1996] Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17.

[Hujeux, 1985] Hujeux, J. (1985). Une loi de comportement pour le chargement cyclique des sols. *Génie parasismique*, pages 278–302.

[Igusa et al., 2002] Igusa, T., Buonopane, S., and Ellingwood, B. (2002). Bayesian analysis of uncertainty for structural engineering applications. *Structural Safety*, 24(2):165–186.

[Iooss, 2011] Iooss, B. (2011). Revue sur l'analyse de sensibilité globale de modèles numériques. *Journal de la Société Française de Statistique*, 152(1):3–25.

[Jacques et al., 2006] Jacques, J., Lavergne, C., and Devictor, N. (2006). Sensitivity analysis in presence of model uncertainty and correlated inputs. *Reliability Engineering & System Safety*, 91(10):1126–1134.

[Janon et al., 2014] Janon, A., Nodet, M., and Prieur, C. (2014). Uncertainties assessment in global sensitivity indices estimation from metamodels. *International Journal for Uncertainty Quantification*, 4(1).

[Jolliffe, 2005] Jolliffe, I. (2005). *Principal component analysis*. Wiley Online Library.

[Jones et al., 1998] Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492.

[Kaplan and Garrick, 1981] Kaplan, S. and Garrick, B. J. (1981). On the quantitative definition of risk. *Risk analysis*, 1(1):11–27.

[Karimi and Hüllermeier, 2007] Karimi, I. and Hüllermeier, E. (2007). Risk assessment system of natural hazards: A new approach based on fuzzy probability. *Fuzzy Sets and Systems*, 158(9):987–999.

[Kennedy and O'Hagan, 2001] Kennedy, M. C. and O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464.

[Keynes, 1921] Keynes, J. M. (1921). A treatise on probability.

[Kiureghian and Ditlevsen, 2009] Kiureghian, A. D. and Ditlevsen, O. (2009). Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112.

[Kleijnen, 2005] Kleijnen, J. P. (2005). An overview of the design and analysis of simulation experiments for sensitivity analysis. *European Journal of Operational Research*, 164(2):287–300.

[Klir, 1989] Klir, G. J. (1989). Is there more to uncertainty than some probability theorists might have us believe?*. *International Journal Of General System*, 15(4):347–378.

[Klir, 1994] Klir, G. J. (1994). On the alleged superiority of probabilistic representation of uncertainty. *IEEE Transactions on Fuzzy Systems*, 2(1):27–31.

[Klir and Wierman, 1999] Klir, G. J. and Wierman, M. J. (1999). *Uncertainty-based information: elements of generalized information theory*, volume 15. Springer.

[Knight, 1921] Knight, F. H. (1921). Risk, uncertainty and profit. *New York: Hart, Schaffner and Marx*.

[Koehler and Owen, 1996] Koehler, J. and Owen, A. (1996). Computer experiments. *Handbook of statistics*, 13(13):261–308.

[Konikow and Bredehoeft, 1992] Konikow, L. F. and Bredehoeft, J. D. (1992). Ground-water models cannot be validated. *Advances in water resources*, 15(1):75–83.

[Krige, 1951] Krige, D. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of Chemical, Metallurgical, and Mining Society of South Africa*.

[Kucherenko et al., 2011] Kucherenko, S., Feil, B., Shah, N., and Mauntz, W. (2011). The identification of model effective dimensions using global sensitivity analysis. *Reliability Engineering & System Safety*, 96(4):440–449.

[Lagomarsino and Giovinazzi, 2006] Lagomarsino, S. and Giovinazzi, S. (2006). Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings. *Bulletin of Earthquake Engineering*, 4(4):415–443.

[Laloui et al., 2004] Laloui, L., Tacher, L., Moreni, M., and Bonnard, C. (2004). Hydro-mechanical modeling of crises of large landslides: application to the la frasse landslide. In *9th International Symp. on Landslides*, number LMS-CONF-2004-008, pages 1103–1110. Balkema.

[Lamboni et al., 2011] Lamboni, M., Monod, H., and Makowski, D. (2011). Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliability Engineering & System Safety*, 96(4):450–459.

[Langewisch and Apostolakis, 2010] Langewisch, D. and Apostolakis, G. (2010). A comparison of polynomial response surfaces and gaussian processes as metamodels for uncertainty analysis with long-running computer codes. In *the 10th International Probabilistic Safety Assessment & Management Conference, Seattle, WA, USA*.

[Le Gratiet, 2013] Le Gratiet, L. (2013). *Multi-fidelity Gaussian process regression for computer experiments*. PhD thesis, Université Paris-Diderot-Paris VII.

[Le Gratiet et al., 2014] Le Gratiet, L., Cannamela, C., and Iooss, B. (2014). A bayesian approach for global sensitivity analysis of (multifidelity) computer codes. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):336–363.

[Li et al., 2014] Li, G., Lu, Z., Lu, Z., and Xu, J. (2014). Regional sensitivity analysis of aleatory and epistemic uncertainties on failure probability. *Mechanical Systems and Signal Processing*, 46(2):209–226.

[Li and Lu, 2013] Li, L. and Lu, Z. (2013). Regional importance effect analysis of the input variables on failure probability. *Computers & Structures*, 125:74–85.

[Lilburne and Tarantola, 2009] Lilburne, L. and Tarantola, S. (2009). Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, 23(2):151–168.

[Lindley, 2006] Lindley, D. (2006). Understanding uncertainty? john wiley and sons. *New Jersey*.

[Lindley, 2000] Lindley, D. V. (2000). The philosophy of statistics. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(3):293–337.

[Lockwood et al., 2012] Lockwood, B., Anitescu, M., and Mavriplis, D. J. (2012). Mixed aleatory/epistemic uncertainty quantification for hypersonic flows via gradient-based optimization and surrogate models. In *the 50th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Nashville, Tennessee*, pages 9–12.

[Lopez-Caballero and Modaressi Farahmand-Razavi, 2008] Lopez-Caballero, F. and Modaressi Farahmand-Razavi, A. (2008). Numerical simulation of liquefaction effects on seismic ssi. *Soil Dynamics and Earthquake Engineering*, 28(2):85–98.

[Lopez-Caballero et al., 2007] Lopez-Caballero, F., Razavi, A. M.-F., and Modaressi, H. (2007). Nonlinear numerical method for earthquake site response analysis i—elastoplastic cyclic model and parameter identification strategy. *Bulletin of Earthquake Engineering*, 5(3):303–323.

[Lophaven et al., 2002] Lophaven, S. N., Nielsen, H. B., and Søndergaard, J. (2002). Dace-a matlab kriging toolbox, version 2.0. Technical report.

[Loschetter et al., 2015] Loschetter, A., Rohmer, J., de Lary, L., and Manceau, J. (2015). Dealing with uncertainty in risk assessments in early stages of a co2 geological storage project: comparison of pure-probabilistic and fuzzy-probabilistic frameworks. *Stochastic Environmental Research and Risk Assessment*, pages 1–17.

[Ma and Zabaras, 2011] Ma, X. and Zabaras, N. (2011). Kernel principal component analysis for stochastic input model generation. *Journal of Computational Physics*, 230(19):7311–7331.

[Marrel et al., 2009] Marrel, A., Iooss, B., Laurent, B., and Roustant, O. (2009). Calculations of sobol indices for the gaussian process metamodel. *Reliability Engineering & System Safety*, 94(3):742–751.

[Marrel et al., 2008] Marrel, A., Iooss, B., Van Dorpe, F., and Volkova, E. (2008). An efficient methodology for modeling complex computer codes with gaussian processes. *Computational Statistics & Data Analysis*, 52(10):4731–4744.

[Marzocchi et al., 2004] Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C., and Boschi, E. (2004). Quantifying probabilities of volcanic events: the example of volcanic hazard at mount vesuvius. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 109(B11).

[Masson and Denœux, 2006] Masson, M.-H. and Denœux, T. (2006). Inferring a possibility distribution from empirical data. *Fuzzy sets and systems*, 157(3):319–340.

[MATE, 1999] MATE (1999). (ministere de l'aménagement du territoire et de l'environnement). plans de prévention des risques naturels (ppr) — risques de mouvements de terrains. guide méthodologique. Technical report, La Documentation Francaise, Paris.

[Matti, 2008] Matti, B. (2008). *Geological heterogeneity in landslides: characterization and flow modelling*. PhD thesis, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE.

[McKay et al., 1979] McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

[Metropolis et al., 1953] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.

[Mishra, 2002] Mishra, S. (2002). Assigning probability distributions to input parameters of performance assessment models. Technical report, Swedish Nuclear Fuel and Waste Management Co., Stockholm (Sweden).

[Modaressi et al., 1997] Modaressi, H., Foerster, E., and Mellal, A. (1997). Computer-aided seismic analysis of soils. In *Proc. Of th 6th Int. Symp. On Numerical models in Geomechanics, NUMOG VI, Montréal, Québec, CCanada July*, pages 2–4.

[Morio, 2011] Morio, J. (2011). Influence of input pdf parameters of a model on a failure probability estimation. *Simulation Modelling Practice and Theory*, 19(10):2244–2255.

[Morris, 1991] Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174.

[Nachbaur and Rohmer, 2011] Nachbaur, A. and Rohmer, J. (2011). Managing expert-information uncertainties for assessing collapse susceptibility of abandoned underground structures. *Engineering Geology*, 123(3):166–178.

[Neal, 1997] Neal, R. M. (1997). Monte carlo implementation of gaussian process models for bayesian regression and classification. *Technical Report No. 9702, Department of Statistics, University of Toronto*.

[Nilsen, 2000] Nilsen, B. (2000). New trends in rock slope stability analyses. *Bulletin of Engineering Geology and the Environment*, 58(3):173–178.

[Nilsen and Aven, 2003] Nilsen, T. and Aven, T. (2003). Models and model uncertainty in the context of risk analysis. *Reliability Engineering & System Safety*, 79(3):309–317.

[Oakley and O'Hagan, 2004] Oakley, J. E. and O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: a bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):751–769.

[Oberguggenberger and Fellin, 2002] Oberguggenberger, M. and Fellin, W. (2002). From probability to fuzzy sets: the struggle for meaning in geotechnical risk assessment. In *Conference Report*, pages 1–10.

[Oberguggenberger et al., 2009] Oberguggenberger, M., King, J., and Schmelzer, B. (2009). Classical and imprecise probability methods for sensitivity analysis in engineering: A case study. *International Journal of Approximate Reasoning*, 50(4):680–693.

[O'Hagan, 2006] O'Hagan, A. (2006). Bayesian analysis of computer code outputs: a tutorial. *Reliability Engineering & System Safety*, 91(10):1290–1300.

[O'Hagan and Oakley, 2004] O'Hagan, A. and Oakley, J. E. (2004). Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering & System Safety*, 85(1):239–248.

[Oliver and Chen, 2011] Oliver, D. S. and Chen, Y. (2011). Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1):185–221.

[Olivier et al., 2013] Olivier, M., Sedan, O., and Monod, B. (2013). Contribution of physical modelling to landslide hazard mapping: case of the french basque coast. In *Landslide Science and Practice*, pages 109–118. Springer.

[Papadrakakis and Lagaros, 2002] Papadrakakis, M. and Lagaros, N. D. (2002). Reliability-based structural optimization using neural networks and monte carlo simulation. *Computer methods in applied mechanics and engineering*, 191(32):3491–3507.

[Pappenberger and Beven, 2006] Pappenberger, F. and Beven, K. J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, 42(5).

[Pappenberger et al., 2010] Pappenberger, F., Ratto, M., and Vandenberghe, V. (2010). Review of sensitivity analysis methods. In *P.A. Vanrolleghem (Ed): Modelling aspects of water framework directive implementation*, pages 191–265. IWA Publishing.

[Park et al., 2005] Park, H.-J., West, T. R., and Woo, I. (2005). Probabilistic analysis of rock slope stability and random properties of discontinuity parameters, interstate highway 40, western north carolina, usa. *Engineering Geology*, 79(3):230–250.

[Park and Baek, 2001] Park, J.-S. and Baek, J. (2001). Efficient computation of maximum likelihood estimators in a spatial linear model with power exponential covariogram. *Computers & Geosciences*, 27(1):1–7.

[Parry and Drouin, 2009] Parry, G. and Drouin, M. (2009). Risk-informed regulatory decision-making at the us nrc: dealing with model uncertainty. *Nuclear Regulatory Commission*, page 2009.

[Paté-Cornell, 2002] Paté-Cornell, E. (2002). Risk and uncertainty analysis in government safety decisions. *Risk Analysis*, 22(3):633–646.

[Patt and Dessai, 2005] Patt, A. and Dessai, S. (2005). Communicating uncertainty: lessons learned and suggestions for climate change assessment. *Comptes Rendus Geoscience*, 337(4):425–441.

[Pedroni et al., 2013] Pedroni, N., Zio, E., Ferrario, E., Pasanisi, A., and Couplet, M. (2013). Hierarchical propagation of probabilistic and non-probabilistic uncertainty in the parameters of a risk model. *Computers & Structures*, 126:199–213.

[Piedra-Morales, 1991] Piedra-Morales, I. (1991). Carrière de beauregard, à pont evêque (38), analyse de la stabilité et définition des moyens à mettre en œuvre pour assurer la sécurité. Technical report, BRGM (French geological survey) technical report RR.34089-FR (in French).

[R Core Team, 2014] R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[Rabinowitz and Steinberg, 1991] Rabinowitz, N. and Steinberg, D. M. (1991). Seismic hazard sensitivity analysis: a multi-parameter approach. *Bulletin of the Seismological Society of America*, 81(3):796–817.

[Ramsay and Silverman, 2005] Ramsay, J. and Silverman, B. (2005). *Functional data analysis*. Springer, New York.

[Rasmussen and Williams, 2006] Rasmussen, C. and Williams, C. (2006). *Gaussian Processes for Machine Learning*. MIT Press.

[Ratto et al., 2001] Ratto, M., Tarantola, S., and Saltelli, A. (2001). Sensitivity analysis in model calibration: Gsa-glue approach. *Computer Physics Communications*, 136(3):212–224.

[Raufaste et al., 2003] Raufaste, E., da Silva Neves, R., and Mariné, C. (2003). Testing the descriptive validity of possibility theory in human judgments of uncertainty. *Artificial Intelligence*, 148(1):197–218.

[Rogers, 2003] Rogers, M. D. (2003). Risk analysis under uncertainty, the precautionary principle, and the new eu chemicals strategy. *Regulatory Toxicology and Pharmacology*, 37(3):370–381.

[Rohmer, 2013] Rohmer, J. (2013). Dynamic sensitivity analysis of long-running landslide models through basis set expansion and meta-modelling. *Natural Hazards*, pages 1–18.

[Rohmer and Baudrit, 2011] Rohmer, J. and Baudrit, C. (2011). The use of the possibility theory to investigate the epistemic uncertainties within scenario-based earthquake risk assessments. *Natural hazards*, 56(3):613–632.

[Rohmer and Foerster, 2011] Rohmer, J. and Foerster, E. (2011). Global sensitivity analysis of large-scale numerical landslide models based on gaussian-process meta-modeling. *Computers & Geosciences*, 37(7):917–927.

[Rohmer and Verdel, 2014] Rohmer, J. and Verdel, T. (2014). Joint exploration of regional importance of possibilistic and probabilistic uncertainty in stability analysis. *Computers and Geotechnics*, 61:308–315.

[Rougier and Beven, 2013] Rougier, J. and Beven, K. J. (2013). Model and data limitations: the sources and implications of epistemic uncertainty. *Risk and Uncertainty Assessment for Natural Hazards*, page 40.

[Roustant et al., 2012] Roustant, O., Ginsbourger, D., and Deville, Y. (2012). Dicekriging, diceoptim: Two r packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, 51:1–55.

[Sacks et al., 1989] Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical science*, pages 409–423.

[Saeidi et al., 2013] Saeidi, A., Deck, O., and Verdel, T. (2013). Comparison of building damage assessment methods for risk analysis in mining subsidence regions. *Geotechnical and Geological Engineering*, 31(4):1073–1088.

[Saint-Geours et al., 2015] Saint-Geours, N., Tarantola, S., Kopustinskas, V., and Bolado-Lavin, R. (2015). Computing first-order sensitivity indices with contribution to the sample mean plot. *Journal of Statistical Computation and Simulation*, 85(7):1334–1357.

[Saltelli, 2002] Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297.

[Saltelli and Annoni, 2010] Saltelli, A. and Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12):1508–1517.

[Saltelli et al., 2010] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270.

[Saltelli et al., 2008] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.

[Saltelli et al., 1999] Saltelli, A., Tarantola, S., and Chan, K.-S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56.

[Santner et al., 2003] Santner, T. J., Williams, B. J., and Notz, W. (2003). *The design and analysis of computer experiments*. Springer.

[Saporta, 2011] Saporta, G. (2011). *Probabilités, analyse des données et statistique*. Editions Technip.

[Sarma et al., 2008] Sarma, P., Durlofsky, L. J., and Aziz, K. (2008). Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics. *Mathematical Geosciences*, 40(1):3–32.

[Savage, 1954] Savage, L. (1954). *The Foundations of Statistics*. New York: Wiley.

[Schmidt et al., 2008] Schmidt, J., Turek, G., Clark, M., Uddstrom, M., and Dymond, J. (2008). Probabilistic forecasting of shallow, rainfall-triggered landslides using real-time numerical weather predictions. *Natural Hazards and Earth System Science*, 8(2):349–357.

[Schneider and Moss, 1999] Schneider, S. and Moss, R. (1999). Uncertainties in the ipcc tar. recommendations to lead authors for more consistent assessment and reporting. *Unpublished document*.

[Shackle, 1955] Shackle, G. L. S. (1955). *Uncertainty in Economics and other Reflections*. CUP Archive.

[Shafer et al., 1976] Shafer, G. et al. (1976). *A mathematical theory of evidence*, volume 1. Princeton university press Princeton.

[Smets, 2005] Smets, P. (2005). Belief functions on real numbers. *International journal of approximate reasoning*, 40(3):181–223.

[Sobol', 1990] Sobol', I. M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 2(1):112–118.

[Sobol et al., 2007] Sobol, I. M., Tarantola, S., Gatelli, D., Kucherenko, S., and Mauntz, W. (2007). Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability Engineering & System Safety*, 92(7):957–960.

[Song et al., 2014] Song, S., Lu, Z., Li, W., and Cui, L. (2014). The uncertainty importance measures of the structural system in view of mixed uncertain variables. *Fuzzy Sets and Systems*, 243:25–35.

[Steimen, 2004] Steimen, S. (2004). *Uncertainties in earthquake Scenarios.* PhD thesis, Swiss Federal Institute of Technology Zürich.

[Stein, 1999] Stein, M. L. (1999). *Interpolation of spatial data: some theory for kriging.* Springer.

[Storlie et al., 2009] Storlie, C. B., Swiler, L. P., Helton, J. C., and Sallaberry, C. J. (2009). Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering & System Safety*, 94(11):1735–1763.

[Straub and Schubert, 2008] Straub, D. and Schubert, M. (2008). Modeling and managing uncertainties in rock-fall hazards. *Georisk*, 2(1):1–15.

[Sumner et al., 2012] Sumner, T., Shephard, E., and Bogle, I. (2012). A methodology for global-sensitivity analysis of time-dependent outputs in systems biology modelling. *Journal of The Royal Society Interface*, 9(74):2156–2166.

[Tacher et al., 2005] Tacher, L., Bonnard, C., Laloui, L., and Parriaux, A. (2005). Modelling the behaviour of a large landslide with respect to hydrogeological and geomechanical parameter heterogeneity. *Landslides*, 2(1):3–14.

[Tarantola et al., 2012] Tarantola, S., Kopustinskas, V., Bolado-Lavin, R., Kaliatka, A., Uščuras, E., and Vaišnoras, M. (2012). Sensitivity analysis using contribution to sample variance plot: Application to a water hammer model. *Reliability Engineering & System Safety*, 99:62–73.

[Thierry et al., 2009] Thierry, P., Prunier-Leparmentier, A.-M., Lembezat, C., Vanoudheusden, E., and Vernoux, J.-F. (2009). 3d geological modelling at urban scale and mapping of ground movement susceptibility from gypsum dissolution: The paris example (france). *Engineering Geology*, 105(1):51–64.

[Thywissen, 2006] Thywissen, K. (2006). Components of risk - a comparative glossary. Technical report, United Nations University, UNU-EHS, 2/2006.

[Tritsch et al., 2002] Tritsch, J., Toulemont, M., Durville, J., and Pothérat, P. (2002). Guide technique: Evaluation des aléas liés aux cavités souterraines. collection environnement-risques naturels. Technical report, LCPC/INERIS/MATE.

[Tukey, 1954] Tukey, J. W. (1954). Unsolved problems of experimental statistics*. *Journal of the American Statistical Association*, 49(268):706–731.

[UN/ISDR, 2004] UN/ISDR (2004). (united nations international strategy for disaster reduction). living with risk. a global review of disaster reduction initiatives. Technical report, United Nations, Geneva.

[van Asselt and Rotmans, 2002] van Asselt, M. and Rotmans, J. (2002). Uncertainty in integrated assessment modelling. *Climatic Change*, 54:75–105.

[Von Neumann and Morgenstern, 1944] Von Neumann, J. and Morgenstern, O. (1944). Theory of games and economic behavior.

[Walker et al., 2003] Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B., Janssen, P., and Krayer von Krauss, M. P. (2003). Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated assessment*, 4(1):5–17.

[Walley, 1991] Walley, P. (1991). *Statistical reasoning with imprecise probabilities.* Chapman and Hall London.

[Waltham et al., 2005] Waltham, T., Bell, F. G., and Culshaw, M. (2005). *Sinkholes and subsidence.* Springer.

[Wieland et al., 2012] Wieland, M., Pittore, M., Parolai, S., Zschau, J., Moldobekov, B., and Begaliev, U. (2012). Estimating building inventory for rapid seismic vulnerability assessment: Towards an integrated approach based on multi-source imaging. *Soil Dynamics and Earthquake Engineering*, 36:70–83.

[Winckel et al., 2004] Winckel, A., Petitjean, J., Borie, M., Mallet, C., and Aubié, S. (2004). Etat des connaissances hydrologiques et hydrogéologiques de la côte basque. Technical report, BRGM/RP-53372-FR, 113p.

[Zadeh, 1978] Zadeh, L. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 1:3–28.

[Zadeh, 1965] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.

[Zadeh, 1975] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—i. *Information sciences*, 8(3):199–249.

[Zadeh, 2008] Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information Sciences*, 178(13):2751–2779.

[Zhang et al., 2010] Zhang, H., Mullen, R. L., and Muhanna, R. L. (2010). Interval monte carlo methods for structural reliability. *Structural Safety*, 32(3):183–190.

[Zimmerman and Cressie, 1992] Zimmerman, D. L. and Cressie, N. (1992). Mean squared prediction error in the spatial linear model with estimated covariance parameters. *Annals of the institute of statistical mathematics*, 44(1):27–43.

Importance ranking of parameter uncertainties in geo-hazard assessments

Epistemic uncertainty can be reduced via additional lab or in site measurements or additional numerical simulations. We focused here on parameter uncertainty: this corresponds to the incomplete knowledge of the correct setting of the input parameters (like values of soil properties) of the model supporting the geo-hazard assessment. A possible option to manage it is via sensitivity analysis, which aims at identifying the contribution (i.e. the importance) of the different input parameters in the uncertainty on the final hazard outcome. For this purpose, advanced techniques exist, namely variance-based global sensitivity analysis. Yet, their practical implementation faces three major limitations related to the specificities of the geo-hazard domain: 1. the large computation time cost (several hours if not days) of numerical models; 2. the parameters are complex functions of time and space; 3. data are often scarce, limited if not vague. In the present PhD thesis, statistical approaches were developed, tested and adapted to overcome those limits. A special attention was paid to test the feasibility of those statistical tools by confronting them to real cases (natural hazards related to earthquakes, cavities and landslides).

Key words: Sensitivity Analysis; Lack of Knowledge; Model Parameter; Long-running Simulators, Time-varying Parameters; Fuzzy Sets.

Analyse de sensibilité des incertitudes paramétriques dans les évaluations d'aléas géotechniques

Les incertitudes épistémiques peuvent être réduites via des études supplémentaires (mesures labo, in situ, ou modélisations numériques, etc.). Nous nous concentrons ici sur celle "paramétrique" liée aux difficultés à évaluer quantitativement les paramètres d'entrée du modèle utilisé pour l'analyse des aléas géotechniques. Une stratégie de gestion possible est l'analyse de sensibilité, qui consiste à identifier la contribution (i.e. l'importance) des paramètres dans l'incertitude de l'évaluation de l'aléa. Des approches avancées existent pour conduire une telle analyse. Toutefois, leur application au domaine des aléas géotechniques se confronte à plusieurs contraintes : 1. le coût calculatoire des modèles numériques (plusieurs heures voire jours) ; 2. les paramètres sont souvent des fonctions complexes du temps et de l'espace ; 3. les données sont souvent limitées, imprécises voire vagues. Dans cette thèse, nous avons testé et adapté des outils statistiques pour surmonter ces limites. Une attention toute particulière a été portée sur le test de faisabilité de ces procédures et sur la confrontation à des cas réels (aléas naturels liés aux séismes, cavités et glissements de terrain).

Mots clefs : Analyse de sensibilité ; Manque de connaissances ; Paramètre de modèle ; Simulateur couteux en temps de calcul ; Série temporelle ; Ensemble flou.