



HAL
open science

Couplage données historiques - modélisation numérique pour la prédétermination des avalanches : une approche bayésienne

Nicolas Eckert

► **To cite this version:**

Nicolas Eckert. Couplage données historiques - modélisation numérique pour la prédétermination des avalanches : une approche bayésienne. Mathematics [math]. AgroParisTech, 2007. English. NNT : 07AGPT0060 . pastel-00003404

HAL Id: pastel-00003404

<https://pastel.hal.science/pastel-00003404>

Submitted on 2 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

pour obtenir le grade de

Docteur

de

**l'Institut des Sciences et Industries du Vivant et de l'Environnement
(Agro Paris Tech)**

Spécialité : Statistique appliquée aux sciences de l'Environnement

présentée et soutenue publiquement par

Nicolas ECKERT

le 20 décembre 2007

COUPLAGE DONNEES HISTORIQUES - MODELISATION NUMERIQUE POUR LA PREDETERMINATION DES AVALANCHES : UNE APPROCHE BAYESIENNE

Directeur de thèse : Eric PARENT

*Travail réalisé au sein du Cemagref, Unité Erosion Torrentielle, Neige et Avalanches, 38402 Saint
Martin d'Hères*

*Avec une collaboration active de l'ENGREF, équipe MOdélisation et Risque en Statistique
Environnementale, 75732 Paris*

Devant le jury :

M. Gérard DEGOUTTE, IGGREF, CEMAGREF AIX EN PROVENCE-CGGREF, **Président**
Mme. Anne-Catherine FAVRE, Professeure, INRS DE QUEBEC, **Rapporteur**
M. Christopher KEYLOCK, Enseignant-chercheur, UNIVERSITE DE LEEDS, **Rapporteur**
M. Jean-Noël BACRO, Professeur, UNIVERSITE DE MONTPELLIER, **Examineur**
M. Jean-Jacques BOREUX, Enseignant-chercheur, UNIVERSITE DE LIEGE, **Examineur**
M. Mohamed NAAIM, Directeur de recherche IGRF, CEMAGREF GRENOBLE, **Examineur**
M. Eric PARENT, ICGREF HDR, ENGREF-AGRO PARIS TECH, **Examineur**
M. Didier RICHARD, IGRF chef d'unité, CEMAGREF GRENOBLE, **Examineur**

Remerciements

Jusqu'à il y a peu, je me disais chaque fois que je voyais la page « remerciements » d'une thèse que c'était juste un élément du plan type au même titre que le sommaire ou le résumé. Mais maintenant que j'ai moi-même « souffert » pour produire quelque chose de personnel, je me rends compte que les remerciements à la fin d'une thèse, c'est beaucoup plus qu'un passage obligé. C'est une façon de rendre un peu aux autres ce qu'ils vous ont donné au cours de trois ans de travail, voire d'un cheminement encore plus long. J'ai donc moi aussi envie de dire ma gratitude sincère à tous ceux qui ont permis à ce travail d'aboutir, et tant pis si celui qui ouvrira ce manuscrit passe directement à la suite!

Mes remerciements vont en premier lieu à Eric Parent qui a été successivement mon tuteur à l'ENGREF, puis celui qui m'a fait découvrir que la statistique, c'est autre chose que calculer des moyennes et des barres d'erreur, puis enfin un directeur de thèse lointain (près de 600 km) mais en même temps toujours proche. Merci Eric d'avoir pris le temps de donner à un ingénieur les clés pour comprendre la modélisation. Merci aussi de m'avoir fait rencontrer dans des endroits agréables (Londres, Megève, Niolon, etc.) toute une communauté de chercheurs compétents et sympathiques.

Mes remerciements vont également à Mohamed Naaim et Didier Richard qui m'ont encadré tout au long de ce travail au sein de l'équipe ETNA du Cemagref. Ils m'ont fait découvrir les risques en montagne et les avalanches, m'ont intégré dans une équipe dynamique et m'ont considéré dès le départ comme un collaborateur sur lequel on pouvait compter, y compris pour une expertise.

La transition est toute trouvée vers l'ensemble des équipes ETNA et MORSE. J'en remercie tous les membres pour ces années agréables, et en particulier pour l'ensemble des discussions enrichissantes que l'on a pu avoir ensemble ainsi que pour les moments de convivialité que l'on a partagé. Je pense en particulier aux personnels administratifs qui ont su me guider dans l'univers terrifiant pour le non initié des ordres de missions, états de frais et autres indicateurs d'activité. Merci donc à Claude Pigeon et Paola Pentier à l'ENGREF ainsi qu'à Martine Girier et Corinne Manquat au Cemagref.

Je remercie aussi les institutions qui ont permis à ce travail de recherche d'être accompli : l'ENGREF pour m'avoir autorisé à suivre un cursus personnalisé au cours de ma dernière année de scolarité, le Cemagref pour m'avoir accueilli, tout d'abord sous la forme d'un projet de Formation Complémentaire par la Recherche puis d'une affectation définitive, ainsi que le corps du GREF pour m'avoir jugé apte à réaliser le projet qui me motivait. J'espère avoir été digne de cette confiance et je suis très content de pouvoir continuer dans les années à venir à travailler dans le domaine des risques en montagne.

J'ai également une pensée affectueuse pour tous ceux qui, très récemment ou il y a bien plus longtemps, ont permis à leur manière l'aboutissement de ce travail : l'ensemble des observateurs de terrain de l'Enquête Permanente sur les Avalanches sans le patient travail desquels rien n'aurait été possible, les enseignants m'ayant donné envie de faire des sciences, les personnes m'ayant fait partager leur expérience et/ou guidé vers ce sujet et, plus simplement, les collègues avec qui j'ai pu échanger des idées. En en oubliant sans doute beaucoup, je pense ainsi à Claude Schneider, Fabrice Brunet et Bruno Goffé, Claude Michel et Vazken Andreassian, Ghislain De Marsily et Gérard Degoutte, Nicky Best et Sylvia Richardson, Urs Gruber, Maurice Meunier et Christophe Ancy, mes stagiaires Serigne, Julien et Robert, etc.

Je terminerai en remerciant tous ceux qui m'ont toujours soutenu, que ce soit dans mon travail ou dans la vie de tous les jours : mes parents, ceux d'Irène, Catherine et Christophe, Marianne, mes amis Alain, Alexis, Jeanne, Julie, Olivier, Amandine, Franck, Bertrand, Seb et tous les autres ainsi que mes collègues de labeur du couloir de la soufflerie et d'ailleurs. Enfin, je ne peux qu'offrir une dédicace spéciale à Irène pour m'avoir supporté tout au long de ces trois ans, y compris les jours où j'étais dépité ou perdu dans mes pensées parce que « ca ne marchait pas ». Merci Irène pour ta présence et ta joie de vivre. Parfois la science » doit s'arrêter à la porte du bureau !

Sommaire

Remerciements	3
Sommaire	5
1. Introduction	7
1.1. Des méthodes standardisées pour les risques naturels	7
1.2. Dans le cas des avalanches, l'ingénieur est démuni.....	8
1.3. Comment progresser ?	10
1.4. Le mémoire de thèse est organisé en huit chapitres.....	11
2. Synthèse des connaissances disponibles	14
2.1. Connaître les avalanches.....	14
2.2. Se protéger des avalanches	28
2.3. Modélisation statistique et inférence bayésienne.....	32
2.4. Conclusion du chapitre 2	40
3. Assembler les connaissances	41
3.1. Revisiter et approfondir les approches statistiques-dynamiques.....	41
3.2. Modélisation du fonctionnement d'un site avalancheux.....	43
3.3. Inférence bayésienne du modèle proposé	51
3.4. Prédiction par mise en œuvre des méthodes de Monte Carlo	52
3.5. Remettre en cause les hypothèses	56
3.6. Construire un cadre décisionnel.....	57
3.7. Conclusion du chapitre 3	64
4. Construction et estimation du modèle d'intensité	65
4.1. Un modèle simplifié inversible.....	65
4.2. Un modèle intermédiaire non inversible.....	82
4.3. Un modèle de type fluide.....	102
4.4. Conclusion du chapitre 4	110
5. Construction et estimation du modèle fréquentiel	112
5.1. Un modèle poissonien local.....	112
5.2. Intégrer la dimension spatiale	114
5.3. Améliorer la modélisation spatiale	127
5.4. Améliorer la modélisation spatio-temporelle.....	136
5.5. Bilan et application du modèle final	146
5.6. Conclusion du chapitre 5	151
6. Calcul des aléas de référence	152
6.1. Moyenne <i>a posteriori</i> et période de retour.....	152
6.2. Lois prédictives.....	165
6.3. Analyse de sensibilité	173
6.4. Calcul de pression.....	188
6.5. Conclusion du chapitre 6	195
7. Dimensionnement d'un ouvrage de protection	198
7.1. Définir un aléa de référence bivarié pour le dimensionnement	198
7.2. Un modèle d'optimisation simplifié analytique.....	205
7.3. Un modèle d'optimisation numérique	216
7.4. Conclusion du chapitre 7	225
8. Synthèse, conclusion et perspectives	227
8.1. Bilan des résultats obtenus.....	227
8.2. Avalanches et période de retour.....	231
8.3. L'importance de la topographie	233
8.4. Pourquoi une approche bayésienne ?.....	234
8.5. Comment situer l'approche proposée ?.....	236
8.6. Conclusion et perspectives.....	238
Références	241
Et sur le web :	254
Liste des Figures	255
Liste des Tableaux	258
Liste des Variables et notations	260
Liste des abréviations et sigles	265
Annexe : jeux de données utilisés	266

1. Introduction

1.1. Des méthodes standardisées pour les risques naturels

1.1.1. La prédétermination : une réponse sur le long terme

La notion de risque naturel s'applique à une grande variété de phénomènes caractérisés par des échelles de temps et d'espace différentes : séismes, écroulements rocheux, crues, avalanches, tempêtes...etc (Besson, 1996). Elle est usuellement découpée en une composante d'aléa et une composante de vulnérabilité. La première correspond à la réalisation du phénomène. La seconde, dont la définition est plus floue, traduit l'existence d'enjeux ainsi que leur fragilité, en particulier l'intensité des dommages qu'ils sont susceptibles de subir en fonction de l'intensité de l'aléa. Différentes vulnérabilités sont souvent distinguées : humaine, physique, sociale, organisationnelle, voire environnementale.

En toute rigueur, le travail de l'ingénieur en charge de l'évaluation du risque se décompose donc en deux étapes : quantification de l'aléa en extension et en intensité puis identification des enjeux concernés et des dommages prévisionnels en fonction de l'intensité de l'aléa. Si le niveau de risque est jugé inacceptable, des mesures de mitigation temporaires ou permanentes peuvent alors être proposées. La sécurité étant au cœur des missions régaliennes de l'Etat, la puissance publique est en première ligne pour garantir la protection des populations et des actifs menacés. Les interventions s'effectuent selon trois axes :

- La gestion de crise consiste en un ensemble d'actions entreprises en contexte de risque imminent ou juste après une catastrophe et destinées à minimiser des dommages avérés ou prévus à très court terme. Parmi elles, citons l'évacuation, l'interdiction temporaire d'accès à certaines zones, l'affichage et la diffusion de consignes de prudence ou encore l'organisation des secours.
- La prévision : il s'agit de prévoir l'occurrence et si possible l'intensité d'un phénomène dommageable avant qu'il ne se produise. Son intérêt est de permettre l'enclenchement des mesures de gestion de crise dès qu'un seuil d'alerte est dépassé et plus généralement l'affichage du niveau d'aléa attendu selon une échelle allant, pour de nombreux aléas, de 1 à 5. La base de la prévision est l'analyse en temps réel de l'évolution des variables pertinentes pour le phénomène étudié.
- La prédétermination à laquelle s'intéresse cette thèse constitue, au contraire de la prévision, une réponse à long terme. Son objectif est d'envisager toutes les occurrences possibles d'un phénomène sur un site donné afin d'organiser l'aménagement de l'espace en conséquence et de mettre en place, si nécessaire, des mesures de protection permanentes. Pour la prédétermination, se pose donc de façon aiguë le problème de la détermination des zones menacées par les événements les plus extrêmes seulement.

1.1.2. En théorie, un cadre rigoureux : aléa de référence et période de retour

En ingénierie, le problème de la prédétermination est très souvent traité par le biais de la définition d'un aléa de référence. Ce dernier correspond à une réalisation du phénomène étudié aux conséquences potentiellement dommageables du fait de son intensité élevée et de sa rareté : débit de crue de projet, extension maximale d'une lave torrentielle, etc. L'aléa de référence constitue l'élément de base incontournable pour le dimensionnement d'éventuels ouvrages de protection et le zonage, c'est-à-dire le découpage de l'espace en secteurs assortis si nécessaire de prescriptions en matière d'urbanisme. L'intérêt d'une telle démarche est de remplacer l'analyse complexe d'un phénomène intrinsèquement variable par celle, plus simple, d'un événement unique, que celui-ci soit totalement fictif ou dérivé d'une observation.

L'aléa de référence est souvent exprimé sous la forme d'une période de retour pour une ou plusieurs variables d'intérêt. Celle-ci quantifie, sous hypothèse de répétabilité du phénomène et d'indépendance des occurrences, l'intervalle de temps moyen en années séparant deux événements d'intensité au

moins égale à celle de l'aléa de référence. Ainsi, si $F(z)$ est la fonction de répartition annuelle du phénomène Z considéré, la période de retour T_z associée à l'aléa de référence z est donnée par l'équation 1.1. C'est donc simplement l'inverse de la probabilité annuelle de dépassement de la valeur z .

$$T_z = \frac{1}{1 - F(z)} \text{ (eq. 1.1)}$$

La période de retour permet d'exprimer dans la même unité, l'année, différents phénomènes se produisant à des échelles de temps et d'espace différentes : crue éclair sur un petit bassin versant montagnard ou crue de plaine d'un grand fleuve par exemple. Du point de vue de la puissance publique, elle constitue donc un outil normatif et réglementaire de premier choix. En effet, elle permet de définir des seuils d'aléas unificateurs pour le zonage et le dimensionnement d'ouvrages de protection. Ainsi, au moins en théorie, l'égalité des citoyens devant différents risques naturels est garantie et un égal effort d'investissement public en termes de mesures de protection est possible. Par exemple, les zones soumises à un risque d'inondation fort correspondent en France aux terrains concernés par l'extension d'une crue au plus centennale, et ce quel que soit le bassin versant considéré. De même, les grands ouvrages hydrauliques sont presque toujours dimensionnés pour une crue de projet décennale.

1.1.3. En pratique, des méthodes discutables mais ayant le mérite d'exister

Malheureusement, la rigueur de ce schéma théorique se heurte à la complexité des phénomènes naturels. Les connaissances disponibles sont généralement insuffisantes pour permettre le calcul d'aléas de référence sur la base de la seule connaissance physique. Ainsi les lois d'écoulement des modèles hydrauliques font appel à un ou plusieurs coefficients partiellement inconnus caractérisant la rugosité du lit. De même, le comportement de la neige en écoulement reste trop imparfaitement connu pour éviter le recours à des coefficients tabulés traduisant les propriétés de la neige en mouvement. En ingénierie des risques naturels, la prédétermination utilise donc des procédures normalisées s'appuyant au moins pour partie sur des connaissances empiriques et/ou de l'apprentissage à partir des données disponibles.

En outre, un jeu d'hypothèses souvent invérifiables, voire parfois erronées, est nécessaire à l'obtention des aléas de référence. Ainsi en va-t-il des modèles de valeurs extrêmes très couramment utilisés en hydrologie pour la définition des crues de projet décennales. Ceux-ci postulent la max-stabilité et la stationnarité sur le très long terme des séries de débits de crue. La max stabilité est pourtant quasi invérifiable avec des séries courtes de données tandis que la stationnarité sur le très long terme est plus que douteuse étant donné les fluctuations du climat au cours des derniers millénaires. Les procédures normalisées de calcul d'aléa de référence correspondent donc davantage à la nécessité d'un compromis entre connaissances disponibles et impératifs opérationnels qu'à « la vérité » des phénomènes étudiés. Mais elles ont le mérite d'exister et de proposer à l'ingénieur des solutions permettant de mettre en œuvre la législation, ce qui garantit, à défaut de mieux, une « égalité méthodologique théorique » des citoyens face au risque.

1.2. Dans le cas des avalanches, l'ingénieur est démuni

1.2.1. Un risque bien spécifique

Le risque avalanche est l'un des principaux risques naturels en montagne. Il possède la particularité d'être fortement limité dans l'espace et le temps. Ainsi, en France, un peu plus de 600 communes sont identifiées comme menacées, principalement dans les départements de haute montagne des Alpes et des Pyrénées mais également dans les Vosges, le Massif Central et le Jura, et ce de la mi-novembre à la mi-mai. A l'échelle européenne, le risque avalanche existe dans l'ensemble des pays de l'arc alpin et des Carpates : Suisse, Autriche, Italie, Slovénie, Slovaquie, Pologne, etc. ainsi qu'en Scandinavie et en Islande.

Les avalanches concernent en premier lieu les pratiquants des sports de montagne hivernaux : ski de randonnée, ski hors piste, raquette..., etc. Mais, plus ponctuellement, lorsque des conditions climatiques particulièrement défavorables s'installent, elles sont également responsables de pertes

humaines importantes dans les vallées et de dégâts économiques conséquents. Ainsi l'Europe a-t-elle connu un hiver 1998-99 particulièrement coûteux et meurtrier (SLF Davos, 2000). Les dommages matériels ont été évalués à environ 500 millions d'euros (Fuchs et McAlpin, 2005) et plusieurs catastrophes majeures ont occasionné de nombreuses victimes, notamment à Galthur, Autriche, où 31 personnes sont décédées et 22 autres ont été blessées (Barbolini et al., 2004b).

L'aléa avalancheux est caractérisé par sa soudaineté et sa brutalité. Contrairement aux crues de plaine par exemple, la durée d'une avalanche ne permet pas d'imaginer une évacuation des populations après déclenchement du phénomène. De plus, tout impact d'une avalanche sur une construction ou un véhicule a des conséquences potentiellement désastreuses, de sorte que des mesures de mitigation temporaires destinées à réduire les dommages sont extrêmement difficiles à mettre en œuvre. La prévision au jour le jour est donc très importante afin de détecter les situations critiques et de procéder à d'éventuelles évacuations préventives. De plus, la gestion à long terme de l'urbanisme se doit d'être particulièrement soignée afin de limiter au maximum la construction d'habitations ou d'infrastructures dans des zones potentiellement touchées et d'envisager, si nécessaire, la construction de protections paravalanches permanentes (digues, tunnels,... etc.).

Une spécificité supplémentaire de l'aléa avalancheux est son caractère réellement multivarié. Pour évaluer le risque, l'extension de l'aléa, les champs de vitesse et de pression, le volume de dépôt, la hauteur de l'écoulement, etc. doivent tous être évalués, et ce alors que les observations montrent que ces variables sont souvent mal corrélées. Une avalanche à faible distance d'arrêt peut par exemple infliger de très fortes pressions d'impact avec une neige dense et très cohésive.

1.2.2. Pas de réponses toute faites

Parmi les aléas naturels, les avalanches sont particulièrement mal connues en tant que phénomène physique. La connaissance des processus élémentaires, érosion/dépôt ou comportement de la neige en écoulement par exemple, reste encore très imparfaite et progresse lentement. De même, la répartition spatio-temporelle des avalanches à l'échelle d'un massif montagneux reste souvent mal connue du fait d'observations de terrain rares et entachées de fortes incertitudes. Les effets du changement climatique sur l'activité avalancheuse sont en particulier encore très mal appréhendés.

Cette méconnaissance physique a pour corollaire un très faible développement des procédures normalisées de définition d'un aléa de référence pour le dimensionnement et/ou le zonage (Ancey et Richard, 2000). Quelques méthodes existent, elles seront évoquées au chapitre 2. Cependant, elles utilisent toutes des hypothèses très discutables et se heurtent au caractère multivarié de l'aléa. Par conséquent, même si un certain consensus existe à l'échelle européenne sur le découpage en trois zones d'aléa (faible, moyen, fort), aucune procédure systématisée pour leur obtention ne se dégage réellement.

C'est particulièrement vrai en France où la prise en compte du risque avalanche dans l'aménagement du territoire montagnard repose pour l'heure essentiellement sur l'expertise, c'est-à-dire la capacité de l'ingénieur en charge de l'évaluation de l'aléa à effectuer, sur la base de son expérience, une synthèse entre différentes sources d'information éparses (données historiques, témoignages, analyse forestière, approches géomorphologiques...) pour définir un ou plusieurs aléa(s) de référence (Jarry et Ancey, 2006). La modélisation numérique n'est utilisée que dans certains cas particuliers car elle est souvent jugée lourde à mettre en œuvre.

1.2.3. Mais une nécessité d'action rapide

Dans des zones de montagne de l'Europe, la population permanente et surtout temporaire est en forte croissance depuis le milieu du vingtième siècle du fait du développement des activités touristiques, notamment hivernales. La place y étant par définition réduite, la pression foncière est très forte et de nouveaux espaces tendent à être progressivement urbanisés. Cette évolution rapide doit être suivie malgré l'état actuel des connaissances et des procédures réglementaires. Hors d'Europe, et particulièrement en Amérique du Nord, c'est moins vrai parce que les zones de montagne sont souvent caractérisées par des densités de population plus faibles limitant ainsi le risque avalanche aux seuls pratiquants de sports d'hiver et aux routes.

La nécessité d'action est renforcée par les crises exceptionnelles qui sont fortement médiatisées et difficilement acceptables aux yeux de la population et des décideurs. Ainsi l'avalanche de Montroc du

9 février 1999 (Ancey et al., 2000 ; Rapin et Ancey, 2000) est venue rappeler en France le risque que représentent les grandes avalanches pour les habitants et le patrimoine bâti. Quatorze chalets, tous situés en zone bleue ou blanche du Plan de Prévention des Risques (PPR) de la vallée de Chamonix, ont été rasés et douze décès ont été déplorés (Figure 1.1).

Le rapport de l'Inspection Générale de l'Environnement (Glass et al., 2000) consécutif à cette catastrophe a constaté qu'« une réactualisation de la politique publique de prévention des avalanches » était nécessaire pour, en particulier, définir plus rigoureusement un aléa de référence. Dans les autres pays européens, un constat et des directives similaires ont émergé à la suite de l'hiver 1999. Ils ont entraîné une réflexion importante ainsi que des évolutions à la fois théoriques et opérationnelles dans les méthodes de caractérisation de l'aléa avalancheux.



Figure 1.1 : Catastrophe de Montroc

9 février 1999. Commune de Chamonix-Mont-Blanc, Haute-Savoie France.

Le lendemain de l'avalanche, les secouristes sont encore au travail et fouillent le dépôt. Au second plan, partie supérieure d'un chalet brisée et déplacée par l'avalanche.

Photo F. Rapin/Cemagref.

1.3. Comment progresser ?

1.3.1. Sur le terrain

A l'échelle européenne, une dynamique de modernisation des méthodes utilisées en ingénierie a été enclenchée. Néanmoins, des solutions assez différentes ont été apportées par les différents pays. Ainsi la Suisse a renforcé le recours à une méthode systématisée faisant appel à une modélisation numérique relativement avancée. L'Islande a de son côté privilégié la prise en compte des enjeux à travers le développement d'une méthode de calcul de risque. En France, l'accent a été mis sur l'inventaire des sites à risques, leur évaluation et la mise à la disposition des praticiens de l'ensemble des connaissances disponibles. Ainsi une vigoureuse action de restauration des bases de données avalancheuses a été lancée (Bélangier, 2004). De même une méthode permettant l'évaluation de la dangerosité des couloirs a été développée (Rapin et al., 2004). Enfin un guide méthodologique pour la constitution du volet avalanche des PPR a été rédigé (MEDD et al., 2004). Celui-ci, inspiré des directives utilisées pour le risque d'inondation, insiste fortement sur la prise en compte pour chaque site d'un aléa de référence, en particulier centennal. Toutefois, même s'il propose une revue des méthodes d'analyse utilisables, il n'offre guère d'éléments permettant son obtention de manière systématique et uniformisée.

1.3.2. En recherche fondamentale

En matière de recherche fondamentale, les efforts des différents pays européens ont été plus coordonnés, notamment au travers des Programmes Cadres de Recherche et Développement (PCRD) successifs. Leur intérêt a été de créer un réseau de partenaires efficace qui permet d'avancer sur des sujets difficiles en mettant en commun les moyens d'équipes de tailles relativement modestes. L'essentiel de l'effort de recherche a concerné les processus élémentaires, notamment au niveau de l'interaction des écoulements avec des obstacles. Plus récemment, l'accent a également été mis sur

l'amélioration des méthodes d'évaluation de la vulnérabilité. Néanmoins, la vitesse des progrès en matière de connaissances physiques reste insuffisante pour proposer rapidement des solutions à l'ensemble des questions auxquelles doit répondre l'ingénieur sur le terrain.

1.3.3. A l'interface : la place de cette thèse

Le fossé existant entre la recherche fondamentale et les préoccupations de terrain a donné naissance au début des années 2000 à une thématique de recherche appliquée. Celle-ci vise à utiliser au mieux l'ensemble des connaissances disponibles pour la caractérisation des aléas de référence. Elle a été portée notamment par Bozhinsky et al. (2001) en Russie, Barbolini et Savi. (2001) en Italie et Meunier et al. (2001) au Cemagref de Grenoble. Son principe est d'intégrer la modélisation physique des avalanches dans un cadre stochastique multivarié. Ce faisant, le phénomène avalancheux est envisagé dans sa variabilité sur le site étudié et l'apport de la modélisation numérique déterministe est valorisé. Les travaux effectués au début des années 2000 sont basés essentiellement sur des approches de Monte Carlo simples. Ils ont mis en évidence le potentiel de cette approche mais aussi plusieurs difficultés techniques majeures : modélisation de la dépendance inter-variable, formalisation de l'expertise, calibration d'un modèle numérique possédant un nombre important de paramètres et prise en compte des différentes sources d'incertitudes en prédiction.

Cette thèse s'est inscrite dans la droite ligne de cette thématique de recherche, à l'interface entre préoccupations opérationnelles et recherche fondamentale. Son objectif était tout d'abord de reprendre les travaux effectués et de les intégrer dans un cadre théorique plus rigoureux. Ensuite, il s'agissait de construire des outils de prédétermination permettant le zonage et le dimensionnement des ouvrages susceptibles d'être affectés par les avalanches en traitant rigoureusement les différents types d'incertitudes. Le travail de thèse a donc consisté dans un premier temps en l'élaboration d'un cadre stochastique intégrant les phénomènes prépondérants se produisant à l'échelle du couloir avalancheux depuis la rupture jusqu'à l'arrêt de l'écoulement. Les modèles ainsi obtenus ont ensuite été calibrés sur des sites choisis et utilisés en prévision (au sens statistique) dans une optique de gestion à long terme plutôt que de connaissance physique de l'aléa. Enfin, afin de pallier certaines des limites inhérentes à la notion d'aléa de référence, la modélisation stochastique a été intégrée dans un cadre décisionnel permettant le choix optimal d'une stratégie de protection paravalanche.

De façon pratique, l'angle d'attaque retenu a été celui de la modélisation bayésienne hiérarchique dont les méthodes modernes d'inférence par simulation constituent le prolongement naturel (Clark, 2005 ; Parent et Bernier, 2007). En premier lieu ont été explorées les solutions offertes en matière d'estimation de modèles complexes. Toutefois, la transposition au domaine des avalanches des techniques de quantification de l'expertise, de prise en compte des différentes sources d'incertitudes en prédiction, de formalisation de la dépendance spatiale et d'analyse décisionnelle ont également été abordées. Pour ce faire, la thèse a été effectuée au travers d'un partenariat entre le Cemagref et le laboratoire MORSE de l'ENGREF qui disposait de la compétence méthodologique nécessaire.

1.4. Le mémoire de thèse est organisé en huit chapitres

L'objectif du chapitre 2 est d'effectuer un bref tour d'horizon de la connaissance disponible pour la construction d'un modèle bayésien de prédétermination des avalanches et, plus généralement, pour la modélisation probabiliste des avalanches dans une perspective de gestion à long terme. Du point de vue phénoménologique sont concernées l'ensemble des approches développées pour la connaissance des avalanches : analyse naturaliste, modélisation numérique déterministe et données historiques. D'un point de vue plus méthodologique, la connaissance disponible inclut l'ensemble des développements en matière de modélisation statistique et d'inférence bayésienne. Enfin, à l'interface entre connaissance du phénomène et connaissance disciplinaire, se situent les approches de prédétermination déjà existantes ainsi que les techniques développées dans les domaines du génie paravalanche et de l'analyse du risque.

Le chapitre 3, le plus méthodologique, construit le cadre formel proposé par la thèse. Les connaissances sont assemblées de façon à constituer un ensemble cohérent pour le zonage et le dimensionnement d'ouvrages. Trois phases sont distinguées : la modélisation stochastique, l'inférence et la prédiction. La modélisation stochastique correspond à l'intégration d'un modèle de propagation d'avalanche dans une structure probabiliste décrivant la variabilité du phénomène étudié. L'inférence

correspond à la calibration probabiliste du modèle à l'aide des données disponibles. La prédiction au sens statistique du terme correspond à l'obtention des aléas de référence recherchés. L'échelle d'étude est celle d'un site avalancheux de l'Enquête Permanente sur les Avalanches (EPA). L'intensité des avalanches, c'est-à-dire l'ensemble des variables décrivant l'écoulement, est modélisée indépendamment de leur fréquence définie comme le nombre d'événements par hiver. Le choix du formalisme bayésien permet la combinaison pondérée des données locales par l'expertise naturaliste préalable ainsi que la distinction explicite des différentes sources d'incertitudes. Les bases permettant l'extension spatio-temporelle et décisionnelle du cadre proposé sont également présentées.

Le chapitre 4 est consacré au modèle d'intensité décrivant après déclenchement les caractéristiques quantitatives de l'écoulement. A partir du cadre défini dans le chapitre 3, trois jeux d'hypothèses de modélisation correspondant à différents compromis entre disponibilité des données et description de la propagation de l'avalanche sont proposés. Ce chapitre s'attaque également à la principale difficulté technique abordée par la thèse en détaillant l'inférence des trois modèles de propagation choisis à l'aide d'algorithmes de simulation ad hoc. L'objectif de ce chapitre est donc, d'un point de vue pratique, de proposer différents outils de prédétermination et, d'un point de vue méthodologique, de montrer que le problème de la calibration d'un modèle de propagation d'avalanche quelconque peut être surmonté dans un cadre bayésien.

Le chapitre 5 est consacré au modèle fréquentiel décrivant les occurrences avalancheuses à des fins de prédétermination. Celui-ci est plus « classique » que le modèle d'intensité dans la mesure où seule la variabilité des occurrences avalancheuses y est décrite, et ce par un modèle statistique explicite. Après présentation d'un modèle local très simple, ce chapitre aborde également deux questions supplémentaires : l'interpolation des données ponctuelles disponibles vers les sites non documentés et la modélisation en contexte non stationnaire. Pour ce faire, un modèle spatial puis spatio-temporel des occurrences avalancheuses est proposé. L'objectif de ce chapitre est donc, d'un point de vue pratique, de compléter le modèle d'intensité pour permettre le calcul des aléas de référence. D'un point de vue méthodologique, il s'agit de donner une dimension spatiale et spatio-temporelle à une approche jusque là purement locale et stationnaire en mettant à profit les mêmes outils d'inférence que pour le modèle d'intensité du chapitre précédent.

Le chapitre 6 combine les modèles locaux d'intensité et de fréquence pour le calcul des aléas de référence. Les trois modèles stochastiques proposés sont donc utilisés en prédiction sur les cas d'étude choisis afin de déterminer les distances d'arrêt de référence. De plus, l'incertitude liée à la méconnaissance locale du phénomène y est combinée avec sa variabilité intrinsèque pour obtenir la loi de probabilité prédictive de n'importe quelle période de retour. Les distributions des autres variables de sortie fournies par les différents modèles sont également étudiées et le calcul des pressions et des incertitudes additionnelles associées est mené. Une analyse de sensibilité aux différentes hypothèses de modélisation est proposée afin de tenter d'évaluer lesquelles sont les plus lourdes de conséquence sur le zonage de l'aléa. Au final, les différents modèles de prédétermination proposés sont comparés de façon à guider le choix de l'ingénieur intéressé par leur utilisation en fonction des contraintes du cas d'étude qu'il souhaite traiter.

Le chapitre 7 aborde le problème du dimensionnement sous incertitude d'un ouvrage de protection. En guise d'introduction, un vrai cas d'étude illustre la difficulté à obtenir des aléas de référence multivariés. Le problème du dimensionnement est traité ensuite dans un cadre décisionnel. Les effets sur l'écoulement avalancheux d'une digue verticale sont inclus dans le modèle stochastique d'aléa sous la forme d'une loi semi-empirique simple issue d'expériences de laboratoire sur les milieux granulaires. En outre, le modèle stochastique d'aléa est complété par une fonction de coût permettant l'optimisation de la hauteur de l'ouvrage. Les calculs sont tout d'abord menés analytiquement avec un modèle d'aléa monodimensionnel explicite, puis numériquement à l'aide de l'un des modèles d'aléa plus réaliste mais plus complexe décrits au chapitre 4. Le modèle simplifié met en lumière les effets principaux en permettant un calcul rapide et grossier. Le modèle numérique permet quant à lui une quantification plus fine de l'aléa et des dommages prévisionnels. L'objectif méthodologique de ce chapitre est donc d'aller jusqu'à la construction d'un cadre de travail d'aide à la décision complet. Le formalisme proposé y est illustré par un cas de figure simple mais pouvant néanmoins être rencontré en ingénierie paravalanche.

Le chapitre 8 propose un regard critique sur les résultats obtenus : synthèse des avancées, limites apparues au cours de la thèse, inflexions par rapport au projet initial, etc. Le choix d'une approche

bayésienne y est particulièrement discuté, de même que le positionnement de l'approche proposée par rapport à la pratique et aux connaissances disponibles pour d'autres risques naturels. Enfin, des perspectives pour de futurs développements y sont proposées.

Une partie de ces développements a déjà fait l'objet de publications : le cadre général du chapitre 3 dans Eckert et al. (2007a), le modèle de prédétermination locale simplifié décrit aux chapitres 4.1 et 6 dans Eckert et al. (2007b), le modèle de prédétermination locale intermédiaire décrit aux chapitre 4.2 dans Eckert et Parent (2006) et le modèle spatial de base du chapitre 5.2 dans Eckert et al. (2007c). Deux articles sur l'approche décisionnelle décrite au chapitre 7 sont en voie de finalisation (Eckert et al., 2007d ; Eckert et al., 2007e). Enfin, la méthode de calibration du modèle de prédétermination fluide et son application aux calculs de pressions d'impact décrite aux chapitres 4.3 et 6.5 reste à publier, de même que la modélisation spatio-temporelle des occurrences avalanches décrite aux chapitres 5.3 à 5.5.

2. Synthèse des connaissances disponibles

Préambule

Dans ce chapitre, la connaissance disponible pour la construction d'un modèle bayésien de prédétermination des avalanches est rapidement synthétisée. D'un point de vue phénoménologique, la connaissance naturaliste, la modélisation numérique déterministe et les données historiques sont présentées. D'un point de vue plus méthodologique, la modélisation statistique, en particulier bayésienne, et son application aux risques naturels sont rapidement décrites. Enfin, l'ensemble des approches déjà existantes en termes de génie paravalanche, de prédétermination et d'analyse de risque sont introduites

2.1. Connaître les avalanches

2.1.1. Connaissance naturaliste « macroscopique »

Le terme "avalanche de neige" désigne un écoulement rapide d'une masse importante de neige sous l'effet de la gravité faisant suite à une rupture d'équilibre au sein du manteau neigeux. La connaissance naturaliste concernant ce phénomène spectaculaire comprend la description qualitative de l'écoulement ainsi que ses liens avec la topographie et les propriétés du matériau. D'abord diffuse et confondue avec la mémoire des événements passés chez les habitants des zones de montagne, la connaissance naturaliste des avalanches s'est peu à peu singularisée et stabilisée. De nombreux ouvrages généraux concernant la neige (Llamas, 1993) et les avalanches (McClung et Schearer, 1993 ; Ancey, 1998 ; Ancey et al., 2006) ainsi que des sites web de référence (<http://www.anena.org/>) en exposent les fondements, de sorte que seuls quelques points essentiels sont rappelés.



Figure 2.1 : Avalanche dense sur site expérimental

Col du Lautaret, Hautes Alpes, France.

L'avalanche percute une entrave équipée de capteurs. L'écoulement présente une certaine cohésion permettant la formation de « boules ».

Photo O. Marco/Cemagref.

Usuellement, un site avalancheux se décompose en une zone de départ où s'initie le mouvement, une zone d'écoulement plus ou moins chenalisée et une zone d'arrêt de pente faible voire négative où sont presque toujours situés les éventuels enjeux. La zone de départ a été étudiée afin de préciser les conditions propices au déclenchement (Mases Cobero, 1997 ; Fyffe et Zaiser, 2004) et d'évaluer les surfaces de départ potentielles (Maggioni et Gruber, 2003 ; Maggioni, 2005). De telles recherches se heurtent à la difficulté d'acquisition des données en haute montagne (Nespor et Sevruk, 1998) et à la multiplicité des processus physiques en jeu : cumuls de neige, transformations thermodynamiques du manteau neigeux, redistribution de la neige par le vent, initiation et propagation de la fracture... Néanmoins, de façon simplifiée, on peut considérer le manteau neigeux comme un corps en équilibre soumis à des forces l'entraînant vers le bas (poids, suraccumulation de neige due au transport de neige par le vent, surcharge ponctuelle due au passage d'un skieur ou à une chute de corniche ou de séracs) et à des forces le maintenant en place (frottement du manteau neigeux sur le sol, ancrages latéraux sur des rochers, cohésion de la neige à l'intérieur d'une même strate et cohésion entre les différentes

strates constituant le manteau neigeux). L'équilibre est rompu et l'avalanche se déclenche lorsque la composante tangentielle du poids est supérieure à la somme des frottements et de la cohésion. Plus la pente est raide et plus la composante tangentielle est importante pour une même valeur du poids. Ainsi, les déclenchements se produisent presque uniquement sur des pentes supérieures à 30°. Toutefois, lorsque la pente est trop forte, la neige ne peut pas s'accumuler, de sorte qu'on ne rencontre que peu de déclenchements sur des pentes supérieures à 50°.

La zone d'écoulement retient l'attention des spécialistes de la mécanique des fluides du fait des processus complexes mis en jeu : lien contrainte/déformation, dépôt et reprise de neige au cours de l'écoulement, incorporation d'air, effet de souffle, etc. Une terminologie variée conduisant souvent à une certaine confusion permet de décrire la masse de neige mise en mouvement. Ainsi les avalanches peuvent-elles être classées selon le type de neige (avalanche de neige poudreuse, humide, récente...), le type de départ (avalanche de plaques, cassure ponctuelle ou « en poire »...), de saison (avalanches de printemps), etc. Du point de vue de la physique du phénomène, la classification la plus pertinente repose toutefois sur le régime d'écoulement. Le plus fréquent en climat tempéré, est l'écoulement dense (on parle également d'avalanche coulante). D'une épaisseur de quelques mètres en moyenne, il suit relativement bien les couloirs naturels et est caractérisé par une vitesse de 20 à 30 m/s. L'incorporation d'air dans l'écoulement est faible, de sorte que la masse volumique reste forte, allant de 150 kg/m³ à plus de 450 kg/m³ (Figure 2.1).



Figure 2.2 : Avalanche aérosol atteignant la forêt

Minage de la corniche sommitale de la pointe du Dard en Vanoise, Savoie, France.

Photo C. Vion/Cemagref.



Figure 2.3 : Dégâts dus à une avalanche aérosol

Chalet totalement détruit par une avalanche de poudreuse exceptionnelle le 20 janvier 1981. Commune de La Morte, Isère, France.

Photo F. Valla/Cemagref.

Le second archétype, appelé avalanche aérosol, est caractérisé par une forte incorporation d'air dans l'écoulement permettant la formation d'un nuage de particules de glace en suspension. Les avalanches aérosol présentent un écoulement turbulent composé de grandes volutes, des vitesses rapides pouvant atteindre 100 m/s, des hauteurs d'écoulement de plusieurs dizaines de mètres et des masses volumiques faibles de l'ordre de 2 à 5 kg/m³. Elles n'apparaissent généralement que sous certaines

conditions : neige froide, sèche et faiblement cohésive ainsi que pente forte. Toutefois, des exemples historiques montrent que des accumulations de neige lourde peuvent aussi parfois engendrer des aérosols. L'avalanche est considérée comme un aérosol « pur » lorsque la phase dense est inexistante. Un tel cas de figure est cependant assez rare car l'incorporation continue de neige dans la suspension turbulente au niveau d'une couche de saltation est nécessaire pour que l'aérosol ne s'atténue pas très rapidement. Les avalanches aérosols sont donc en majorité des avalanches « mixtes », même si la phase dense reste souvent invisible (Figure 2.2).

Les avalanches de neige provoquent des dégâts importants. Ces dommages dépendent fortement du type d'avalanche, du volume de neige mis en mouvement et des caractéristiques géométriques et topographiques du couloir contraignant les propriétés de l'écoulement au moment de l'impact. Les avalanches denses provoquent des dégâts liés à leurs fortes densités (Figure 1.1) alors que ceux produits par les aérosols sont essentiellement dus à l'effet de souffle et à la turbulence (Figure 2.3).

2.1.2. Processus physiques élémentaires et modélisation numérique déterministe

2.1.2.1. De l'observation à la modélisation

Dans une optique de prévention du risque et de compréhension du phénomène, des informations de plus en plus fines ont été progressivement accumulées par les ingénieurs et les scientifiques. Les observations de terrain fortuites ou *a posteriori* ont progressivement été complétées grâce à des déclenchements artificiels sur des sites expérimentaux équipés de capteurs (Gubler et al., 1986 ; Takeuchi et al. ; 2003). La difficulté de l'observation d'avalanches réelles en haute montagne a également conduit à la mise en œuvre d'expériences sur modèles réduits sous forme de neige (Bouchet, 2003 ; Bouchet et al., 2004) ou de matériaux analogiques souvent granulaires (Hutter et al. 1995).

Ces informations ont permis le développement progressif de modèles numériques déterministes de plus en plus sophistiqués. Basés sur les observations et les équations élémentaires de la mécanique, ils sont souvent qualifiés de « modèles physiques », ce qui rend leurs paramètres théoriquement extrapolables d'un site avalancheux à l'autre puisque supposés « vrais ». Cependant, les processus élémentaires restent encore largement méconnus, rendant incontournable une part au moins résiduelle d'empirisme dans la modélisation.

La démarche de modélisation poursuit dans le domaine des avalanches des objectifs variés : compréhension du phénomène d'écoulement, détermination des conditions d'occurrence, caractérisation des zones à risque, etc. (Kleemayr, 2004). De ce fait, une multitude de modèles d'avalanches fondés chacun sur un jeu d'hypothèses plus ou moins discutables existent. Effectuer un tour d'horizon, même rapide, de l'ensemble des modèles existant serait donc fastidieux. En outre, il est assez difficile de se faire une idée de l'efficacité relative des différents modèles puisque les études comparatives se bornent souvent à des analyses de sensibilité (Barbolini et al., 2000 ; McClung, 2001). Seules quelques clés de compréhension sont donc présentées.

2.1.2.2. Déclenchement et propagation

Les modèles déterministes de déclenchement visent à prévoir l'initiation du mouvement au niveau de la zone de départ. Ils s'intéressent donc à l'occurrence de l'avalanche sans prendre en compte son intensité. Ils peuvent être qualifiés de « tactiques » dans le sens où ils visent à prédire l'initiation de la fracture au niveau de la zone de départ de l'avalanche en fonction des conditions nivo-météorologiques mesurées en temps réel. Leur intérêt est de permettre un suivi de l'évolution du manteau neigeux, utile par exemple dans les stations de ski quand il s'agit de décider de l'ouverture du domaine skiable. La modélisation mécanique s'appuie sur un bilan des forces s'exerçant sur une portion de manteau neigeux. Grossièrement, son principe est le calcul du quotient entre forces résistantes et forces cisailantes d'entraînement, l'avalanche étant supposée se produire lorsqu'il est inférieur à 1 (McClung, 1981 ; Schweizer, 1999). Cependant, la méconnaissance des processus élémentaires, notamment en termes de thermodynamique de la neige, rend difficile ce type d'approche. Il a ainsi été montré qu'avec les valeurs caractéristiques de masse volumique, hauteur, pente et volume de neige mises en jeu dans une avalanche, le quotient entre forces résistantes et d'entraînement pouvait aisément varier de 0.72 à plus de six (Harbitz et al. 2001), ce qui illustre la difficulté à prédire le déclenchement de façon déterministe. Concrètement, la prévision s'effectue donc le plus souvent à l'échelle d'un massif plutôt que d'un couloir unique (CEN, 2000).

Les modèles déterministes de propagation visent à reproduire l'écoulement de l'avalanche sur un site précis. Ils s'appuient sur les équations élémentaires de la mécanique de façon à obtenir par intégration sur la topographie les champs de vitesse et de pression à l'intérieur de l'avalanche. Les plus simples fournissent en sortie uniquement la distance d'arrêt. Les plus sophistiqués proposent des jeux de paramètres dynamiques, éventuellement pluridimensionnels, aux pas d'espace et de temps utilisés par le schéma numérique de calcul. Elaborés en premier lieu par des hydrauliciens, les modèles de propagation ont focalisé les efforts de recherche dans le domaine des avalanches (Salm, 2004), de sorte qu'ils ont atteint un degré de complexité élevé et qu'il en existe actuellement une bonne cinquantaine (Harbitz, 1999). Néanmoins, ces modèles souffrent tout autant que les modèles de déclenchement du manque de connaissances du manteau neigeux. Ils font donc appel à des lois au moins partiellement empiriques, notamment en termes de lien entre vitesse et pression et de loi de frottement dont les coefficients doivent être calibrés sur des événements réels (Buser et Frutiger, 1980 ; Martinelli et al., 1980 ; Blagovechshenskiy et al. 2002).

2.1.2.3. Modélisation de la dynamique de l'écoulement

- *Echelle de description*

Les mesures par RADAR sur sites expérimentaux (Sovilla, 2004) ont mis en évidence la stratification verticale des avalanches avec une phase aérosol surmontant, lorsqu'elle existe, une phase dense. La couche basale correspondant à la phase dense de l'avalanche est caractérisée par un écoulement quasi laminaire. La dissipation d'énergie y est dominée par le frottement à l'interface avec le sol. Le rôle de la cohésion et des agrégats sur le profil de vitesse (Bouchet, 2003) et de la reprise de neige sur le volume de matériau (Sovilla et al., 2001 ; Sovilla et Bartelt, 2002) a toutefois été mis en évidence. La couche sommitale de l'écoulement représentant l'aérosol présente une densité au moins dix fois inférieure à celle de la couche basale. La dissipation d'énergie y est dominée par la turbulence et l'incorporation d'air dans l'écoulement. Les lois d'échelle ont été quantifiées par analogie avec des courants de gravité (Beghin et al., 1981). Entre les couches dense et aérosol, l'échange de masse et d'énergie s'effectue au niveau d'une interface dominée par la saltation des grains de neige et des agrégats (Issler, 1997; Naaim, 1998).

Pour s'approcher au plus près des observations, les modélisateurs représentent l'avalanche dans le cadre de la mécanique des milieux continus (Savage et Hutter, 1989) par des modèles à plusieurs couches possédant chacune des propriétés particulières (Issler, 1997; Naaim, 1998; Zwinger, 2000) et où la densification du matériau au cours de l'écoulement est quantifiée grâce à des relations empiriques. De tels modèles sont aussi réalistes que possible étant donné l'état actuel des connaissances, mais très coûteux du point de vue du temps de calcul nécessaire, l'ordre de grandeur pour une seule avalanche étant celui du jour. Ils sont de ce fait uniquement utilisés pour la rétro analyse d'événements bien documentés ou la simulation de scénarii de référence sur des sites particulièrement délicats.

Une approche pouvant être qualifiée d'intermédiaire et ne considérant que la phase dense de l'écoulement a été proposée par plusieurs auteurs (Brugnot et Pochat, 1981 ; Christen et al 2002, Naaim et al. 2004) : la hauteur de l'écoulement est supposée faible par rapport à sa longueur, de sorte que les variations verticales de la vitesse peuvent être négligées. Cela conduit au formalisme dit de Barré de Saint-Venant. La formation de l'aérosol n'est ainsi pas prise en compte. En revanche, l'érosion et le dépôt de neige au cours de l'écoulement peuvent être quantifiés de façon réaliste. L'ordre de grandeur pour un calcul de propagation utilisant le formalisme de Saint-Venant est celui de la minute, parfois moins en fonction des hypothèses simplificatrices effectuées. Les modèles de propagation de type Saint-Venant sont ce de fait maintenant assez abondamment utilisés pour le zonage sur de larges régions (Gruber et Bartelt, 2007) et la possibilité d'effectuer un nombre important de simulations en un temps raisonnable permet de les calibrer de façon automatisée sur l'ensemble des événements ayant affecté un site étudié.

Les approches les plus simples considèrent la déformation de la neige en mouvement comme négligeable et représentent de ce fait l'avalanche comme un bloc glissant rigide (Voellmy, 1955 ; Perla et al. 1980). Des processus physiques importants sont ainsi négligés, mais le temps de calcul devient si faible que des milliers de simulations peuvent être effectués en des temps très réduits (Meunier et Ancey, 2004). De plus, ces modèles sont utilisés depuis plus de 50 ans par les praticiens, de sorte que leur robustesse en matière de calcul des distances d'arrêt est bien établie.

- *Loi de frottement*

Le principal obstacle à la compréhension de la dynamique des avalanches, et ce quelle que soit l'échelle de description choisie, est la méconnaissance de la rhéologie de la neige, c'est-à-dire du lien entre contrainte et déformation au cours de l'écoulement (Dent et Lang, 1980). En termes de modélisation, ce problème se traduit par la nécessité de choisir une loi de frottement au moins partiellement spéculative. De nombreuses formulations ont été proposées sur la base de considérations théoriques et/ou d'observations de terrain (Cassassa et al., 1989 ; Cassassa et al., 1990).

Parmi elles, la proposition de Voellmy (1955) revêt un caractère particulier du fait de son ancienneté et du grand succès qu'elle a rencontré depuis son apparition. Une analogie hydraulique a conduit cet auteur précurseur à proposer une loi de frottement associant un coefficient de frottement coulombien μ_{voel} et un terme « turbulent » dépendant d'un second coefficient ξ_{voel} et du carré de la vitesse. Il s'agit en fait de considérer le frottement comme un développement limité en vitesse à l'ordre 2 en négligeant le terme « laminaire », c'est-à-dire le terme dépendant linéairement de la vitesse. On y reviendra au chapitre 3 (eq. 3.5). L'interprétation traditionnelle de la force de frottement de Voellmy fait correspondre le coefficient μ_{voel} à la qualité de la neige au moment de l'événement (valeurs faibles pour de la neige froide et sèche, plus forte pour de la neige transformée et/ou humide), et le coefficient ξ_{voel} à la rugosité du couloir (valeur forte pour un couloir pentu et lisse, plus faible pour un couloir moins pentu et/ou boisé). Salm (1993) a montré que cette relation possédait une justification théorique partielle si la surface du couloir était assimilée à un ensemble d'hémisphères. Plus récemment, des mesures de profil vertical de vitesse ont également apporté des arguments en faveur de cette loi de frottement (Dent, 1993 ; Bouchet et al., 2004).

Néanmoins, la question du choix de la loi de frottement reste ouverte au débat et d'autres formulations sont toujours utilisées. Dans le modèle de Perla et ses dérivés, le terme turbulent dépend de la masse de neige en plus du carré de la vitesse et de la hauteur de l'écoulement (Perla et al., 1980 ; Perla et al 1984). Le terme « turbulent » est parfois considéré comme négligeable, conduisant à un frottement constant de type coulombien (Savage et Hutter, 1991). Au contraire, le terme « laminaire » est parfois considéré comme non négligeable, voire prépondérant, ce qui conduit à l'utilisation d'une loi de frottement de type granulaire (Nishimura et Maeno, 1989).

- *Calcul de la pression*

La pression est une variable cruciale pour le zonage et le dimensionnement des ouvrages de protection. Pourtant, elle n'est une sortie directe d'aucun modèle de propagation et son calcul à partir des autres variables dynamiques reste un problème très délicat. Longtemps, la pratique courante en ingénierie a été d'utiliser les lois de l'hydrodynamique pour calculer la pression théorique en écoulement libre à partir de la masse volumique et de la vitesse de l'écoulement sans prendre en compte la rhéologie du matériau. Cependant, des mesures expérimentales convergentes sur des sites avalancheux documentés (Sovilla et al., 2007) et sur des expériences de laboratoires utilisant des matériaux granulaires sont récemment venus remettre en cause cette pratique. Il a en effet été montré que la pression en écoulement libre sous-estime la pression d'impact subie par un obstacle pour des écoulements caractérisés par des nombres de Froude faibles. Des approches théoriques commencent à confirmer ces résultats expérimentaux. Ainsi Naaim (2007) a-t-il obtenu par analogie avec un fluide visqueux une formulation de la pression d'impact compatible avec les observations, c'est-à-dire prédisant de fortes pressions pour les écoulements peu dynamiques.

2.1.3. Les données historiques

2.1.3.1. Intérêt et disponibilité des données historiques

Les données obtenues sur les sites instrumentés réels sont irremplaçables car elles permettent d'obtenir une information très abondante et détaillée sur quelques événements. En contrepartie, elles sont coûteuses et par conséquent rares. Les expériences de laboratoire sur la neige ou les matériaux granulaires sont évidemment plus faciles à répéter dans des conditions contrôlables mais elles ne correspondent pas complètement à de « vraies » avalanches. Les données historiques recueillies systématiquement et selon un protocole standardisé concernent quant à elles un nombre restreint de variables (extension, date, ...etc.) mais concernent de vraies avalanches, et ce à l'échelle de régions

entières. La collecte systématique et à large échelle de données historiques apporte donc une information que l'on pourrait situer comme étant « à mi-chemin » entre les données issues de sites instrumentés réels et celles issues d'expériences sur modèles réduits. Outre leur intérêt pour la gestion du risque, elles permettent de mieux connaître le phénomène à une échelle spatiale assez large correspondant à celle des massifs montagneux et du forçage climatique.

Malgré l'émergence des modèles numériques, la mémoire des événements passés reste donc cruciale pour l'urbanisation des zones menacées et l'anticipation des événements futurs. Tout comme la connaissance naturaliste, l'origine du recueil d'information historique concernant les avalanches se confond avec les débuts de l'occupation sédentaire des vallées de montagne. Retrouver la trace des événements les plus anciens passe par un travail d'archives nécessitant l'interprétation de termes ayant pu disparaître du vocabulaire courant. Cœur et al. (1998) mentionnent ainsi des archives faisant état d'événements avalancheux au 17^{ème} et 18^{ème} siècle sur la commune d'Allemont (Isère, France). De tels documents sont utiles pour une expertise ponctuelle. Cependant, pour la prédétermination à l'échelle d'une vaste zone, des données bien plus exhaustives et abondantes sont indispensables afin de pouvoir organiser l'urbanisation de façon systématisée.

Dans la plupart des pays européens exposés au risque d'avalanche, des bases de données avalancheuses ont donc été constituées à des fins de prédétermination. Elles ont été inventoriées et comparées dans le cadre du projet européen SAME (Burnet et Marti, 1998). Une base de données d'une centaine d'événements extrêmes bien documentés a également été construite dans le cadre du projet européen CADZIE (Domaas et al., 2002). En Amérique du Nord, des bases de données avalancheuses existent également (McClung et al., 1989), mais elles sont essentiellement utilisées pour la prévision et la sécurisation des routes et des domaines skiables car les besoins en matière de prédétermination sont moins aigus du fait d'une pression démographique moindre. McCollister et al. (2003) mentionnent par exemple une base de données de 10 000 événements utilisée pour la mise en œuvre d'une méthode de prédiction de type « plus proches voisins » à Jackson Hole, Wyoming. Quelques bases de données avalancheuses ont été entièrement publiées, en particulier dans les pays du nord de l'Europe (Bakkehoi et al. 1983 en Norvège; Johanesson 1998 en Islande), mais le plus souvent chaque pays garde jalousement les informations à sa disposition.

En France, l'information historique sur l'aléa avalancheux repose sur deux bases de données anciennes : une chronique enregistrant les événements successifs se produisant sur des sites sélectionnés, l'Enquête Permanente sur les Avalanches (EPA), et une carte d'avalanches couvrant des zones plus larges avec un objectif d'exhaustivité, la Carte de Localisation des Phénomènes d'Avalanche (CLPA). EPA et CLPA offrent donc des visons temporelle et spatiale complémentaires d'un même phénomène. A noter que les avalanches majeures sont également enregistrées dans la base « événements » du service de Restauration des Terrains en Montagne (RTM) qui s'intéresse à l'ensemble des phénomènes dommageables en zone de montagne.

Le fort développement de la collecte de la donnée historique en France est intimement lié aux choix réglementaires qui ont été progressivement effectués. Les zones d'aléa ont été introduites dans la législation en 1974 et, dès le départ, le rôle central des bases de données avalancheuses pour leur élaboration a été souligné (De Crecy, 1980 ; Cemagref division Nivologie, 1983). Depuis, la législation s'est renforcée avec notamment l'intégration des PZEA (Plans de Zones Exposées aux Avalanches) dans les PER (Plans d'Exposition aux Risques naturels prévisibles) puis les PPR (MATE, 1997) et l'affirmation plus nette des prescriptions urbanistiques en zones d'aléa fort et moyen. Néanmoins, son esprit ne s'est pas modifié et l'information historique prime toujours sur l'emploi de méthodes systématisées de type "directive Suisse" (Jarry et Ancey, 2006). De plus, le système judiciaire français n'accepte pas, lorsqu'il est consulté après une catastrophe, qu'une donnée disponible quelconque ait été négligée s'il est avéré que sa prise en compte aurait permis une réduction des dommages.

2.1.3.2. L'information historique en France

L'EPA a été initiée dès le début du 20^{ème} siècle par l'inspecteur général des eaux et forêts Mougin (Mougin, 1922). Elle a d'abord concerné la Savoie et la Haute-Savoie puis a été progressivement étendue à l'ensemble des Alpes et des Pyrénées. La collecte de données est assurée selon un protocole précis (Cemagref et al., 2003) par des agents de l'Office National des Forêts (ONF) appelés les « observateurs », chaque observateur étant responsable d'un nombre précis de sites (Nuiry, 2004). Les observations recueillies sont tout d'abord consignées par l'observateur dans un carnet puis centralisées

dans le fichier EPA proprement dit dont la gestion est assurée par le Cemagref de Grenoble (CTGREF division Nivologie, 1974 ; Strazzeri et Manche, 1998). Depuis le début de l'observation, environ 80 000 événements ont été enregistrés sur près de 5 800 sites, avec de 1 à plus de 100 événements par site. Les données disponibles ainsi que leur utilisation seront présentées à la section 2.1.3.4.

La CLPA a été créée sur recommandation de la mission interministérielle d'étude sur la sécurité des stations de montagne réunie suite à une catastrophe majeure survenue à Val d'Isère en 1970 (Saunier, 1970). Son objectif est de proposer une information cartographique à l'échelle du territoire entier afin de servir de base de discussion pour l'aménagement du territoire et la construction d'ouvrages de protection. Elle est constituée d'un fond topographique à l'échelle 1/25 000 sur lequel sont reportées pour tous les sites avalancheux les limites extrêmes des avalanches connues. La responsabilité de la constitution de cette carte est confiée au Cemagref de Grenoble et à l'ONF depuis 1972 (CTGREF division Nivologie, 1975). Pour l'élaboration de la carte, la collecte d'information s'effectue selon deux méthodes complémentaires (Borrel, 1993 ; Borrel, 1994). D'un côté, une enquête de terrain comprenant recueil de témoignages et dépouillement des archives existantes (photographies, articles de journaux, expertises,... etc.) est conduite. L'ensemble des informations est synthétisé sur des fiches et l'extension maximale de l'avalanche est tracée en magenta sur un fond topographique et numérotée. Parallèlement, une analyse de stéréophotographies aériennes est effectuée afin de détecter les marques du passage des avalanches sur la végétation et le relief. Les surfaces ainsi obtenues sont reportées en orange sur le fond topographique. Les ouvrages de protection existants y sont également reportés (Figure 2.4). Depuis son origine, la CLPA constitue l'outil de base pour l'établissement des cartes d'aléas et est appréciée par tous les praticiens pour sa précision et son exhaustivité. Une étude a évalué à 10 000 le nombre de constructions situées dans les emprises CLPA (Cabos, 2003 ; Cabos et Bélanger, 2003).

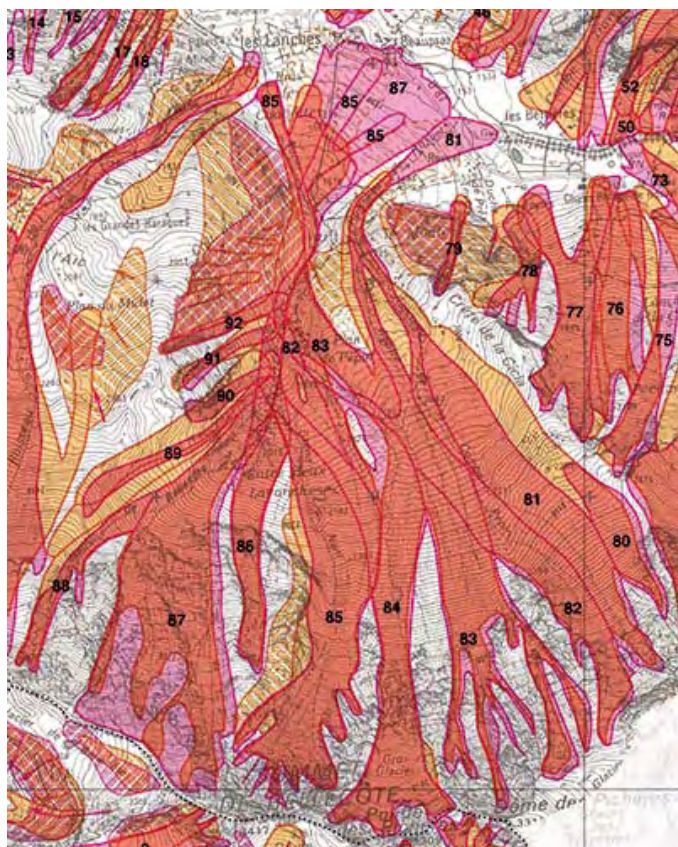


Figure 2.4 : Extrait de la CLPA

« Les Lanches », commune de Peseix-Nancroix, Savoie, France.

En noir, l'identifiant CLPA. En magenta, les emprises issues de témoignages. En orange, les emprises issues de l'analyse de stéréophotographies et d'enquêtes de terrain. En rouge, les emprises pour lesquelles les deux types d'information sont concordantes.

2.1.3.3. La rénovation des bases de données avalancheuses

Suite à la catastrophe de Montroc, la rénovation des bases de données françaises a été effectuée entre 2001 et 2006 au travers d'un projet de 5 ans dit « de toilettage ». Celui-ci a été financé par le Ministère chargé de l'environnement ainsi que certaines collectivités locales et conduit par l'ONF et le Cemagref. Son objectif était la rénovation tant quantitative que qualitative des bases de données afin d'améliorer la connaissance de l'aléa à l'échelle du territoire entier (Bélanger et Cassayre, 2004). De

plus, les différents publics concernés, c'est-à-dire bureaux d'études spécialisés, collectivités locales, particuliers, etc., devaient avoir un accès plus facile aux données. Les protocoles d'observation ont donc été entièrement revus et une grande quantité d'informations nouvelles a été acquise sur le terrain. L'ensemble des informations a été numérisé et le site internet www.avalanches.fr a été développé de façon à permettre un accès rapide à la donnée ainsi que la diffusion d'une documentation technique adaptée.

- *Rénovation de l'EPA*

La rénovation de l'EPA s'est essentiellement traduite par une mise à jour du réseau d'observation et de la liste des sites suivis ainsi que par une redéfinition plus rigoureuse du protocole d'observation. Chaque site présent dans la chronique a de ce fait été étudié individuellement par une équipe de trois agents composée de l'observateur, d'un technicien de l'ONF et d'un technicien spécialisé du Cemagref. Il s'agissait de décider si l'observation devait être poursuivie ou arrêtée en fonction du nombre et de la qualité des données collectées depuis le début de l'observation ainsi que du niveau d'exposition d'éventuels enjeux. Quelques nouveaux sites présentant de forts enjeux ont également été ajoutés à la chronique, de sorte qu'à présent l'EPA compte 3 839 sites en cours d'observation. Pour chacun de ces sites, une description topographique incluant carte et photographie a été effectuée. La photographie permet à l'observateur de localiser précisément le site sur la carte (Figure 2.5). La carte du site est tracée sur un fond topographique noir et blanc à l'échelle 1/25 000^{ème}. Elle comprend l'emprise du site, le chenal d'écoulement principal, un seuil d'observation, un seuil d'alerte et un point d'observation (Figure 2.6). L'observateur doit systématiquement renseigner les événements dont l'altitude d'arrêt est inférieure au seuil d'observation. Il peut également renseigner des événements plus modestes s'il considère que l'information apportée présente un intérêt quelconque (victimes, trajectoire ou volume atypique par exemple). Le seuil d'alerte, atteint par des avalanches extrêmes seulement, est utilisé pour le déclenchement d'une procédure de sécurité visant à enregistrer l'événement dans la base de données du service RTM. Le point d'observation est un endroit sûr même par conditions météorologiques difficiles et offrant une bonne visibilité sur le site. L'ensemble des informations cartographiques, textuelles et quantitatives a été numérisé et est accessible via www.avalanches.fr.

- *Rénovation de la CLPA*

Toutes les données ayant servi à l'élaboration de la CLPA depuis ses débuts (témoignages, archives, ... etc.) ont été numérisées et regroupées dans une base de données. La photo-interprétation a été complétée par des analyses de terrain. De nouvelles zones ont été cartographiées, de sorte qu'au total environ 800 000 ha sont couverts à la fin 2006, dont environ 30 000 ha d'emprises avalancheuses. Comme pour l'EPA, toutes les données cartographiques ont été numérisées, géoréférencées et converties au format SIG sous la forme de deux espaces de travail correspondant aux Alpes et aux Pyrénées et accessibles via www.avalanches.fr.

Pour chaque massif montagneux, une notice a été rédigée avec l'aide de Météo France. Celle-ci inclut une description des caractéristiques météorologiques, topographiques, géomorphologiques et environnementales du massif ainsi qu'une description des principaux événements historiques.

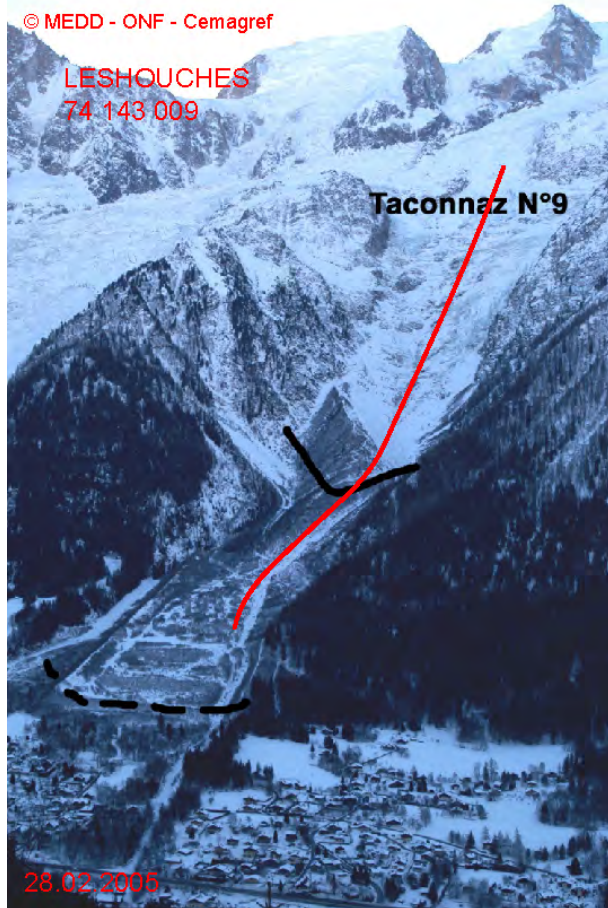
Enfin, un processus de mise à jour régulier de la carte a été mis en place. En effet, depuis sa première édition, de nouveaux événements se sont produits dépassant les emprises historiques. De plus, de "vieux" événements ont parfois été redécouverts. Chaque année, 10% de la superficie totale de la carte sont donc revus, ce qui conduit à une mise à jour décennale de l'ensemble de la carte. Annuellement, quelques gros événements peuvent également entraîner une mise à jour locale de la carte.

- *Création de la base Sites Sensibles d'Avalanche*

Durant le projet de rénovation des bases de données avalanches, une enquête nommée Sites Sensibles d'Avalanche (SSA) a également été conduite. Son principe était d'effectuer une évaluation homogène du niveau de risque des couloirs avalancheux français afin d'identifier rigoureusement les sites présentant les niveaux de risque les plus élevés et d'y concentrer les efforts de mitigation. Une méthode de classification combinant aléa et vulnérabilité a ainsi été développée (Rapin et al., 2004). L'enquête est basée sur 38 questions concernant les enjeux, la morphologie du site, l'information historique disponible et la météorologie locale, avec à chaque fois un nombre fini de réponses possibles. L'ensemble des réponses est pondéré et combiné de façon à ranger le site étudié dans la

catégorie risque faible, moyen ou fort. Pour un technicien connaissant la méthode et le site, compléter le questionnaire est relativement simple et rapide (pas plus d'une demi journée). L'enquête a démarré par un inventaire de 1 200 sites présumés à risque en 2002. Depuis 2004, la méthode a été appliquée à 600 d'entre eux (Rapin et al. 2006). Les résultats de l'enquête ont été transmis aux autorités locales en charge des sites et regroupés dans une base de données gérée par le Cemagref et accessible via www.avalanches.fr. Celle-ci, intermédiaire entre données brutes et résultats de traitements, présente l'intérêt d'intégrer un important volet vulnérabilité en complément de la vision brute de l'aléa donnée par l'EPA et/ou la CLPA.

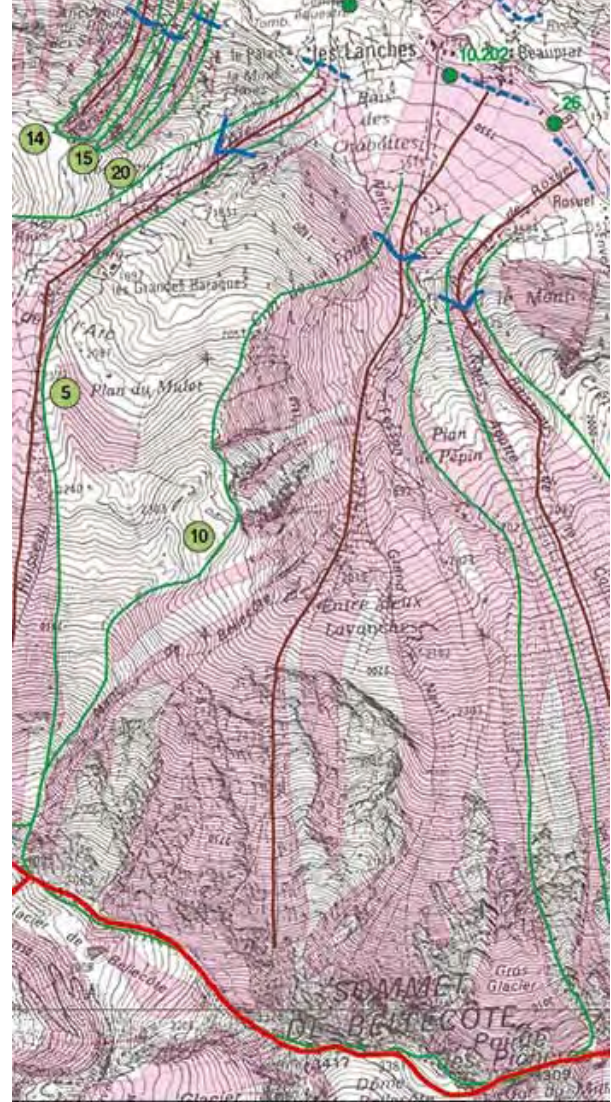
Figure 2.5 : Exemple d'un site EPA



Couloir de Taconnaz, commune des Houches, Haute-Savoie, France.

Le tracé rouge correspond au chenal d'écoulement principal. Le tracé noir correspond au seuil d'observation. Le tracé noir pointillé correspond au seuil d'alerte. Il s'agit pour ce couloir exceptionnel de la digue frontale de l'ouvrage paravalanche installé en début de zone d'arrêt. Cependant, chenal d'écoulement et seuils ne figurent en général pas sur les photographies de site "officielles".

Figure 2.6 : Site EPA sur fond topographique



Site "les Lanches" sur fond topographique, commune de Peseix-Nancroix, Savoie, France.

Pour chaque site EPA, son identifiant et le point d'observation apparaissent en vert, le chenal d'écoulement principal en brun, les seuils d'observation et d'alerte en bleu. Les emprises de CLPA apparaissent en rosé sur fond topographique.

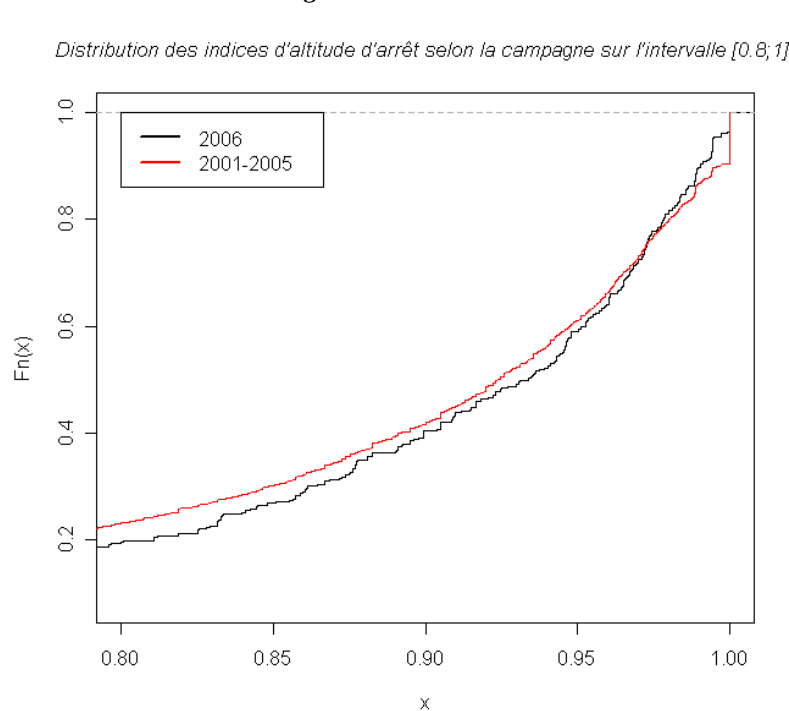
2.1.3.4. Zoom sur les données EPA

- Informations disponibles

Le fichier EPA contient les dates des événements avalancheux observés (Figure 2.10), des informations quantitatives (altitudes de départ et d'arrivée et dimensions du dépôt) ainsi que des informations qualitatives enregistrées sous forme catégorielle souvent binaire (Jamard et al., 2002).

Pour la modélisation des avalanches, les variables quantitatives sont particulièrement utiles à la calibration locale des modèles de propagation, en premier lieu les altitudes de départ et d'arrêt. A titre d'exemple, la Figure 2.7 résume les altitudes d'arrêt enregistrées au cours de la campagne 2006, c'est-à-dire de l'hiver 2006/2007, sous la forme de la fonction de répartition d'un indice permettant la comparaison inter-site grâce à la normalisation par l'altitude minimale du site.

Figure 2.7 : Altitude d'arrêt des avalanches sur sites EPA



Comparaison de l'hiver 2006/07 avec la moyenne des autres campagnes post toilettage.

L'indice

$$x = \frac{1}{e} \exp\left(1 - \frac{z_{stop} - z_{min}}{z_{min}}\right)$$

est égal à 1 si l'altitude minimale z_{min} du site est atteinte et compris entre 0 et 1 sinon, avec une valeur d'autant plus élevée que l'altitude d'arrivée z_{stop} est basse.

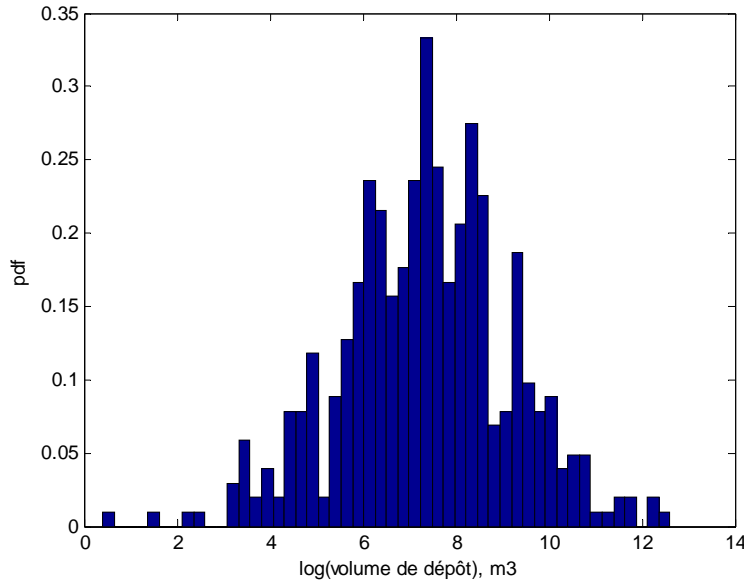
La campagne 2006/07 est significativement différente (au sens du test de Kolmogorov-Smirnov avec un niveau de confiance de 95%) de la moyenne des autres campagnes post toilettage avec à la fois moins d'événements présentant une altitude d'arrêt très supérieure à l'altitude minimale du site et moins d'événements atteignant le fond de vallée.

Les volumes de dépôt fournissent une autre mesure de l'intensité des événements observés. Ils posent cependant des problèmes spécifiques rendant leur utilisation systématique pour la calibration de modèles numériques plus difficile que celle des altitudes. Premièrement, ce ne sont pas les volumes mais les trois dimensions (longueur, largeur et hauteur) du dépôt qui sont enregistrées, de sorte que le calcul du volume de dépôt nécessite l'application d'un coefficient de forme traduisant le fait qu'un dépôt d'avalanche n'a généralement pas une forme de parallélépipède rectangle. En expertise, un coefficient de $\frac{1}{2}$ correspondant à une forme en « triangle camembert » est généralement utilisé. La Figure 2.8 résume, conditionnellement à cette hypothèse, les volumes de dépôt enregistrés au cours de la campagne 2006. Etant donnée la forte asymétrie de la distribution, une présentation en coordonnées logarithmiques a été privilégiée. Deuxièmement, pour les événements antérieurs au toilettage, les dimensions du dépôt ne figurent que sur certains avis datant principalement d'avant 1970, et elles n'ont généralement pas été numérisées. Elles doivent donc être recherchées « manuellement » dans les carnets forestiers, sans aucune garantie de leur existence.

Les variables catégorielles permettent de mieux connaître l'événement en termes de conditions météorologiques au moment du déclenchement, de description sommaire des caractéristiques de l'événement (nature de la zone de départ, mode de propagation, etc.) et de dégâts occasionnés. Les conditions météorologiques de déclenchement (cumul de neige au cours des 3 jours précédant

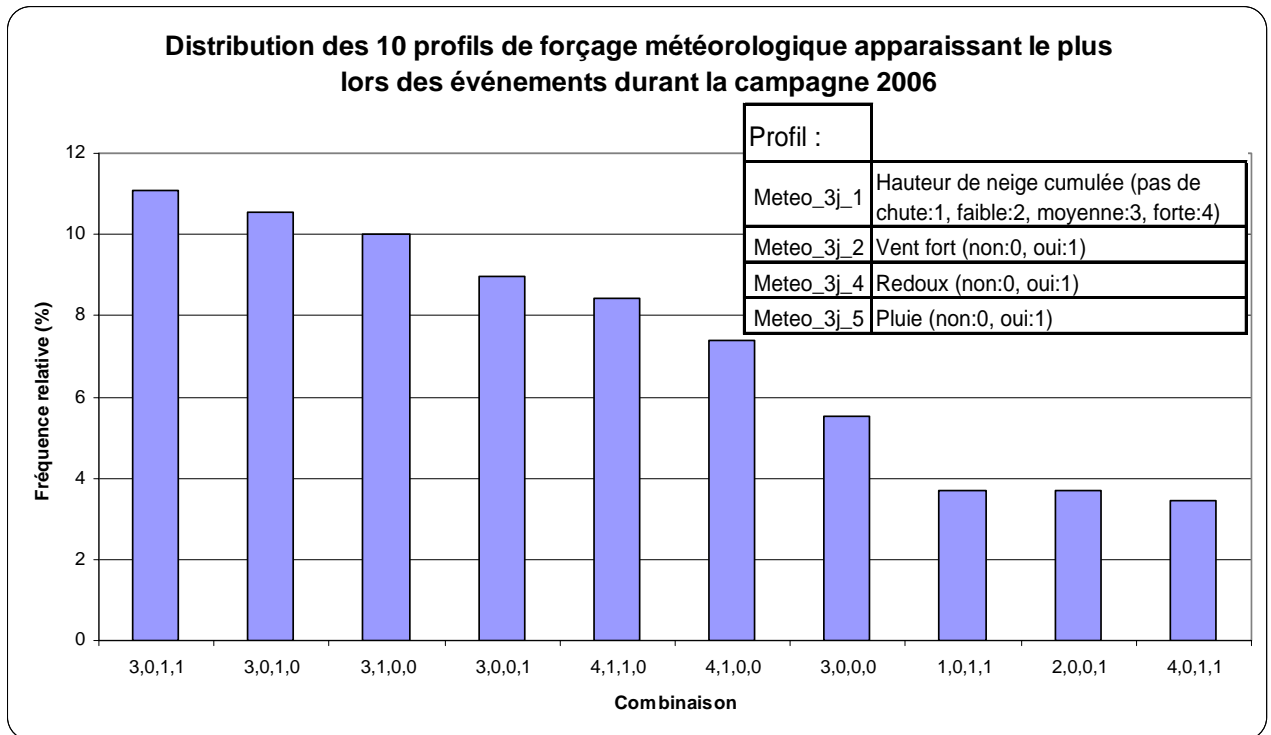
l'événement, force et direction du vent, présence ou absence de redoux et/ou de pluie) permettent par exemple de mieux comprendre les causes du déclenchement. Celles correspondant à la campagne 2006 sont présentées sous la forme d'un histogramme des combinaisons les plus fréquentes en Figure 2.9.

Figure 2.8 : Volumes de dépôt des avalanches sur sites EPA durant l'hiver 2006/07



Volumes disponibles pour 426 événements. 322 d'entre eux correspondent à des avalanches de faible dimension (moins de 5000 m3) et 98 d'entre eux à des avalanches de dimension moyenne (de 5000 à 100 000 m3). Seuls 6 événements de plus de 100 000 m3 ont donc été enregistrés, avec un maximum annuel de 300 000 m3.

Figure 2.9 : Conditions météorologiques de déclenchement sur sites EPA durant l'hiver 2006/07



Les conditions météorologiques au cours des trois jours précédant le déclenchement sont enregistrées grâce à trois variables catégorielles : le cumul de neige avec 4 modalités et les trois autres, i.e. vent fort, redoux et pluie, sous forme de présence/absence.

La majorité des déclenchements se sont produits dans un contexte relativement homogène : chutes de neige moyennes et/ou fortes accompagnées de pluie et/ou de redoux et/ou de vent fort. Les événements ont été concentrés sur deux périodes temporelles très courtes correspondant à des épisodes de mauvais temps ayant affecté les Alpes et les Pyrénées de façon décalée. Par rapport à une campagne « moyenne », on ne retrouve pas les avalanches « de fonte » printanière se produisant par beau temps et sans chute de neige.

- *Qualité des données*

Etant donnée la quantité d'information très importante présente dans l'EPA ainsi que son protocole de collecte original, le problème de la qualité des données se pose de façon très aiguë (Garcia et Bélanger, 2002). Une première source d'erreur est liée à la localisation des sites : pour les données anciennes, celle-ci est en effet parfois incertaine car les noms de certains sites ont changé au cours du temps. Plus généralement, l'information cartographique n'était pas précise avant 1950, de sorte qu'un niveau d'incertitude assez élevé entoure les altitudes de départ et d'arrêt les plus anciennes du fait des références topographiques peu fiables dont disposaient les observateurs de l'époque.

Une autre source d'incertitude est liée à la non exhaustivité dans l'espace et dans le temps de la collecte d'information. L'activité avalancheuse recensée dans l'EPA est en effet le produit de 2 facteurs, l'activité avalancheuse intrinsèque des sites sélectionnés et l'effort d'observation sur ces sites. Or, l'effort d'observation n'a pas été constant partout durant les campagnes passées et les périodes où l'observation a été insuffisante ne sont que rarement connues avec précision. Il n'est de ce fait pas possible de caractériser avec l'EPA l'activité avalancheuse intrinsèque ainsi que ses variations dans l'espace et dans le temps sans examen préalable des séries de données, ce qui rend en particulier difficile l'étude de la stationnarité des occurrences.

Les valeurs enregistrées sont également entachées d'une imprécision irréductible due à la difficulté de l'observation : limites en termes de visibilité, difficulté à évaluer l'altitude de départ depuis le point d'observation, difficulté de l'estimation visuelle des dimensions de dépôt, etc. De plus, lorsque les valeurs ont été bien enregistrées, des erreurs ont pu être introduites dans les fichiers au cours des opérations d'archivage successives.

Malgré ses limites, l'EPA constitue du fait de son ancienneté et de sa collecte généralisée à l'échelle du territoire une source d'information de grande valeur pour la prise en compte du risque avalanche en France. Les irrégularités majeures à l'échelle de la base entière telle l'absence d'observation durant les deux guerres mondiales sont en effet relativement rares (Zambito et al., 2002). De plus, sous réserve de ne considérer que les régions où les événements sont nombreux et enregistrés depuis longtemps et de se placer à des échelles de temps et d'espace assez grandes pour s'affranchir des variations de la qualité de l'observation locale, l'EPA présente tout de même une certaine régularité et une bonne cohérence vis-à-vis des analyses expertes. La fin du toilettage en 2007 et l'engagement récent d'une démarche qualité vont d'ailleurs dans le sens d'un renforcement de la crédibilité de l'information.

- *Utilisation pour la prévision*

L'EPA n'est pas utilisée en France pour la prévision du risque de déclenchement au jour le jour car une approche physique effectuée par Météo France est privilégiée (CEN, 2000). Celle-ci repose sur l'acquisition en temps réel de données météorologiques, la modélisation numérique de l'évolution du manteau neigeux à l'échelle de massifs montagneux et une synthèse experte effectuée par les prévisionnistes. Toutefois, quelques travaux de recherche ont exploré l'étude des corrélations entre le forçage météorologique et l'activité avalancheuse enregistrée dans l'EPA. Grandjean (2000) ainsi que Jomelli et al. (2007) ont en particulier exploré l'utilisation de modèles de régression logistique pour expliquer les événements observés sur quelques sites bien documentés par les cumuls de neige des jours précédents le déclenchement. Les situations météorologiques favorables au déclenchement à l'échelle des massifs ont également été explorées, en particulier sous la forme de corrélations entre les événements observés et le niveau de risque affiché par Météo France (Marchal et al., 2003 ; CEN et Cemagref ETNA, 2006). Enfin Etienne (2006) a analysé le lien entre situations météorologiques synoptiques et déclenchements.

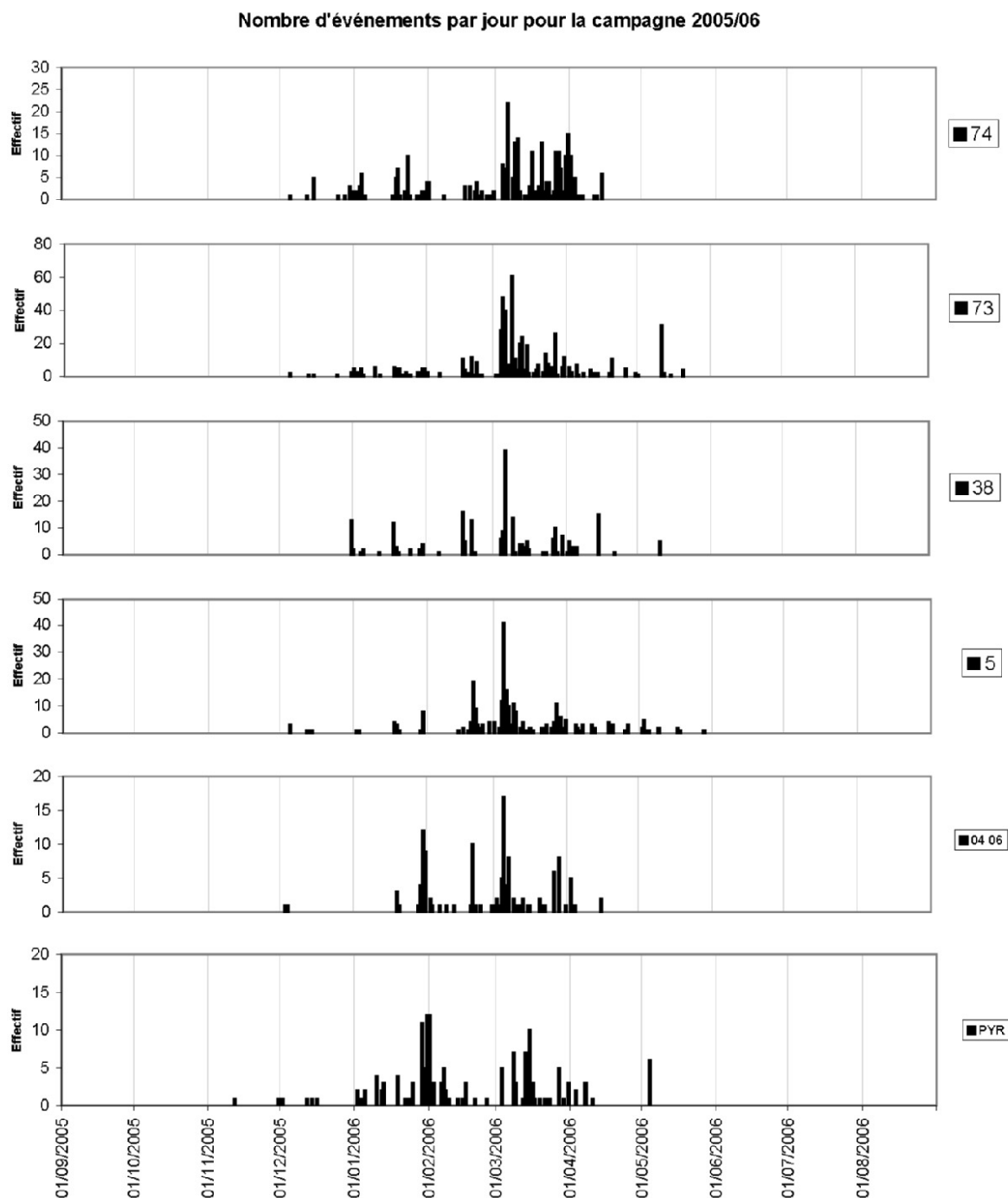
- *Utilisation pour la prédétermination*

La valorisation usuelle des données de l'EPA est la prédétermination locale. Les informations y figurant sont en effet très utiles pour évaluer la variabilité de l'aléa, ses caractéristiques extrêmes et les conditions d'occurrence des événements. Elles s'intègrent donc facilement, au même titre que les autres données historiques (CLPA et/ou SSA et/ou base RTM) et la connaissance naturaliste, dans la synthèse experte généralement utilisée pour dresser les cartes d'aléa (par exemple Cemagref ETNA, 2000 ; Cemagref ETNA, 2002).

Une autre valorisation naturelle des événements successifs enregistrés dans l'EPA pour la prédétermination est la mise en œuvre d'approches statistiques. Les « méthodes norvégiennes » qui seront présentées au paragraphe suivant ne sont pas très utilisées car il a été montré qu'elles n'étaient

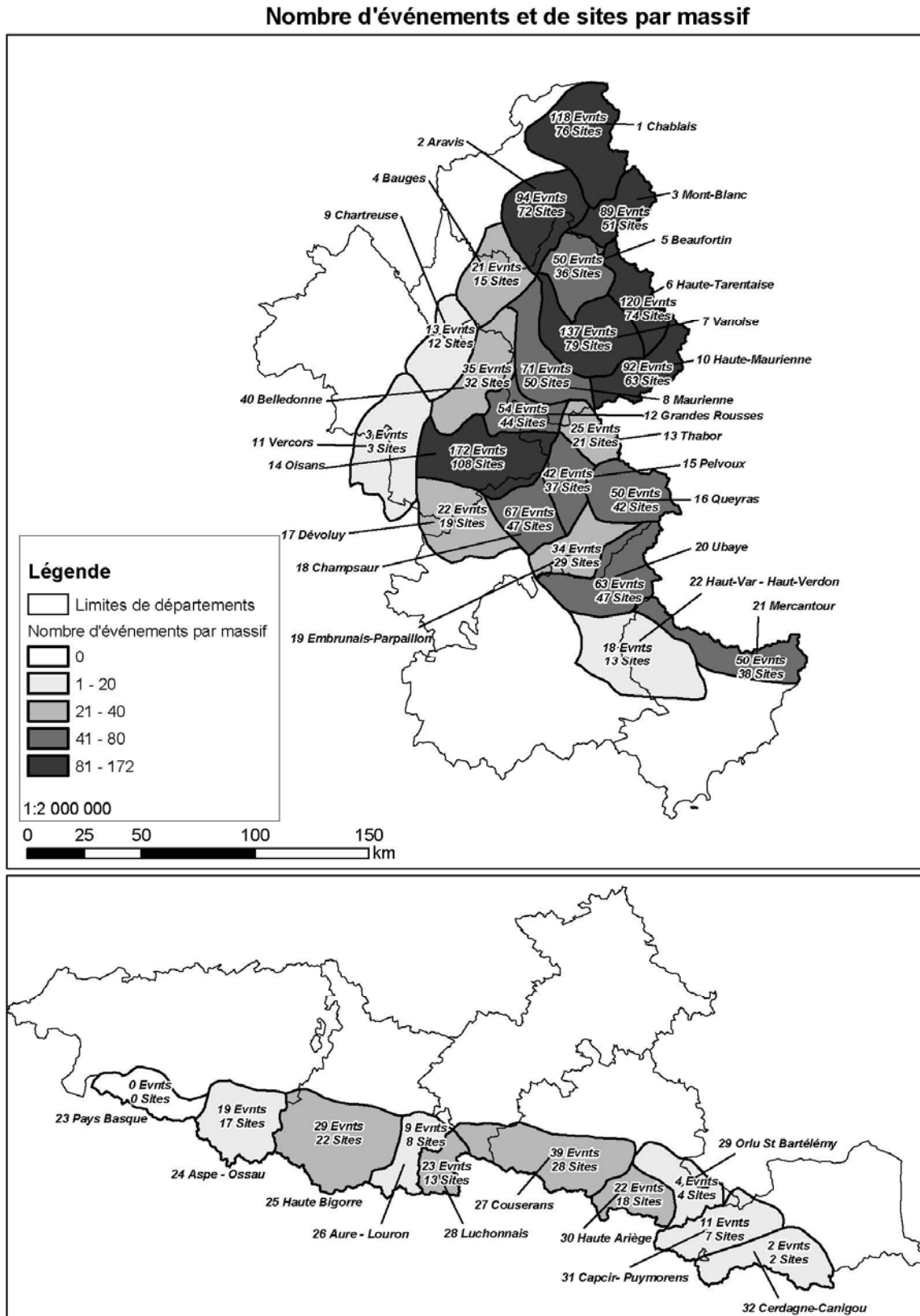
pas bien adaptées aux géométries complexes caractérisant les couloirs des Alpes françaises (Berthier, 1986 ; Adjel, 1995 ; Adjel, 1996). Les données EPA sont en revanche de plus en plus utilisées pour la calibration locale de modèles numériques (Naaim et al., 2004 ; Barbolini et al., 2000). Enfin, l'EPA est parfaitement adaptée aux récentes approches statistiques-dynamiques qui seront elles aussi présentées au paragraphe suivant en permettant de choisir des distributions d'entrées adaptées aux caractéristiques locales de l'aléa (Meunier et al., 2001; Meunier et Ancey, 2004).

Figure 2.10 : Distribution temporelle des avalanches sur sites EPA durant l'hiver 2005/06



74=Haute-Savoie; 73=Savoie; 38=Isère; 05=Hautes Alpes; 04=Alpes Maritimes ; 06=Alpes de Haute Provence ; Pyr= Pyrénées. Dans les quatre départements où il s'est produit le plus d'événements (i.e. 74, 73, 38 et 05), ceux-ci ont principalement eu lieu en mars et avril, avec un pic d'activité marqué durant la seconde semaine de mars. Dans les deux autres départements des Alpes et les Pyrénées, l'activité a été bien moins structurée dans le temps au cours de l'hiver. Quelques événements se sont produits tardivement en Savoie, Hautes Alpes et dans les Pyrénées.

Figure 2.11 : Distribution spatiale des avalanches sur sites EPA durant l'hiver 2005/06



La majeure partie de l'activité a été localisée dans les Alpes du Nord avec quatre massifs enregistrant plus de 100 événements : l'Oisans, la Vanoise, la Haute Tarentaise et le Chablais. Tous les massifs des Alpes ont enregistré au moins dix événements, à l'exception du Vercors. Dans les Pyrénées, l'activité a été nulle à très faible (moins de 5 événements) dans trois massifs : Pays Basque, Cerdagne-Canigou et Orli-St-Barthélémy, faible (5 à 10 événements) dans le massif Aure-Louron, et notable (de 11 à 39 avis) dans les 6 autres massifs plutôt situés au centre de la chaîne.

- *Analyses du phénomène à l'échelle "régionale"*

Les données de l'EPA présentent également un intérêt en termes d'apprentissage sur le phénomène. En effet, même si les sites EPA ne sont pas des index traduisant parfaitement l'activité avalancheuse à l'échelle de la France, ils sont néanmoins suffisamment nombreux pour en donner un bon aperçu et permettre des analyses spatio-temporelles. Les Figures 2.7 à 2.9 ont déjà donné des exemples d'utilisation de l'EPA pour la caractérisation de l'intensité des événements d'un hiver. La Figure 2.10 présente en outre la répartition temporelle des événements durant le même hiver pour 6 départements ou groupes de départements tandis que la Figure 2.11 propose la distribution spatiale des événements de l'hiver 2005/2006 à l'échelle des massifs utilisés pour la prévision du risque d'avalanche (Cemagref ETNA, 2006). Ce type d'analyse descriptive permet d'établir un bilan annuel (par exemple Cemagref ETNA, 2004 ; Cemagref ETNA, 2005) et de mettre en évidence une activité exceptionnelle à une échelle de temps et d'espace donnée. Ainsi Cosse et al. (2005) ont montré que durant quelques jours de janvier 2004, un nombre particulièrement élevé d'événements avait été enregistré dans le sud-est des Alpes en lien avec un forçage météorologique de type "retour d'est" particulièrement intense. Quelques études statistiques basées sur des données plus anciennes ont également été effectuées. Jamard et Bélanger (2002) ont ainsi exploré les similarités de fonctionnement des différents couloirs en termes de dates de déclenchement. Les pics d'activité, appelés communément crues avalancheuses, ont également été étudiés afin de préciser leurs périodes d'occurrence privilégiées et de les caractériser en termes fréquentiels (Diongue, 2006 ; Diongue et Eckert, 2006). De telles approches bénéficient de la grande quantité de données disponibles qui permet une estimation robuste mais pâtissent bien évidemment des limites des données anciennes en termes de qualité.

2.2. Se protéger des avalanches

2.2.1. Méthodes de prédétermination existantes

2.2.1.1. Les méthodes norvégiennes

Les méthodes statistiques dites « norvégiennes » constituent historiquement les premières approches permettant d'affecter de façon systématique des périodes de retour aux avalanches. Leur principe est d'expliquer la distance d'arrêt, c'est-à-dire la distance parcourue mesurée le long d'un axe horizontal, à l'aide de covariables topographiques (Lied et Bakkehoi, 1980 ; Bakkehoi et al., 1983) ou de caler directement sur les distances d'arrêt normalisées observées un modèle paramétrique issu de la théorie des valeurs extrêmes (McClung et Lied, 1987 ; Keylock, 2005) et de considérer les quantiles élevés de la distribution obtenue (Figure 2.12).

Le manque d'observations locales est généralement compensé par une hypothèse d'homogénéité régionale : tous les événements ayant affecté un même massif montagneux sont regroupés au sein d'un unique « échantillon » et l'inférence est effectuée une fois pour toute (McClung, 2000), moyennant éventuellement une transposition sur un couloir modèle moyen (Jonasson et al., 1999 ; Sigurdsson et al., 1998). La variabilité intra couloir n'est alors pas distinguée de la variabilité inter couloir.

Ces méthodes, développées comme leur nom l'indique dans les pays d'Europe du Nord, sont bien adaptées au profil topographique des couloirs scandinaves, majoritairement de type parabolique. Le comportement de la queue de la distribution des distances d'arrêt y est en effet « régulier » de sorte que les valeurs de projet ainsi calculées sont réalistes (Johannesson, 1998). Les méthodes norvégiennes ont également été appliquées avec succès en Amérique du Nord (McClung et al., 1989) et au Japon (Fujizawa et al., 1993). Mais, comme signalé précédemment, leur utilisation dans les Alpes est plus problématique du fait de la topographie généralement moins régulière des couloirs alpins rendant l'extrapolation au-delà de la plus forte valeur observée fallacieuse.

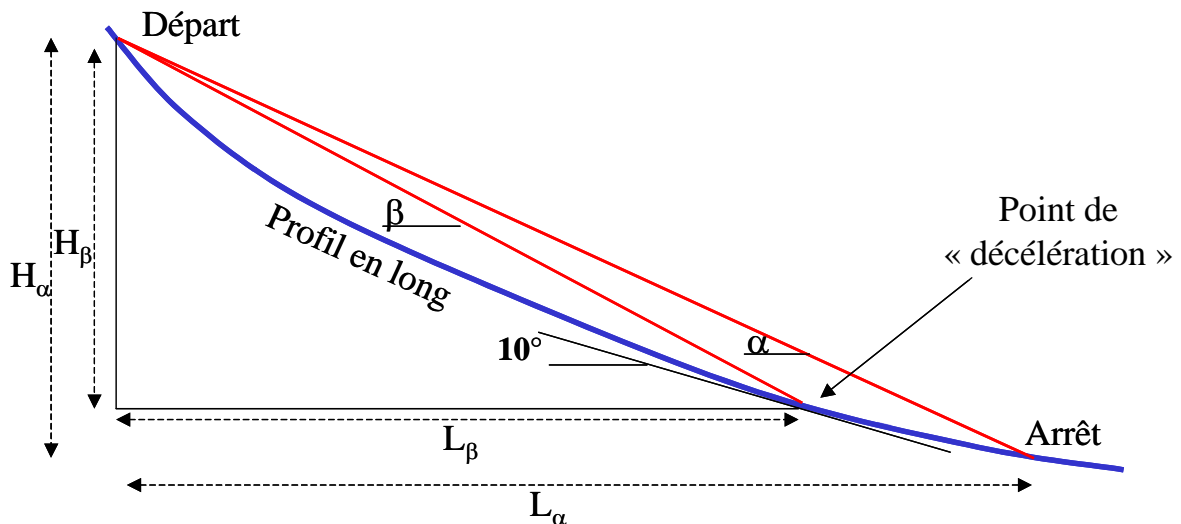
2.2.1.2. La méthode suisse et ses dérivés

La directive suisse (Salm et al., 1990) constitue l'alternative classique aux méthodes norvégiennes pour l'obtention standardisée d'avalanches de période de retour élevées. Son principe est de considérer des hauteurs de neige de projet correspondant à des cumuls de précipitation trentennaux et tricentennaux puis de propager ces volumes initiaux avec un modèle dynamique dont les coefficients de frottements sont tabulés en fonction des caractéristiques du couloir (altitude, pente, rugosité de surface...etc.). Pour chacune des deux hauteurs de neige, une distance d'arrêt et un profil longitudinal

de vitesse sont obtenus. Les pressions exercées par l'écoulement sont déduites de l'application des lois de l'hydraulique à surface libre. La distinction des trois zones d'aléa s'effectue ensuite grâce à une combinaison des distances d'arrêt et des pressions des deux événements de référence considérés. Originellement, le calcul s'effectuait avec un modèle de propagation de type bloc glissant incluant une loi de frottement de Voellmy. Le développement de modèles de propagation de type fluides a toutefois conduit à la mise à jour de la méthode suisse, avec le remplacement de la description « centre de masse » par un modèle basé sur les équations de Saint-Venant (Bartelt et al., 1999).

La directive suisse constitue une approche claire et standardisée permettant d'appréhender ce que peut être une avalanche rare (propagation d'un cumul de neige trentennal) ou exceptionnelle (propagation d'un cumul de neige tricentennal). Elle est de ce fait utilisée de façon systématique pour le zonage et le dimensionnement d'ouvrages en Suisse (Bardou, 2006) mais aussi dans plusieurs autres pays européens tels l'Autriche (Hopf, 1998) avec de légères variations. Mais elle ne résout pas le problème de la caractérisation probabiliste des événements de référence puisque le lien entre la période de retour de la chute de neige et celle de la distance d'arrêt et/ou de la pression de référence n'est pas univoque (Ancey et al., 2004). De plus, la probabilité de déclenchement n'est pas prise en compte, de sorte que les périodes de retour calculées peuvent être sous estimées si l'activité du couloir est très forte ou au contraire surestimées si l'activité est très faible.

Figure 2.12 : Principe des méthodes Norvégiennes



Description topographique du couloir avalanchueux d'après Meunier et al. (2004).

La méthode dite « alpha-beta » consiste en une régression de la distance d'arrêt maximale caractérisée par l'angle α sur le point caractérisé par l'angle β où la pente atteint 10 degrés.

La méthode du rapport d'arrêt consiste en l'ajustement d'une loi de probabilité sur le rapport

$$r_{\alpha/\beta} = \frac{L_{\alpha} - L_{\beta}}{L_{\alpha}}$$

2.2.1.3. Les méthodes statistiques-dynamiques et l'approche conceptuelle

A partir de 2001, des méthodes de simulation de type Monte Carlo ont été proposées pour la prédétermination des avalanches par de nombreux auteurs (Meunier et al. 2001 ; Barbolini et Savi, 2001 ; Bozhinskiy et al., 2001 ; Chernouss et Fedorenko, 2001 ; Barbolini et Keylock, 2002 ; Barbolini et Cappabianca, 2003 ; Meunier et Ancey, 2004 ; Bozhinskiy, 2004). Ces approches partent toutes du constat que la distance d'arrêt n'est pas, contrairement à un débit de crue, une variable intrinsèque et qu'il est par conséquent nécessaire de simuler des avalanches fictives pour pouvoir déterminer un aléa de référence par interpolation. Leur principe est d'alimenter un modèle de propagation déterministe avec des données d'entrée tirées dans des distributions aléatoires supposées parfaitement identifiées et généralement indépendantes et d'obtenir ainsi la distribution des distances d'arrêt et des autres variables de sortie. Les divergences entre les approches proposées concernent l'origine des données (un seul site ou un massif entier) et les hypothèses de modélisation (loi de

frottement et échelle de description pour le modèle de propagation, formes paramétriques des observables). En particulier, la construction des distributions de probabilité des entrées est abordée par Bozhinskiy (1992) et Bozhinskiy et Andreyev (1993). De même, le problème du choix de l'opérateur de propagation en fonction de la facilité du calage est discuté par Meunier et al. (2003) et Ancey et Meunier (2004), tandis que la forme de la distribution de probabilité des coefficients de frottement du modèle de propagation est investiguée dans Ancey et al. (2003).

Outre l'obtention aisée d'avalanches de période de retour élevée, les applications de ce type d'approche en matière d'analyse de sensibilité et de quantification des incertitudes sont très nombreuses. Néanmoins, les résultats obtenus sont étroitement dépendants des hypothèses probabilistes effectuées, notamment le choix de la forme des distributions des données d'entrée et de leur éventuelles corrélations. De plus, les lois statistiques des variables d'entrée ne sont en général pas estimées sur des données provenant du site ou s'effectue la prédiction, soit parce que l'ensemble des données nécessaires (hauteurs de neige, surfaces de départ, coefficients de frottement, etc.) n'est pas disponible, soit parce que le problème de l'inférence n'est même pas abordé.

L'émergence de ces méthodes dites statistiques-dynamiques a conduit Meunier et al. (2004) à introduire dans le domaine des avalanches la notion de modèle conceptuel. Le modèle de propagation n'est en effet alors qu'un simple opérateur mathématique qui relie des variables d'entrée à des variables de sortie grâce à un calage et qui ne traduit donc pas nécessairement une réalité physique sous-jacente. Contrairement à d'autres domaines comme l'hydrologie où les « vrais » modèles conceptuels sont nombreux et couramment utilisés (modèles à réservoirs, réseaux de neurones, etc.), la dynamique des avalanches n'a pas vu pour l'instant émerger de modèles totalement empiriques spécifiques et ce sont souvent des modèles conçus à l'origine comme des modèles déterministes qui sont utilisés selon un angle conceptuel, et vice-versa. Ainsi en est-il du modèle historique de Voellmy (1955). Présenté à l'origine comme une vision simplifiée de la physique de l'avalanche, il a beaucoup servi en tant que modèle conceptuel calé sur un site donné, avant que la « directive suisse » (Salm et al., 1990) ne propose des valeurs de paramètres tabulées et par conséquent supposées extrapolables. La limite entre modèle conceptuel et modèle déterministe reste donc floue et un peu artificielle dans le domaine des avalanches.



Figure 2.13 : Filets paravalanches retenant la neige dans la zone de départ

*Protection active.
Ancien site expérimental du
Cemagref à Chamechaude, Isère,
France.*

Photo F.Valla/Cemagref.

2.2.2. Le génie paravalanche

2.2.2.1. Principe

Lorsqu'un manque d'espace disponible conduit à envisager l'urbanisation de zones potentiellement atteintes par des avalanches ou que, à la lumière de nouvelles investigations, il apparaît que des constructions ou des infrastructures que l'on pensait hors de danger sont exposées, une protection paravalanche peut être recherchée. Nous ne traiterons ici que des protections permanentes par opposition aux protections temporaires de type interdiction de circulation ou évacuation. Jusqu'à la moitié du vingtième siècle, la protection paravalanche fut essentiellement l'affaire des forestiers, le reboisement étant considéré comme la meilleure politique globale de prévention. Néanmoins, dès le

dix-huitième siècle, les habitants des zones exposées ont tenté de limiter les dégâts sur les bâtiments, ce qui a conduit à la construction des premiers ouvrages de génie civil à vocation paravalanche. Il convient de noter que sont classiquement distingués les ouvrages de protection active (Figure 2.13) des ouvrages de protection passive (Figure 2.14). Les protections actives sont installées dans les zones de départ et visent à empêcher le déclenchement des avalanches en modifiant localement la répartition de la neige (ouvrages à vent de type barrières à neige, toit-buse ou vire-vent) et/ou en fixant le manteau neigeux (râteliers, claie, filet plantation...etc.) et/ou encore en modifiant la rugosité du sol (banquette). Les protections passives sont au contraire installées dans la zone d'écoulement ou d'arrêt de l'avalanche et visent à modifier l'écoulement de l'avalanche en dissipant l'énergie (tas freineurs), en déviant la trajectoire (étrave, digue déviatrice, tunnel paravalanche) ou encore en stoppant l'avalanche (digue d'arrêt). On peut éventuellement également considérer dans cette catégorie les dispositions visant à renforcer les constructions (épaississement des murs,... etc.).



Figure 2.14 : Protection paravalanche individuelle

*Protection passive, Suisse.
Ouvrage déviateur en
maçonnerie moderne de forme
arrondie.
Le cours de l'avalanche est dévié
de part et d'autres de la maison.*

Photo F. Rapin/Cemagref.

2.2.2.2. Méthodes de dimensionnement

D'abord choisis et dimensionnés sur la base d'approches expertes de type naturaliste (MEDD et METLTM, 2004), les ouvrages de protection paravalanche ont ensuite bénéficié de procédures normalisées en association avec le développement des premiers modèles de propagation d'avalanches (Salm et al., 1990). A l'heure actuelle, des codes numériques ou des modèles réduits de type analogiques sont de plus en plus utilisés pour le dimensionnement. De plus, des approches économiques commencent à voir le jour afin de prendre en compte le coût de construction dans la décision. Ainsi, des analyses de type coût-bénéfice ont été proposées afin de comparer différents scénarii d'aménagement (Wilhelm, 1997 ; Wilhelm, 1999) ou pour *a posteriori* évaluer l'efficacité de certains ouvrages sur le long terme (Fuchs et McAlpin, 2005).

Néanmoins, bien que des progrès significatifs dans la compréhension des effets des ouvrages de protection passive sur les écoulement avalancheux aient été accomplis au cours des dernières années grâce à des expérimentations en sites réels (Lied et al., 2001), en modèle réduit (Primus et al., 2004 ; Faug, 2004) et grâce au recours à la modélisation numérique (Naaim et al., 2004, Sampl et al., 2004), le dimensionnement des ouvrages paravalanches reste délicat. En particulier, le déficit de connaissance en matière de rhéologie de la neige rend difficile le calcul de la pression d'impact subie par la structure.

2.2.3. Les approches de type risque

2.2.3.1. Principe

Face à la difficulté liée au calcul de périodes de retour dans le domaine des avalanches, des approches visant au calcul de la perte moyenne attendue ont récemment été développées. Puisque les enjeux sont alors pris en compte en plus de l'aléa, il ne s'agit plus à proprement parler de prédétermination, mais bien d'analyse de risque. Cependant, à la différence d'autres domaines d'ingénierie où le choix des

valeurs de projet s'effectue de façon routinière par la minimisation d'une fonction de coût quantifiant la perte moyenne attendue en fonction des valeurs prises par la variable de décision (Jordaan, 2005), les calculs de risque dans le domaine des avalanches restent pour l'instant limités au problème du zonage avec la définition d'un niveau de risque acceptable (Keylock et al., 1999 ; Chernouss et Fedorenko, 2001 ; Barbolini et al., 2004a ; Grêt-Regamey et Straub, 2006) et une définition du risque correspondant à une probabilité annuelle de décès (Jonasson et al., 1999). Encore exploratoire dans la plupart des pays européens, une telle approche a été incluse dans la législation Islandaise pour le zonage (Arnalds et al., 2004).

2.2.3.2. Choix d'une unité de référence

Pour une évaluation réaliste des coûts des dommages, le nombre et la nature des bâtiments exposés, leur mobilier ainsi que le nombre d'habitants et la fraction du temps qu'ils y passent doivent être quantifiés. Ce travail s'effectue généralement au travers d'une analyse socio-économique (Fuchs et al., 2004; Fuchs et Bründl, 2005; Fuchs et al., 2005 ; Keiler, 2004; Keiler et al., 2005) dont le point le plus critique est la prise en compte de la vie humaine. Les techniques assurantielles préconisent de donner une valeur économique à une vie humaine, ce qui présente l'avantage de pouvoir mener les calculs facilement. L'option inverse est de ne considérer que les vies humaines et d'exprimer le coût en termes de probabilité de décès comme dans la législation islandaise (Arnalds et al., 2004).

2.2.3.3. Quantification des dommages

Pour évaluer le coût lié à un événement avalancheux, une relation réaliste entre intensité du phénomène et intensité des dommages doit aussi être proposée. Ce travail de modélisation, déjà difficile dans une discipline comme l'hydrologie où le comportement du fluide est presque parfaitement connu (Merz et al., 2004), est une gageure dans le domaine des avalanches du fait des propriétés rhéologiques particulières de la neige et de la méconnaissance y afférent. Il a en effet déjà été mentionné que la pression d'impact restait difficile à quantifier à partir de la vitesse de l'écoulement (Sovilla et al., 2007). D'un point de vue plus « génie civil », le lien entre pression d'impact et dommages infligé à un matériau de construction donné (bois, maçonnerie, béton, ... etc.) reste tout aussi mal connu (Keylock et Barbolini, 2001 ; Berthet Rambaud, 2004 ; Bertrand 2007). Au final, une quantification des dommages en fonction de l'intensité de l'avalanche et correspondant aux mécanismes physiques élémentaires reste donc pour l'instant inaccessible.

Les approches existantes font de ce fait appel à des formulations empiriques de la vulnérabilité définie comme un coût normalisé dont la valeur est 0 en absence de dommage et 1 pour une destruction totale. Elles ont été obtenues par ajustement de courbes paramétriques simples sur quelques événements bien documentés en Islande, Italie et Autriche. Pour ce faire, les événements ont été rétro-analysés et le niveau de dommage observé a été associé à l'intensité calculée de l'avalanche (Jonasson et al., 1999 ; Barbolini et al., 2004b). L'unité proposée est généralement la probabilité de décès en fonction de la vitesse de l'écoulement mais quelques autres formulations existent, notamment en probabilité de destruction en fonction de la vitesse de l'écoulement pour un bâtiment ou en probabilité de décès en fonction du temps d'enfouissement pour un homme isolé.

2.3. Modélisation statistique et inférence bayésienne

2.3.1. La modélisation statistique

2.3.1.1. Modèle et vérité

La statistique est souvent vue comme la branche des mathématiques permettant d'obtenir des résumés pertinents de grandes quantités de données. Cette définition restrictive renvoie en fait à la statistique descriptive et/ou au data mining. La modélisation statistique a une ambition plus large, celle de fournir un support au raisonnement sous incertitude (O'Hagan et Oakley, 2004; Parent et Bernier, 2007). Les connaissances concernant les inconnues du phénomène étudié sont exprimées sous la forme de distributions de probabilité. Le travail du statisticien est alors de proposer une classe de lois de probabilité indexée par un vecteur de paramètres en faisant l'hypothèse que le "vrai" modèle quantifiant le phénomène étudié est inclus dans cette classe. Puis il lui faut identifier ce modèle grâce à une démarche d'inférence, c'est-à-dire d'apprentissage sur la population sous-jacente à partir de l'échantillon disponible. En statistique paramétrique, on se contente d'un vecteur d'inconnues de

dimension finie, tandis que cette limite disparaît en statistique non paramétrique moyennant des difficultés de maniement des modèles supplémentaires.

Bien entendu, l'hypothèse de l'existence d'un "vrai" modèle est en général fallacieuse et l'on est alors amené à se contenter d'un modèle acceptable sur la base d'un test d'adéquation ou d'une validation croisée ou bien à sélectionner un modèle parmi une classe plus large sur la base d'un ou plusieurs critères. Quelques cas de figures où le modèle est supposé reproduire fidèlement la réalité méritent cependant d'être signalés : celui bien connu de la loi normale vers laquelle converge un ensemble de phénomènes d'ordres de grandeur comparables grâce au théorème central limite, celui de la loi physique inattaquable du type loi de Newton, et celui des lois des valeurs extrêmes.

2.3.1.2. La théorie des valeurs extrêmes en quelques mots

Pour l'étude des événements rares, la théorie des valeurs extrêmes peut être utilisée (Coles, 2001). Celle-ci explicite la convergence des queues de distribution vers des lois limites sous certaines hypothèses. Parmi ces dernières, certaines telle la stationnarité sont vérifiables grâce à des tests (Yue et al., 2002 ; Benestad, 2004) tandis que d'autres telles la max stabilité, c'est-à-dire l'existence de suites normalisatrices pour un échantillon de maxima de sous-échantillons du phénomène, sont toujours supposées correctes.

Pour des variables aléatoires continues et unidimensionnelles, les choses sont assez simples et deux approches sont possibles (Boreux, 2005). La première considère le maximum sur une période fixe d'observations régulièrement espacées, par exemple le maximum annuel des observations journalières. Pourvu que la taille de la période soit assez grande, les maxima peuvent être considérés comme des tirages indépendants dans la loi généralisée des valeurs extrêmes ou modèle GEV (Generalized Extreme Value). La seconde, appelée couramment méthode du renouvellement en hydrologie, modélise les observations qui excèdent un seuil fixé par un processus ponctuel de Poisson marqué par une distribution de Pareto généralisée : c'est le modèle POT (Peak Over Threshold). Ce modèle présente l'avantage, par rapport au modèle GEV, de pouvoir retenir éventuellement plus d'une observation par an et, au contraire, d'exclure une année entière si aucune observation n'a été suffisamment élevée pour se trouver en domaine asymptotique. En contrepartie, le problème du choix du seuil se pose de façon aigüe : celui-ci doit être suffisamment haut pour garantir la convergence vers le modèle limite mais suffisamment bas pour conserver une quantité de données satisfaisante. Une bonne solution empirique est de représenter la moyenne arithmétique des dépassements en fonction du seuil et de sélectionner le seuil qui correspond au début de la partie linéaire du graphique obtenu (Coles, 2001). Les modèles GEV et POT possèdent chacun 3 paramètres et un rôle crucial est joué dans les deux modèles par le paramètre dit de forme qui définit l'accroissement de la période de retour avec le quantile et caractérise ainsi le domaine d'attraction de la queue de distribution. Les deux modèles GEV et POT sont très liés puisqu'on peut passer de l'un à l'autre moyennant une reparamétrisation. Ils sont inversibles, c'est-à-dire que dès lors que les paramètres sont connus une expression analytique de n'importe quel quantile de probabilité et donc de n'importe quelle période de retour est disponible. Ils sont de ce fait l'outil privilégié de définition des aléas de référence en hydrologie (Parent et Bernier, 2003a, Parent et Bernier, 2003b, Renard et al., 2006), mais présentent également un fort intérêt pour l'analyse des séries climatiques (Naveau et al., 2005) ou en théorie de la fiabilité.

En revanche, dès lors que l'on sort du cadre monodimensionnel continu, le problème se complexifie très nettement. Pour les variables aléatoires discrètes, il n'y a pas de convergence vers un modèle limite (Anderson et al., 1997) et seules des approches non paramétriques par encadrement sont théoriquement fondées (Guillou et al., 2006). Pour des variables continues pluridimensionnelles, la convergence existe mais d'une part la correspondance univoque avec la période de retour n'existe plus et d'autre part il existe une infinité de modèles limites. Le traitement de ce type de problème passe alors généralement par la spécification des lois marginales et le calcul de coefficients de dépendance asymptotiques permettant de quantifier dans quelle mesure l'apparition d'une valeur extrême sur l'une des variables marginales est susceptible de favoriser l'apparition de valeurs extrêmes sur les autres variables marginales (Bacro, 2006). Une autre possibilité rencontrant actuellement un succès croissant en hydrologie (Favre et al., 2004) est la modélisation directe de la relation de dépendance par une copule (Sklar, 1959). Enfin, très récemment, une approche semi-paramétrique a été proposée permettant la modélisation des valeurs extrêmes des lois conditionnelles pour un problème de dimension quelconque (Heffernan et Tawn, 2004).

2.3.2. L'approche bayésienne

2.3.2.1. La formule de Bayes et son interprétation « bayésienne »

La vraisemblance d'un modèle peut être maximisée par rapport aux paramètres, ce qui conduit aux estimateurs du maximum de vraisemblance généralement recherchés en statistique classique. La théorie asymptotique de la vraisemblance assure, sous condition de régularité des fonctions en jeu, que l'estimateur converge vers la vraie valeur du paramètre et que l'écart entre estimateur et vraie valeur se distribue comme une loi normale. L'intervalle de confiance, dont la largeur dépend du nombre de données et de la variabilité des observations, quantifie la robustesse de l'estimation sous l'hypothèse de la répétabilité des expériences aléatoires.

La statistique bayésienne propose une vision alternative de l'inférence issue des travaux du révérend Thomas Bayes (1763). Celui-ci a en effet laissé son nom à une formule d'apparence aussi modeste qu'inattaquable car conséquence immédiate des axiomes fondamentaux de la théorie des probabilités (eq. 2.1). Selon l'interprétation bayésienne de cette formule, $p(\theta)$ est la distribution de probabilité *a priori* du paramètre θ , celui-ci pouvant être scalaire ou vectoriel. L'appellation « *a priori* » exprime le fait qu'elle a été établie préalablement à l'observation des données notées *data* et qu'elle reflète ainsi l'opinion « personnelle » du statisticien sur le phénomène à modéliser. $l(\text{data}|\theta)$ est la probabilité des données conditionnellement au modèle statistique utilisé sachant θ , c'est-à-dire la fonction de vraisemblance de la statistique classique. $p(\theta|\text{data})$ est la distribution de probabilité *a posteriori* du paramètre θ du modèle. Le passage de la distribution *a priori* à la distribution *a posteriori* des paramètres du modèle statistique est une mise à jour de la connaissance sur la base des observations. Le dénominateur, indépendant de θ , est uniquement une constante de normalisation assurant que $p(\theta|\text{data})$ est bien une distribution de probabilité.

$$p(\theta|\text{data}) = \frac{l(\text{data}|\theta) \times p(\theta)}{\int l(\text{data}|\theta) \times p(\theta) \times d\theta} \quad (\text{eq. 2.1})$$

L'information contenue dans la loi *a posteriori* peut être résumée de différentes manières mais, habituellement, un estimateur ponctuel associé à un intervalle de crédibilité, pendant bayésien de l'intervalle de confiance, est privilégié. Le choix de l'estimateur ponctuel se fait en toute rigueur sur la base d'une fonction de coût quantifiant la perte moyenne attendue si l'on choisit une mauvaise valeur du paramètre plutôt que sa vraie valeur.

Discuter les avantages et les inconvénients de l'inférence bayésienne est un très vieux débat. L'école dite classique, fondée historiquement par entre autres Fisher, Neyman et Pearson au début du vingtième siècle, a longtemps prévalu, comme en témoignent les très nombreuses monographies qui lui sont consacrées (Montfort, 1982 ; Saporta, 1990). Elle postule que tout résultat statistique ne doit être fondé que sur les données et interprète la probabilité comme une fréquence limite sur la base de la répétabilité des expériences (Von Mises et Geringer, 1984). Ses défenseurs ont de ce fait longtemps refusé par principe l'encodage d'un quelconque savoir *a priori*.

L'approche bayésienne développe au contraire une vision subjective de la notion de probabilité. Celle-ci est interprétée comme un pari associant à un événement donné un certain degré de crédibilité (De Finetti, 1937). Dans cette optique, il n'est pas choquant d'associer dans l'inférence des données et la connaissance extérieure aux données modélisée par la loi *a priori*, la formule de Bayes étant l'instrument qui rend possible le couplage par le calcul probabiliste. Néanmoins, les bayésiens ont dû combattre les objections des statisticiens classiques par la recherche de *priors* non informatifs utiles lorsque la connaissance *a priori* n'est pas disponible ou lorsque l'on ne désire pas l'utiliser pour laisser les données s'exprimer par elles mêmes (Bernardo et Smith, 1994). Il a été montré que de tels *priors* conduisent asymptotiquement aux mêmes estimateurs que ceux du maximum de vraisemblance (Berger, 1985 ; Robert, 1992), de sorte que l'inférence bayésienne a été peu à peu acceptée. Cette pacification du débat a été accélérée par le développement des algorithmes d'estimation bayésiens à partir des années 1970-80 qui a conduit les praticiens de disciplines très différentes à s'emparer avec succès de l'inférence bayésienne pour des raisons de commodité davantage que pour une adhésion à ses concepts. La statistique bayésienne a alors pris un essor considérable comme en témoigne la récente multiplication des monographies qui lui sont consacrées (Box et Tiao, 1973 ; Berger, 1985 ;

Bernardo et Smith, 1994 ; Gelman et al., 1995 ; Lee, 1997 ; Bernier et al., 2000 ; Gilks et al., 2001) et auxquelles on se référera pour une présentation théorique plus approfondie.

2.3.2.2. Le cas des modèles hiérarchiques

Les modèles dits hiérarchiques ou « à couches » sont caractérisés par des variables non observables dites latentes. Celles-ci permettent de représenter une grande diversité de processus et offrent de ce fait une très grande liberté au modélisateur pour construire un modèle collant au plus près à l'idée qu'il se fait du phénomène qu'il étudie (Parent et Bernier, 2007).

Les modèles hiérarchiques ne sont pourtant très utilisés pour des problèmes concrets que depuis une quinzaine d'années du fait des difficultés d'estimation qu'ils peuvent poser. Ces difficultés sont nettement réduites sous le paradigme bayésien puisque l'introduction de variables latentes ne fait que rajouter les inconnues $vlat$ et le terme $p(vlat|data, \theta)$ quantifiant leur distribution sachant les paramètres et les observations au théorème de Bayes (eq. 2.2). L'inférence s'effectue de ce fait grâce aux mêmes algorithmes de simulation que pour les modèles non hiérarchiques.

$$p(\theta, vlat|data) \propto l(data|\theta, vlat) \times p(vlat|data, \theta) \times p(\theta) \quad (\text{eq. 2.2})$$

La modélisation bayésienne sur variable latente a trouvé des terrains d'expression naturels en écologie (Rivot, 2003 ; Clark, 2005), en hydrologie (Perreault, 2000) ou en géophysique (Parent et Bernier, 2001 ; Berliner, 2003) en permettant la séparation d'effets spatio-temporels, l'identification de différents facteurs, de tendances ou de ruptures. Elle est particulièrement efficace dans le cadre de la statistique spatiale en permettant l'inférence de données spatialement corrélées non gaussiennes pour lesquelles l'indexage spatial est utilisé comme une variable explicative de l'observation (Cressie, 1993 ; Banerjee et al., 2003).

2.3.2.3. Calcul bayésien et méthodes MCMC

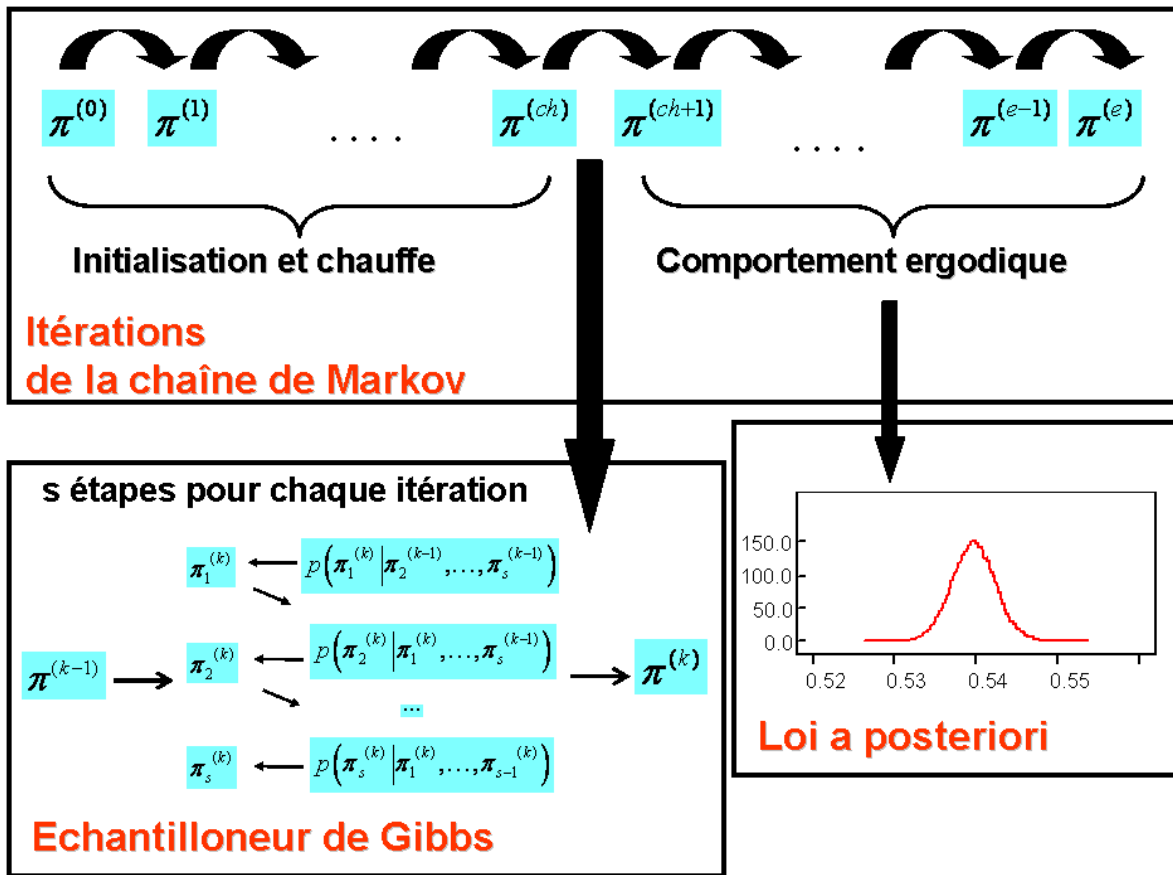
La difficulté technique de l'inférence bayésienne est liée au calcul de l'intégrale au dénominateur de la formule de Bayes. La constante de normalisation n'est en effet généralement pas calculable analytiquement, tout particulièrement lorsque le vecteur des inconnues est de grande dimension. L'analyse bayésienne a de ce fait été longtemps limitée à des modèles simples faisant appel à des lois conjuguées pour lesquels le *posterior* et le *prior* appartiennent à une même classe de lois de probabilité ou pour lesquels des approximations normales peuvent être utilisées.

Le renouveau de l'inférence bayésienne est passé par le développement de méthodes de simulation permettant d'inférer en principe n'importe quel modèle, même de dimension très élevée. Celles-ci sont de deux types : non itératives ou itératives. Les algorithmes non itératifs les plus utiles font appel à l'échantillonnage par importance. Ils sont très efficaces lorsque le *posterior* est assez bien connu avant observation et sont de ce fait bien adaptés à la mise en œuvre de l'apprentissage séquentiel. Ils sont par contre plus difficiles à manier lorsque peu d'arguments permettent de guider le choix de la fonction d'importance. Les algorithmes itératifs, plus connus sous le nom de méthodes de Monte Carlo par Chaînes de Markov (MCMC), utilisent quant à eux une chaîne de Markov construite dans l'espace des paramètres et convergeant sous des hypothèses peu restrictives vers la loi cible recherchée.

Une présentation générale des méthodes MCMC figure dans Robert (1996) tandis que de nombreux articles (Brooks, 1998 ; Brooks et Roberts, 1998) détaillent leurs propriétés de convergence. Le principe général est très simple : un état initial $\pi^{(0)} = (\theta^{(0)}, v_{lat}^{(0)})$ est choisi pour le vecteur des inconnues du modèle, c'est à dire l'ensemble des paramètres et variables latentes. Ensuite l'algorithme MCMC génère une chaîne de Markov, c'est-à-dire une série $(\pi^{(1)}, \pi^{(2)}, \dots, \pi^{(k)}, \dots)$ telle que $\pi^{(k)} | \pi^{(k-1)}$ est indépendante du passé de la chaîne $(\pi^{(1)}, \pi^{(2)}, \dots, \pi^{(k-2)})$, et assure que la loi stationnaire de la chaîne est la loi cible $p(\pi|data)$. Cela implique que pour une période dite de chauffe suffisamment longue de ch itérations, la série $(\pi^{(ch+1)}, \pi^{(ch+2)}, \dots)$ constitue un échantillon du *posterior* recherché. Cet échantillon peut ensuite être utilisé pour évaluer empiriquement les statistiques d'intérêt telles les moyennes, variances et corrélations inter paramètres *a posteriori* (Figure 2.15).

Les deux principales méthodes MCMC sont l'algorithme de Metropolis Hastings (Metropolis et al, 1953 ; Hastings, 1970) et l'échantillonneur de Gibbs (Geman et Geman, 1984). Ils sont désormais bien connus (Tanner, 1992 ; Robert, 1996) et leur utilisation est décrite dans de nombreux domaines d'application (Kuczera et Parent, 1998; Torre et al., 2001), de sorte qu'ils ne seront pas présentés en détail. Le principe de l'échantillonneur de Gibbs est de décomposer chaque itération en tirages dans les différentes lois conditionnelles complètes de chacune des inconnues sachant les autres (Figure 2.15). Pour l'algorithme de Metropolis Hastings, les étapes successives à chaque itération sont la génération d'un candidat à l'aide d'une fonction d'exploration et l'application d'une règle d'acceptation/rejet probabiliste. Signalons enfin que l'algorithme de Gibbs est en fait un cas particulier de celui de Metropolis Hastings, et que ces deux algorithmes peuvent être imbriqués (étape Metropolis à l'intérieur d'un algorithme de Gibbs par exemple) et implémentés de manière très variée, notamment séquentiellement ou par blocs.

Figure 2.15 : Principe des méthodes MCMC et mise en œuvre grâce à l'algorithme de Gibbs



s est la dimension du vecteur des inconnues.

Les e - ch valeurs de la chaîne obtenues après la période de chauffe permettent de reconstituer la loi a posteriori

Le maniement des algorithmes MCMC peut s'avérer très délicat en pratique car il n'existe pas de mesure préalable de la vitesse de convergence, de sorte que cette dernière peut être atteinte au bout d'une période de chauffe très, voire infiniment, longue (Brooks et Gelman, 1998) si l'algorithme est mal construit (choix d'une fonction d'exploration inadaptée, valeurs initiales trop « exotiques »...,etc). Il est donc nécessaire de vérifier à chaque utilisation que la convergence vers la loi cible a été atteinte (Mengersen et al., 1999). Pour ce faire, la seule solution pratique est souvent de lancer simultanément plusieurs chaînes et d'effectuer des tests permettant de vérifier qu'elles se comportent de manière identique, c'est-à-dire que la loi cible est atteinte pour les différentes chaînes. La stratégie la plus courante est alors la comparaison de la variance inter chaîne avec la variance intra chaîne (Gelman, 1996).

L'écriture d'algorithmes MCMC qui convergent à des vitesses raisonnables nécessite donc, de la part du statisticien, une certaine habileté pratique en plus de connaissances théoriques solides. Il doit en effet choisir une exploration de l'espace des paramètres suffisamment « large » pour ne pas rester

bloqué dans un mode local du *posterior* mais également suffisamment fine pour ne pas balayer l'espace des paramètres sans jamais détecter les zones de plus forte probabilité, et ce alors que la surface cible peut posséder, pour des modèles de dimension élevée, une forme complexe rendant difficile son exploration conformément aux probabilités *a posteriori* respective des différentes régions. L'algorithme de Gibbs est à recommander lorsque les lois conditionnelles complètes sont disponibles car le choix de la fonction d'exploration est guidé par l'algorithme. Par contre, lorsque le recours à l'algorithme de Metropolis Hastings est nécessaire, des règles pratiques de réglage de la variance d'exploration en fonction des taux d'acceptation doivent être utilisées. En outre, il est bien souvent nécessaire d'opérer une itération séquentielle sur les différentes inconnues de façon à ne pas se déplacer trop fortement dans l'espace des paramètres à chaque saut.

Signalons enfin que, pour de nombreux modèles, les difficultés liées au maniement des algorithmes MCMC peuvent être évitées grâce au logiciel Winbugs (Spiegelhalter et al., 2000) qui génère automatiquement l'algorithme d'estimation à partir de la déclaration du modèle en reconnaissant lois conjuguées et conditionnelles complètes puis en utilisant des boucles imbriquées de Metropolis Hastings pour les inconnues résiduelles. Son fonctionnement est particulièrement efficace pour les modèles hiérarchiques spatiaux sur grille irrégulière (Thomas et al., 2004) mais il peut également être utilisé pour bien d'autres types de modèles (Congdon, 2001).

2.3.2.4. Prédiction et sélection de modèle

Le formalisme bayésien permet la prédiction en prenant en compte à la fois la variabilité naturelle du phénomène étudiée et l'incertitude liée à la quantité limitée de données sur laquelle a été effectuée l'inférence. La loi prédictive $p(z^{pred} | data)$, c'est-à-dire la probabilité d'observer z^{pred} dans le futur sachant les observations, s'obtient en moyennant le modèle de vraisemblance sur la loi *a posteriori* des paramètres (eq. 2.3). La moyenne prédictive est alors le meilleur pari compte tenu de l'état de connaissance tandis que l'intervalle de crédibilité associé offre des bornes raisonnables pour la prédiction et le contrôle (Girard et Parent, 2004).

La loi prédictive s'obtient analytiquement encore plus rarement que la loi *a posteriori*. Par contre, son obtention par simulation à partir d'un échantillon de la loi *a posteriori* est très simple en échantillonnant le modèle conditionnellement à chaque valeur de la loi *a posteriori*. Cette absence de forme analytique fixe peut d'ailleurs être vue comme un avantage dans la mesure où elle permet une bonne modélisation de la variabilité du prédicteur en mélangeant autant de distributions que nécessaire.

$$p(z^{pred} | data) = \int l(z^{pred} | \theta) \times p(\theta | data) \times d\theta \quad (\text{eq. 2.3})$$

Bien que quantifiant de façon honnête l'erreur de prédiction liée à l'emploi du modèle statistique, la loi prédictive ne prend pas en compte l'erreur de modèle et ne dispense de ce fait pas de tester la validité du choix de modèle effectué. Au delà des méthodes traditionnelles du type séparation des données en échantillons de calage/validation et/ou de validation croisée, le formalisme bayésien propose différents outils de comparaison de modèles permettant la sélection d'un modèle parmi un nombre fini de modèles candidats.

Le plus classique d'entre eux est le facteur de Bayes (Carlin et Chib, 1995), qui est le ratio pour les deux modèles M_1 et M_2 concurrents des vraisemblances moyennées sur les lois *a posteriori* (eq. 2.4). Le calcul de cet analogue bayésien du rapport de vraisemblance pose des problèmes techniques liés notamment à la nécessité d'employer un *prior* propre. Kaas et Raftery (1994) ont cependant montré qu'une approximation relativement robuste pouvait en être obtenue en calculant terme à terme numérateur et dénominateur à partir de la moyenne harmonique par rapport à la loi *a posteriori* de la vraisemblance (eq. 2.5), les calculs pouvant évidemment être effectués numériquement à partir des échantillons obtenus par MCMC. Kaas et Raftery (1994) proposent également des règles pratiques pour l'interprétation du facteur de Bayes indiquant les seuils correspondant à une forte évidence en faveur de l'un des deux modèles.

$$B_{12} = \frac{p(data | M_1)}{p(data | M_2)} \quad (\text{eq. 2.4})$$

$$p(\text{data}|M_1) = \left[\int p(\text{data}|M_1, \theta)^{-1} \times p(\theta|M_1, \text{data}) \times d\theta \right]^{-1} \quad (\text{eq. 2.5})$$

Une autre option pour la comparaison de modèles est le Critère d'Information de la Déviance (DIC) qui propose une pondération de la log vraisemblance moyenne $D_{\text{moy}} = -E_{\theta} \left[\log(p(\text{data}|\theta)) \right]$ par le nombre effectif de paramètres du modèle, noté usuellement pD , afin de pénaliser les modèles surparamétrés (Spiegelhalter et al., 2002). pD est estimé par la différence $D_{\text{moy}} - D_{\text{chap}}$ entre la log vraisemblance moyenne et la log vraisemblance des données au niveau de la moyenne *a posteriori* $D_{\text{chap}} = -\log(\text{data}|E_{\theta}[\theta|\text{data}])$. Le DIC (eq. 2.6), analogue bayésien du critère d'information d'Akaike (AIC), est donc d'autant plus petit que le modèle supporte bien les données.

$$DIC = D_{\text{moy}} + pD = D_{\text{chap}} + 2pD \quad (\text{eq. 2.6})$$

2.3.2.5. Risque et décision

La fonction de coût $C(d, z)$ quantifie le coût de chaque décision d pour chaque réalisation z du phénomène étudié. Ce coût peut éventuellement être pondéré par le comportement du décideur face au risque, on parle alors de fonction d'utilité. Outil de base pour une analyse coût-bénéfice, la fonction de coût est également le support de toute approche d'optimisation visant à minimiser la perte moyenne attendue. Au rang des applications les plus courantes figurent par exemple l'optimisation de plans d'expérience (Atkinson et Donev, 1992), de traitement médicaux (Wakefield, 1994) ou d'ingénierie hydraulique (Bernier, 2003). Dans le domaine des risques naturels, une approche par optimisation économique présente l'avantage de pouvoir éviter le choix d'un aléa de référence lorsque celui-ci est particulièrement difficile à définir, comme par exemple dans le cas d'un phénomène multivarié.

Le risque au sens classique est l'espérance du coût (ou de l'utilité) par rapport au phénomène étudié. C'est donc une fonction de la variable de décision et des paramètres du modèle décrivant le phénomène. Au sens de la théorie de la décision classique, la décision optimale minimise ce risque, mais elle est alors une fonction des paramètres θ décrivant le phénomène. La pratique courante est d'obtenir des estimateurs ponctuels, par exemple par maximum de vraisemblance, puis de recommander par « plug-in » la décision minimisant le risque conditionnellement à ces estimateurs.

Dissocier l'inférence du problème décisionnel peut cependant avoir des conséquences indésirables. En effet, le phénomène est ainsi traité pour la décision comme s'il était parfaitement connu, ce qui est très irréaliste lorsque peu de données sont disponibles. De plus, le cadre classique ne prend pas en compte la différence entre les points de vue du statisticien et du décideur. En général, le statisticien classique effectue en effet l'inférence en minimisant un critère de variance, c'est-à-dire une fonction quadratique symétrique. Le décideur, au contraire, pourra privilégier dans certains cas une pénalisation fortement asymétrique, la submersion d'une digue trop petite ayant par exemple un coût bien plus élevé que la construction d'une digue trop élevée.

Le risque bayésien prend en compte l'incertitude liée à la méconnaissance locale pour la décision en moyennant le risque classique sur toute la loi *a posteriori*. L'erreur d'estimation est ainsi incorporée de façon claire dans la décision. Le risque bayésien peut de ce fait être vu comme une fonction de la variable décisionnelle uniquement ce qui rend la détermination de la décision bayésienne optimale facile. Il convient cependant de garder à l'esprit que le risque bayésien reste également une fonction de l'information apportée par le *prior* et les données. D'importants développements méthodologiques sous la forme d'algorithmes de simulation découlant naturellement des méthodes d'estimation MCMC (Müller, 1999) et IS (Amzal, 2004 ; Amzal et al., 2006) permettent maintenant la mise en œuvre de l'optimisation bayésienne de problèmes multidimensionnels complexes.

Un cas particulier important de la prise de décision bayésienne sous incertitude est celui de l'estimation bayésienne ponctuelle quand un estimateur ponctuel d doit être choisi une fois l'inférence faite. L'espace des décisions se ramène alors au support de la loi *a posteriori* $p(\theta|\text{data})$ et la fonction de coût traduit la pénalité liée à l'erreur d'estimation. En particulier, avec une fonction de coût quadratique $C(\theta, d) = (\theta - d)^2$ souvent proposée à défaut de mieux, l'estimateur bayésien est simplement la moyenne *a posteriori* du paramètre tandis qu'il s'agit de la médiane *a posteriori* avec

un coût linéaire. Cela explique que ces estimateurs sont généralement privilégiés dès lors que la modélisation des coûts n'est pas effectuée. Mais rien n'interdit de construire un cadre décisionnel plus réaliste mais plus complexe conduisant à d'autres estimateurs (Krzysztofowicz, 1983; Berger, 1985).

2.3.2.6. *Peu d'applications au domaine des avalanches*

Pour l'étude des avalanches, la modélisation statistique est peu utilisée. D'abord parce qu'elle reste peu diffusée dans le milieu des avalanches, en particulier auprès des praticiens. Ensuite parce que les données disponibles sont assez peu nombreuses, tout particulièrement à l'échelle d'un massif entier, l'EPA constituant à ce titre une exception remarquable dont la valorisation reste, malgré les travaux déjà évoqués, insuffisante. L'ensemble des approches existantes peut donc être synthétisé rapidement, d'autant qu'elles ont pour la plupart d'entre elles déjà été évoquées :

- L'utilisation la plus routinière de la modélisation statistique dans le domaine des avalanches est la caractérisation fréquentielle des chutes de neige rares au niveau des zones de départ (Parent et Bernier, 2003b). Ce type d'approche inspiré de l'hydrologie stochastique « traditionnelle » (Lang, 2000) vise à quantifier le volume de neige disponible au moment du déclenchement des avalanches les plus importantes.
- En prévision, l'emploi de modèles probabilistes de déclenchement est également assez courant (Gassner et Brabec, 2002; Ghinoi et Chang, 2005), notamment en Amérique du nord où, comme signalé précédemment, des modèles de type plus proches voisins sont souvent utilisés pour définir le risque d'occurrence des avalanches à partir d'analogues nivo météorologiques (McCollister et al., 2003).
- En prédétermination, des modèles statistiques simples de type binomiaux sont parfois utilisés à l'échelle du couloir pour estimer une probabilité de déclenchement moyenne sur une longue période à partir de chroniques historiques (Harbitz et al., 2001). De même, les calculs de période de retour reposent sur des régressions ou des ajustements dans les « méthodes norvégiennes » et sur une estimation de quantile par simulation dans les « méthodes statistiques dynamiques » déjà évoquées.
- D'un point de vue moins pratique, quelques approches ont été tentées à plus large échelle pour étudier les corrélations entre occurrences avalancheuses, topographie (Maggioni et Gruber, 2003 ; Maggioni, 2005) et forçage météorologique (Marchal et al., 2003 ; Etienne, 2006).
- Enfin, l'application des méthodes bayésiennes au domaine des avalanches n'est pas neuve puisque quelques approches assez simples utilisant des lois conjuguées ont été proposées dès le milieu des années 90. Ainsi McClung et Tweedy (1994) utilisent un modèle bayésien d'analyse discriminante. De même, Harbitz et al. (2001) proposent d'estimer les probabilités de déclenchement moyennes à l'aide d'un modèle conjugué beta-binomial. Plus récemment, une approche faisant appel aux réseaux bayésiens a été utilisée pour l'estimation à l'échelle régionale des emprises d'avalanches (Straub et Grêt-Regamey, 2006) et d'un coût global (Grêt-Regamey et Straub, 2006). Enfin Ancey (2005) a été le premier à utiliser les méthodes MCMC en dynamique des avalanches pour l'estimation du coefficient de frottement d'un modèle de propagation coulombien à l'aide d'un algorithme de Metropolis Hastings. Le modèle qu'il propose relie le coefficient de frottement coulombien μ_{coul} au volume de l'avalanche v_{stop} au travers d'un modèle de mélange gaussien de moyenne $c \times \exp(-d \times v_{stop})$ traduisant la diminution moyenne du frottement avec le volume de l'avalanche et de variance σ^2 constante (eq. 2.7). Les trois paramètres (c, d, σ) en sont estimés sur un large échantillon d'avalanches.

$$\left(\mu_{coul} \mid v_{stop}, c, d, \sigma \right) \sim N \left(c \times \exp(-d \times v_{stop}), \sigma^2 \right) \text{ (eq. 2.7)}$$

2.3.2.7. *Mais de nombreuses applications dans des domaines proches*

Au contraire, pour d'autres risques naturels et en sciences de l'ingénieur, la modélisation stochastique revêt une importance croissante. Elle permet en effet la quantification des incertitudes liées à l'utilisation de codes de calculs sophistiqués (Oakley et O'Hagan, 2004a ; Oakley et O'Hagan, 2004b; Karniadakis et Glimm 2006), l'obtention de résultats opérationnels utiles à l'évaluation de l'aléa (Jibson et al., 2000 ; Krzysztofowicz, 2001a ; Krzysztofowicz, 2001b ; Naveau et al., 2005) et une meilleure compréhension du phénomène dont la variabilité est modélisée. Pour ce faire, la théorie des

valeurs extrêmes et/ou la modélisation spatiale hiérarchique sont particulièrement bien adaptées comme en témoigne le nombre de travaux qui leur sont consacrés et dont beaucoup d'exemples ont précédemment été cités.

2.4. Conclusion du chapitre 2

Les avalanches ont été largement étudiées au cours des dernières décennies grâce à des approches autant numériques que naturalistes, mais les processus élémentaires gouvernant le déclenchement et la propagation restent encore mal connus. De ce fait, aucune des méthodes de prédétermination existantes ne néglige le recours à l'information historique. C'est explicite pour les méthodes expertes traditionnelles qui combinent enquête de terrain et recueil de témoignages ainsi que pour les méthodes dites norvégiennes. C'est plus implicite, mais néanmoins indéniable, pour les approches numériques faisant appel à des valeurs tabulées de coefficients de frottement du fait du caractère semi-empirique des lois de frottement proposées.

Pourtant, le lien entre information historique et modélisation numérique reste peu développé, notamment parce que la donnée est rare et que les méthodes probabilistes sont peu diffusées dans le milieu des avalanches. Plus généralement, les applications de la modélisation statistique, en particulier bayésienne, à la valorisation des bases de données avalancheuses restent peu nombreuses, et ce alors qu'elles connaissent un franc succès dans des domaines proches.

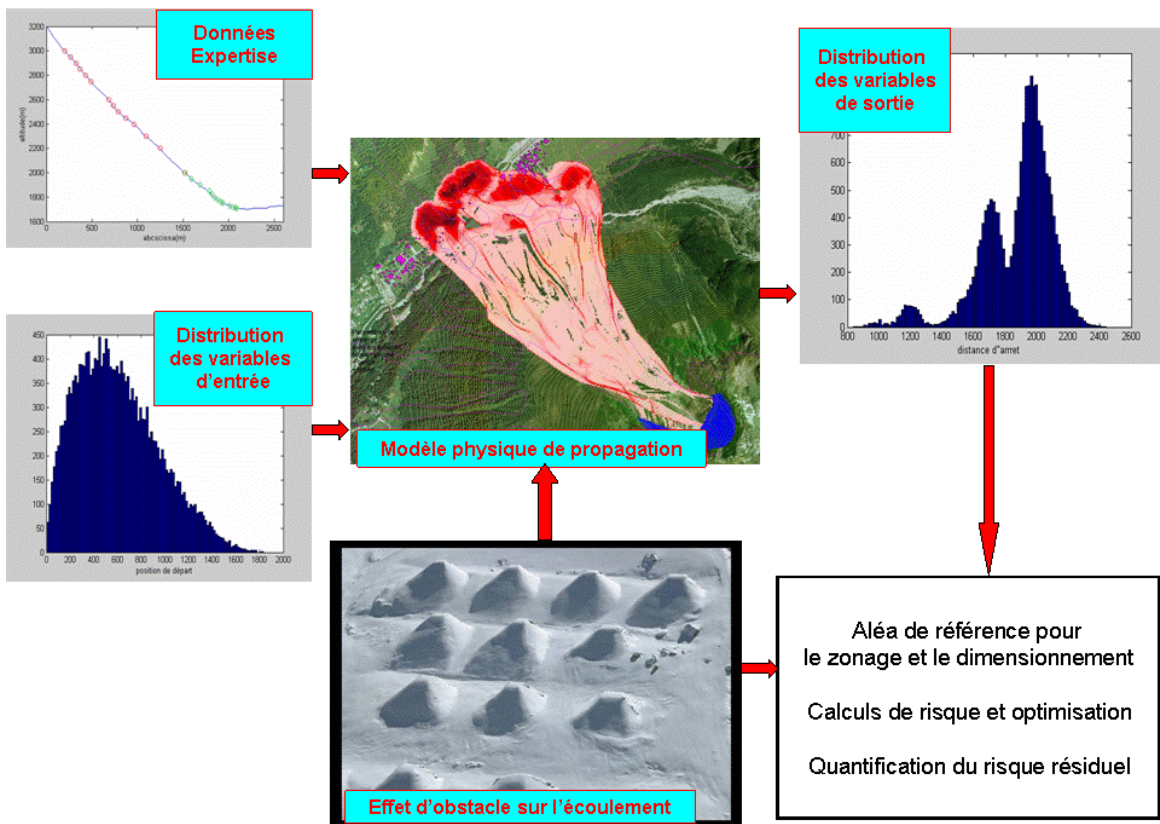
Les outils de modélisation bayésienne choisis dans le cadre de cette thèse peuvent donc vraisemblablement faciliter l'extraction de l'information contenue dans les précieuses données avalancheuses disponibles en France et permettre ainsi l'assemblage des connaissances. En priorité, il s'agit de s'inspirer des approches existantes pour d'autres risques naturels et en sciences de l'ingénieur, à la fois pour obtenir des résultats opérationnels et mieux décrire la variabilité des phénomènes étudiés.

3. Assembler les connaissances

Préambule

L'objectif de ce chapitre est d'entrer dans le vif du sujet du travail effectué au cours de la thèse. Il expose la construction d'un cadre probabiliste générique intégrant les connaissances disponibles présentées au chapitre 2. Il s'agit de montrer comment les outils bayésiens peuvent permettre d'améliorer la coordination entre information historique recueillie dans l'EPA et modèles numériques afin d'effectuer des prédictions pour le zonage et le dimensionnement par simulation (Figure 3.1). Trois phases sont distinguées : la modélisation stochastique, l'inférence et la prédiction. L'échelle d'étude est tout d'abord celle d'un site avalancheux de l'Enquête Permanente sur les Avalanches (EPA). Ensuite, les bases permettant l'extension spatio-temporelle et décisionnelle du cadre proposé sont présentées.

Figure 3.1 : Cadre intégrateur général pour la prédétermination des avalanches.



*Données EPA, distribution des abscisses de départ et d'arrêt sur le site EPA 13 de Bessans, Savoie, France.
Modélisation numérique bidimensionnelle de l'avalanche de Montroc, Haute-Savoie, France, d'après M. Naaim. Tas freineurs de Taconnaz, Haute-Savoie, France, photo F. Valla/ Cemagref.*

3.1. Revisiter et approfondir les approches statistiques-dynamiques

Au-delà du problème de terminologie entre modèle conceptuel et déterministe (De Marsily, 1994) brièvement discuté au chapitre précédent et sur lequel on reviendra au chapitre 8, l'approche retenue dans le cadre de cette thèse s'apparente aux méthodes statistiques-dynamiques. En ce sens, elle prolonge directement la démarche de Meunier et al. (2004) présentée au chapitre 2. Cependant, on cherche à remédier à certaines insuffisances précédemment soulignées (choix de modèle statistiques pour les données d'entrées, mélange de données issues de différents sites, etc.) grâce d'une part à un effort de modélisation accru et d'autre part au recours à des algorithmes de simulation permettant d'effectuer l'inférence sur les données EPA. Pour représenter et estimer des quantités non observées (vraies positions d'arrêt des avalanches par exemple) ou non observables (coefficients de frottement),

la modélisation sur variables latentes est systématiquement utilisée. Plus précisément, les lignes directrices de l'approche proposées sont les suivantes :

- *Distinguer clairement modélisation, inférence et prédiction*

Dans cette thèse, la prédétermination est clairement scindée en trois étapes. En premier lieu, s'effectue le travail de modélisation qui consiste à préciser le système étudié et à traduire les relations inter variables en utilisant au mieux la connaissance disponible. Il s'agit donc d'une étape d'assemblage au moins partiellement subjective. L'inférence constitue ensuite l'étape la plus technique permettant la mise à jour de la connaissance à l'aide des données. Enfin le modèle obtenu peut être utilisé en prédiction. Cette dernière étape est calculatoire mais peu technique. Conditionnellement aux choix de modélisation effectués et à l'information utilisée pour l'inférence, des simulations sont effectuées de façon à obtenir les résultats opérationnels désirés : aléas de référence, quantification du risque résiduel après construction d'un éventuel ouvrage de protection, voire optimisation de l'ouvrage lui-même. Ces trois étapes de la prédétermination sont introduites de façon générique dans ce chapitre puis développées plus concrètement au cours des chapitres suivants.

- *Ne pas se contenter de distributions standard*

Dans les méthodes statistiques-dynamiques, les coefficients de frottement ainsi que les autres entrées du modèle de propagation sont modélisés par des distributions standard telles des lois normales ou de Gumbel. Ces choix de commodité mathématique ne permettent pas réellement de s'adapter à la variabilité des données, en particulier pour les variables latentes. Dans cette thèse, on privilégie donc des lois de mélange qui, comme dit au chapitre 2, présentent l'avantage d'être très souples et de s'exprimer simplement dans le cadre des modèles hiérarchiques.

- *Prendre en compte la corrélation inter variables*

Dans les méthodes statistiques de prédétermination, les variables d'entrée sont généralement supposées indépendantes. Cette hypothèse est possible pour une analyse de sensibilité visant à évaluer l'influence d'une variable d'entrée sur la variabilité des sorties. Elle est par contre fallacieuse si des quantiles élevés des variables de sortie sont estimés car les événements extrêmes catastrophiques sont souvent dus à une combinaison de valeurs extrêmes de différents facteurs pouvant interagir. Dans cette thèse, on propose donc des structures de modèles autorisant systématiquement l'existence de corrélations inter variables quitte à établir *a posteriori* qu'elles sont inexistantes.

- *Evaluer aléas de référence et périodes de retour à partir des lois de probabilité marginales*

Comme on l'a vu au chapitre 2, les aléas de référence dans le domaine des avalanches sont définis en termes probabilisés de manière parfois un peu « fantaisiste ». Ainsi la méthode suisse considère que la période de retour se conserve à travers le modèle de propagation et ne prend pas en compte la probabilité de déclenchement. Dans cette thèse, on s'en tient au contraire à une définition stricte de la période de retour que l'on définit comme le temps moyen en années séparant deux avalanches pour lesquelles la variable aléatoire d'intérêt dépasse le quantile considéré. Les périodes de retour ne sont ainsi calculées qu'à partir des distributions de probabilité annuelle marginales en combinant probabilité de déclenchement et dynamique de l'écoulement.

- *Prédire en tenant compte de l'erreur d'estimation*

En hydrologie, le calcul d'intervalles de confiance sur les valeurs de projet est courant. Cela n'est par contre quasiment jamais effectué dans le domaine des avalanches. Les sites EPA n'ont pourtant enregistré en général que quelques dizaines d'événements depuis qu'ils sont suivis, et très rarement plus de cinquante. Les intervalles de confiance après estimation sont donc nécessairement assez larges, de sorte que la définition des aléas de référence ne peut se faire sans tenir compte de l'erreur d'estimation. Dans cette thèse, des intervalles de confiance et/ou de crédibilité sur les valeurs de projet proposées sont donc systématiquement calculés.

- *Privilégier en premier lieu l'information locale*

Dans cette thèse, on développe une approche n'utilisant que l'information historique disponible sur le site étudié pour effectuer l'inférence. Des modèles parcimonieux ne nécessitant en entrée que les

données enregistrées dans l'EPA sont donc privilégiés par rapport à des modèles plus réalistes du point de vue de la physique mais nécessitant des données qui ne sont jamais disponibles en pratique pour l'ingénierie. Comme dans les méthodes norvégiennes présentées au chapitre 2, les calculs des aléas de référence sont donc basés avant tout sur des données de distances d'arrêt. Mais contrairement aux méthodes norvégiennes, l'hypothèse d'homogénéité régionale n'est pas effectuée.

- *Savoir remettre en cause les hypothèses*

Le travail de modélisation effectué dans le cadre de cette thèse est caractérisé par un nombre important d'hypothèses. Le recours à des lois de mélange, même s'il offre une certaine souplesse, reste en particulier un engagement subjectif fort. Une fois qu'une structure de base et des premiers résultats sont obtenus, les hypothèses doivent donc être rediscutées afin de vérifier la robustesse du modèle proposé et de tenter de l'améliorer. Les directions explorées seront présentées de façon générique en section 3.5 puis seront développées plus concrètement sur des cas d'étude avec les différents modèles proposés.

- *Etre conscient des limites de la période de retour et aller au-delà*

Pour des aléas de référence multivariés, la notion de période de retour trouve vite ses limites. Il convient donc d'évaluer dans quelle mesure cet outil de zonage et de dimensionnement peut s'appliquer avec profit au domaine des avalanches. Ce sera l'objet d'une discussion au chapitre 8. Il convient également de développer des outils alternatifs pour les situations où elle est mal adaptée. Le choix effectué dans cette thèse est de compléter la modélisation du phénomène par une modélisation économique permettant de décider en minimisant le risque. Le potentiel d'un cadre décisionnel est illustré en traitant en détail le problème de l'optimisation d'une digue paravalanche. Le principe en est présenté à la fin de ce chapitre et l'ensemble des calculs est détaillé au chapitre 7.

3.2. Modélisation du fonctionnement d'un site avalancheux

3.2.1. Construction d'un cadre probabiliste rigoureux

3.2.1.1. Terminologie et notations

Pour la prédétermination locale, le système étudié est un site avalancheux dont la topographie est supposée connue de façon exacte. Le choix de modélisation effectué est de privilégier un cadre mimant grossièrement la succession des phénomènes depuis la chute de neige jusqu'à l'arrivée de l'avalanche dans sa zone d'arrêt. La variabilité à l'échelle du site des observables caractérisant le forçage nivo-météorologique (cumuls de neige, température, vent, etc.), le déclenchement (surface de la zone de départ, volume de neige mobilisé, etc.) et la propagation (extension, vitesse, pression, etc.) est décrite par des lois de probabilité indexées par des paramètres. Les données de l'EPA sont considérées comme des réalisations effectives de ces distributions et peuvent par conséquent servir de support à l'inférence.

Chacune des avalanches observées sur le site étudié constitue un *événement* et l'ensemble des avalanches susceptibles de se produire constitue un *univers* au sens probabiliste du terme. Nous appellerons caractéristiques l'ensemble des grandeurs numériques ou catégorielles associées à un événement : date, altitude d'arrêt, qualité de la neige, etc. Elles sont supposées *mesurables* au sens mathématique du terme lors de l'observation d'un événement. D'un point de vue probabiliste, on peut donc considérer que l'archivage d'un jeu de valeurs lors d'un événement revient à observer la réalisation d'une *variable aléatoire pluridimensionnelle de dimension k*, chaque caractéristique correspondant à une *variable aléatoire marginale*. Par convention, la variable elle-même est notée en majuscule latine $Z = (Z_1, Z_2, \dots, Z_k)$ et sa réalisation en minuscule latine $z = (z_1, z_2, \dots, z_k)$.

On supposera que la variable aléatoire Z peut être représentée par le modèle de vraisemblance $l(z|\theta)$, $Z|\theta \sim p(\cdot|\theta)$ indiquant alors que conditionnellement à la connaissance du vecteur θ des paramètres du modèle, la variable aléatoire Z est caractérisée par la loi de probabilité $p(\cdot)$ discrète ou continue

indexée par θ . Finalement, les lois jointes, conditionnelles et marginales des variables aléatoires Z_1 et Z_2 sachant les paramètres sont notées respectivement $p(z_1, z_2 | \theta)$, $p(z_1 | z_2, \theta)$ et $p(z_1 | \theta)$.

L'objet de la démarche inférentielle est d'apprendre sur θ grâce aux données, tandis que la prédétermination a pour objectif, une fois θ connu, de préciser les aléas de référence, c'est-à-dire les valeurs des caractéristiques correspondant à des probabilités de dépassement faibles généralement fixées par le décideur.

3.2.1.2. Description des observables

On suppose que N avalanches ont fait l'objet d'une description quantitative sur le site étudié durant une période d'observation de T_{obs} années. Une année est ici une année avalancheuse, c'est-à-dire un « hiver » à cheval sur deux années civiles et correspondant à une campagne de l'EPA. On note P le nombre total d'avalanches qui se sont produites sur le site étudié durant les T_{obs} années d'observation.

P peut éventuellement être supérieur à N si les caractéristiques quantitatives de tous les événements ne sont pas enregistrées, par exemple parce qu'un seuil minimal en altitude ou en volume de dépôt a été défini par le protocole d'observation.

Chaque avalanche peut être décrite par un ensemble de variables (x_i, y_i) , $i \in [1, N]$. Les x_i sont les variables « d'entrées » caractérisant le déclenchement et la propagation : surface de départ, altitude de départ, volume de neige mobilisé, qualité de la neige..., etc. Les y_i correspondent aux variables dites « de sortie » présentant un intérêt pour le zonage et le dimensionnement : distance d'arrêt, extension, champs de vitesse et de pression..., etc. On utilise la notion d'intensité de l'avalanche pour caractériser chaque réalisation (x_i, y_i) , $i \in [1, N]$. La mise en œuvre d'une approche statistique-dynamique suppose alors la définition d'un modèle d'intensité $l(x, y | \theta_M)$. Celui-ci, indexé par l'inconnue θ_M , relie les variables x_i aux variables y_i et offre ainsi une vision systémique du fonctionnement d'un couloir avalancheux.

Du fait de la nature discrète du phénomène avalancheux, toutes les lois de probabilité pouvant être inférées à partir des données (x_i, y_i) , $i \in [1, N]$ sont conditionnelles au déclenchement. Par souci de simplicité, ce conditionnement n'apparaît pas dans la notation choisie pour le modèle d'intensité. Cependant, la variable aléatoire marginale date de l'événement est fondamentalement différente des autres caractéristiques de l'événement. En effet, en prédétermination, la date des événements passés est uniquement nécessaire pour évaluer la fréquence moyenne d'occurrence des avalanches, - i.e. le nombre d'avalanches moyen par couloir et par an- et permettre le passage de la loi marginale d'une caractéristique telle la distance d'arrêt à la période de retour. Plutôt que de s'intéresser à la variable « date » difficile à modéliser on ne considère donc que les nombres d'avalanches annuels

(a_t) , $t \in [1, T_{obs}]$ vérifiant $\sum_t^{T_{obs}} a_t = P$. On modélise ce jeu de données par un modèle fréquentiel

$l(a | \theta_F)$ indexé par l'inconnue θ_F et dont l'espérance $E[A | \theta_F]$ est nécessaire au calcul de périodes de retour.

Au cours de l'hiver t , le couloir avalancheux est soumis à un forçage climatique x_o . Celui-ci peut être vu comme la réalisation d'un vecteur de variables conditionnantes X_o quantifiant les conditions climatiques au cours de l'hiver. Il inclut de ce fait des quantités cruciales pour l'activité avalancheuse telles les températures, les précipitations, les vitesses et directions des vents..., etc. On modélise le nombre d'avalanches au cours de l'hiver t par le modèle probabiliste $A | \theta_F, x_o \sim p(\theta_F, x_o)$. Sachant l'état x_o des variables conditionnantes X_o , un mécanisme aléatoire indexé par θ_F génère donc la partie non expliquée du nombre annuel d'avalanches a_t .

En particulier, l'instant t_d du déclenchement est partiellement expliqué par l'état des variables conditionnantes au moment du déclenchement $x_o(t_d)$. X_o quantifie en effet le climat tout au long de l'année et inclut de ce fait le forçage climatique "instantané" (température journalière, chutes de neige récentes..., etc.) à l'origine du déclenchement. Pour simplifier les notations, l'indexage par t_d est cependant abandonné. x_o explique également partiellement le vecteur x des variables d'entrée de

l'avalanche puisque les conditions météorologiques au moment du déclenchement déterminent également les caractéristiques d'entrée de l'avalanche (surface de départ, volume et qualité de la neige, ... etc.). Comme précédemment, un mécanisme aléatoire est introduit pour générer la partie non expliquée de X , de sorte que $X | \theta_o, x_o \sim p(\theta_o, x_o)$.

La propagation de l'avalanche transforme le vecteur d'entrée x en un vecteur de sortie y considéré comme une réalisation d'un vecteur aléatoire Y dont la distribution caractérise les aléas de référence nécessaires à la prédétermination. Un modèle probabiliste de type $Y | \theta_G, x \sim p(x, \theta_G)$ pourrait être utilisé. Cependant, les modèles de propagation déterministes sont supposés suffisants pour représenter le phénomène, de sorte que l'on n'introduit pas d'effet aléatoire dans la propagation et que l'on privilégie un modèle de type $y | \theta_G = G_{\theta_G}(x)$. Cette notation indique que le modèle de propagation déterministe G est choisi parmi une famille paramétrique $G = G_{\theta_G}$ avec le paramètre θ_G qui doit être estimé à partir des données.

3.2.1.3. Description des variables latentes

Le frottement de la neige au cours de son écoulement dépend de la topographie du site (pente, état de la surface) mais également de la qualité de la neige et de l'état de la surface au moment du déclenchement (neige sur toute la longueur du couloir ou pas, présence de glace, etc.). Par conséquent, les coefficients de frottement du modèle de propagation ne peuvent être tous des paramètres au sens statistique du terme et l'un d'entre eux au moins doit être représenté par une variable latente. Le jeu de variables d'entrée doit donc être partagé entre des variables observables de terrain x' (altitude et surface de départ par exemple) et le ou les coefficient(s) de frottement latent(s) x'' , de sorte que le vecteur des entrées devient $X = X', X'' | \theta_o, x_o \sim p(\theta_o, x_o)$. Le paramètre θ_G du modèle de propagation correspond alors au(x) coefficient(s) de frottement non latent(s) et à d'éventuels autres paramètres.

Enfin, l'introduction d'opérateurs stochastiques décrivant les erreurs d'observation est nécessaire pour rendre compte des différences entre les variables phénoménologiques et les observations. On note $X_{obs} | \theta_x, x \sim p(\theta_x, x)$, $Y_{obs} | \theta_y, y \sim p(\theta_y, y)$ et $A_{obs} | \theta_a, a \sim p(\theta_a, a)$, ce qui veut dire que sachant la réalisation x (respectivement y , a) de la variable aléatoire X (respectivement Y , A), la distribution de l'observation autour de la vraie valeur est indexée par l'inconnue θ_x (respectivement θ_y , θ_a).

L'ensemble du cadre formel proposé est résumé en Figure 3.2 sous la forme d'un graphique acyclique orienté (DAG, Lauritzen, 1996) dont les flèches expriment la dépendance conditionnelle, les ellipses représentent les variables latentes et les rectangles les variables observées (Spiegelhalter et al., 1996). Celui-ci permet de bien visualiser que la spécification du modèle se fait sur la base de lois conditionnelles avec un conditionnement correspondant à l'enchaînement des mécanismes physiques : forçage météorologique puis déclenchement, propagation et arrêt. La loi jointe des variables du modèle s'obtient grâce au calcul des probabilités composées à partir des lois conditionnelles. Soit ici $p(x', x'', y) = p(y | x', x'') \times p(x'' | x') \times p(x')$.

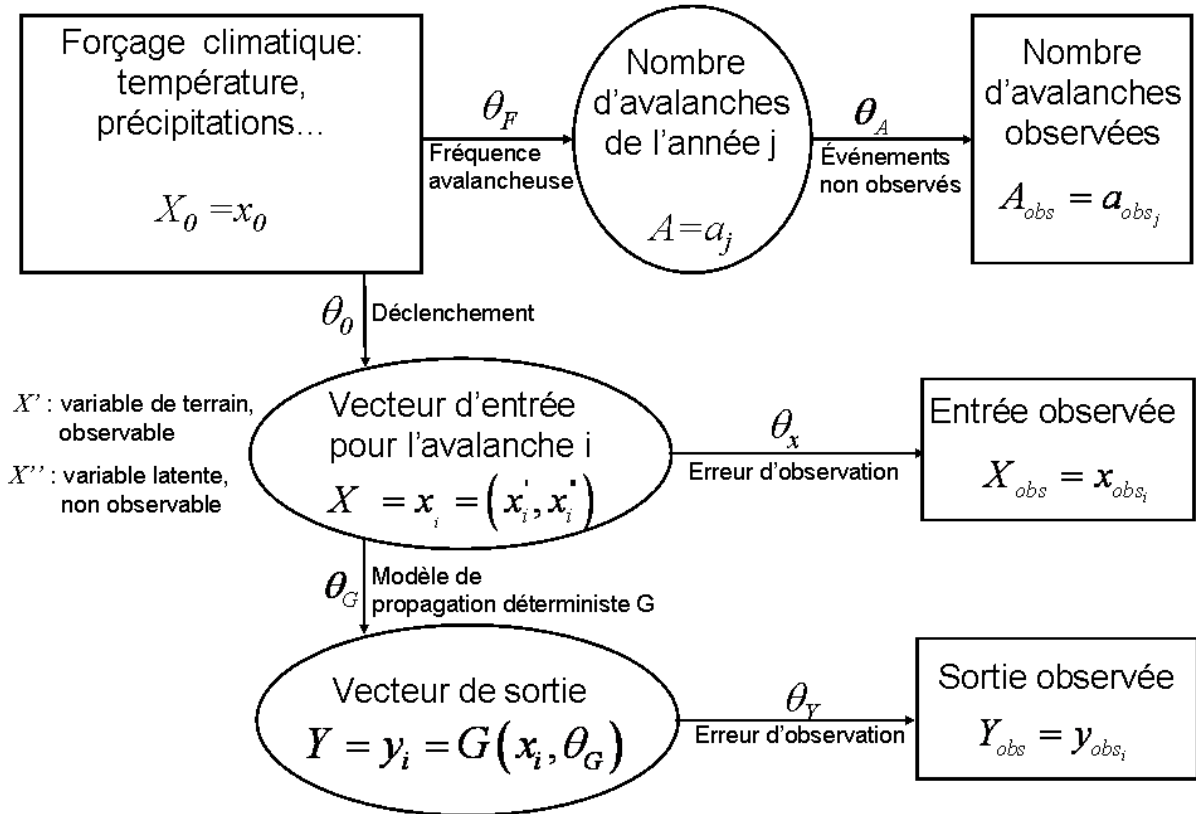
3.2.1.4. Encodage du savoir a priori

La mise en œuvre de l'inférence bayésienne du modèle nécessite la disponibilité d'une loi *a priori* $p(\theta) = p(\theta_M, \theta_F, \theta_X, \theta_Y, \theta_A)$. Celle-ci, comme indiqué au chapitre 2, traduit la connaissance de l'ensemble des inconnues du problème préalablement à l'analyse des données. Le mode de la distribution correspond au meilleur pari avant observation tandis que la dispersion de $p(\theta)$ traduit le degré de certitude *a priori*. Par rapport à l'inférence classique, il s'agit d'un travail de modélisation supplémentaire souvent critiqué pour son caractère non objectif, mais offrant en contrepartie la possibilité d'inclure dans le modèle la connaissance qualitative et diffuse issue de l'expertise (Kadane et Wolfson, 1998).

Dans le cadre de cette thèse, l'incorporation de connaissances expertes grâce à une élicitation de croyance rudimentaire est employée. Cependant, deux autres approches sont également utilisées : le recours à des *priors* non informatifs et l'apprentissage séquentiel, c'est à dire le recours au *posterior*

issu de l'inférence sur un couloir d'apprentissage comme loi *a priori* sur un autre couloir « proche ». Les *priors* non informatifs, abondamment utilisée dans l'ensemble de la thèse, permettent de s'affranchir en partie de la critique de la subjectivité de l'approche bayésienne. L'apprentissage séquentiel, d'une subjectivité « intermédiaire » entre loi *a priori* non informative et loi *a priori* dérivée de l'expertise, permet le transfert d'information d'un couloir à un autre de manière moins brutale et plus rigoureuse que le mélange de données issues de sites différents.

Figure 3.2 : Description systémique du fonctionnement d'un site avalancheux



Le vecteur des paramètres $\theta = (\theta_M, \theta_F, \theta_x, \theta_y, \theta_A)$ correspond aux paramètres des modèles d'intensité, de fréquence et aux modèles d'erreurs d'observation. Les paramètres $\theta_M = (\theta_0, \theta_G)$ du modèle d'intensité se décomposent en paramètres de déclenchement et de propagation. Les vraies valeurs (x, y, a) , bruitées par l'observation, ainsi que le(s) coefficient(s) de frottement variant d'une avalanche à l'autre sont des variables latentes.

3.2.2. Effectuer des hypothèses simplificatrices

3.2.2.1. Négliger les erreurs d'observation dans la modélisation et l'inférence

Les opérateurs stochastiques décrivant les erreurs d'observation sont négligés dans la modélisation et l'inférence, de sorte que $x_i = x_{obs_i}, i \in [1, N]$, $y_i = y_{obs_i}, i \in [1, N]$ et $a_i = a_{obs_i}, t \in [1, T_{obs}]$. Les paramètres $(\theta_x, \theta_y, \theta_A)$ disparaissent donc tandis que les variables latentes (X', Y, A) se confondent avec les observations $(X_{obs}, Y_{obs}, A_{obs})$. Le vecteur $(\theta, x) = (\theta_0, \theta_G, \theta_F, x)$ résume alors l'ensemble des inconnues du problème qui s'en trouve grandement simplifié.

Ce choix est effectué pour garantir l'identifiabilité des différentes inconnues, les variables latentes et les paramètres des modèles d'erreur ne pouvant être inférés simultanément. Afin de le rendre tenable pour des données aussi bruitées que l'EPA, un prétraitement est nécessaire pour éliminer avant inférence les données trop douteuses susceptibles de biaiser les résultats. L'hypothèse qui est effectuée

revient donc à considérer que les erreurs résiduelles après ce prétraitement sont négligeables par rapport à la variabilité naturelle du phénomène étudié.

Pour le modèle fréquentiel, la seule source d'erreur à considérer est la non exhaustivité de la collecte, c'est-à-dire que d'éventuels événements manquants doivent être détectés. Un prétraitement automatisé basé sur des tests de stationnarité de la série des nombres d'observations a été mis en place. Il sera détaillé au chapitre 5 qui décrit le modèle fréquentiel et son implémentation.

En revanche, pour le modèle d'intensité, une multitude de sources d'erreurs détaillées au chapitre 2.1.3.4 doivent être considérées. La détection automatique des erreurs dans les données d'intensité est de ce fait extrêmement difficile. Plutôt que d'essayer de formuler un modèle d'erreur offrant une bonne représentation de la combinaison des différentes sources d'erreur pour chacune des variables marginales, il a été fait l'hypothèse que quelques séries d'événements peu bruitées pouvaient être sélectionnées sur la base d'une analyse experte au cas par cas. Les critères de sélection seront présentés au chapitre 3.3 et les jeux de données retenus en Annexe. La validité de l'approche sera discutée au fur et à mesure de l'exploitation des cas d'étude choisis.

3.2.2.2. Indépendance intensité/fréquence

Une hypothèse simplificatrice importante pour la prédétermination est l'indépendance entre l'intensité des avalanches et le nombre d'avalanches se produisant au cours d'un hiver. L'intérêt de cette hypothèse est de pouvoir séparer la modélisation de l'intensité de la modélisation de la fréquence (eq 3.1). Ainsi l'inférence des deux modèles peut être effectuée séparément. Ensuite, la période de retour $T(y)$ peut être calculée en combinant simplement l'espérance du modèle fréquentiel $E[A|\theta_F]$ avec la fonction de répartition $F(y|\theta_M) = P(Y \leq y|\theta_M)$ de la variable marginale Y du modèle d'intensité (eq 3.2).

$$l(x, y, a|\theta) = l(a|\theta_F) \times l(x, y|\theta_M) \quad (\text{eq 3.1})$$

$$T(y) = \frac{1}{E[A|\theta_F] \times (1 - F(y|\theta_M))} \quad (\text{eq. 3.2})$$

Cette hypothèse est vérifiée si le nombre total d'avalanches se produisant au cours d'un même hiver sur un site donné est faible. En effet, les événements successifs se produisent alors dans des contextes (météorologie, état de couloir, matériau neigeux disponible, etc.) différents, le couloir ayant à chaque fois le temps d'« oublier » le précédent événement. Au contraire, pour un couloir très actif, la neige disponible et l'état du couloir n'ont pas le temps de se reconstituer entre deux événements, de sorte que les caractéristiques quantitatives des événements successifs d'un hiver dépendent du nombre total d'avalanches enregistrées au cours de l'hiver. Pour qu'en pratique l'hypothèse soit vérifiée, on a donc en priorité travaillé sur des couloirs pas trop actifs.

3.2.2.3. Déclenchements aléatoires plutôt que variables conditionnantes climatiques

Les liens déterministes entre les variables climatiques conditionnantes journalières (chute de neige, température, etc.) et les variables d'entrée (x_i', x_i'') , $i \in [1, N]$ ainsi qu'entre les variables conditionnantes moyennes annuelles et les nombres d'avalanches annuels a_t , $t \in [1, T_{obs}]$ restent, comme on l'a vu au chapitre 2, mal connus. On peut donc penser que ne pas modéliser explicitement le lien physique entre climat et avalanche n'affecte pas beaucoup l'efficacité opérationnelle d'un modèle de prédétermination. En conséquence, on fait l'hypothèse que des opérateurs purement aléatoires peuvent être utilisés pour représenter le forçage climatique, ce qui se traduit par $a_t \sim l(\theta_F)$, $t \in [1, T_{obs}]$ et $x_i \sim l(\theta_o)$, $i \in [1, N]$.

3.2.2.4. Echelle de description

On se limite à une approche bidimensionnelle : l'avalanche se déplace le long d'un profil curvilinéaire dont l'équation dans un repère cartésien est de type $z=f(x)$ où z est l'altitude et x la distance mesurée le long d'un axe horizontal dont le zéro est fixé à l'abscisse correspondant au sommet du couloir. Le lecteur est invité à prendre garde à une possible confusion entre cette notation et les variables aléatoires introduites précédemment.

La principale limite de l'hypothèse bidimensionnelle est la non prise en compte des variations de l'extension latérale de l'avalanche. Cette hypothèse est toutefois effectuée couramment en ingénierie paravalanche car elle présente l'avantage d'assurer un lien univoque entre altitude et distance horizontale. Une seule coordonnée suffit alors à identifier la position de l'avalanche si l'on néglige son épaisseur par rapport au dénivelé du couloir. De plus, une description bidimensionnelle de l'écoulement réduit considérablement le temps de calcul. Notons qu'elle n'affecte en rien le cadre formel de l'approche proposée et qu'une extension tridimensionnelle basée sur les méthodes numériques (Gruber et al., 1998) ou statistiques (Keylock et Barbolini, 2001) existantes pourrait être proposée sans difficultés conceptuelles majeures, mais au prix de temps de calculs nettement augmentés.

Une échelle de description doit ensuite être choisie. Deux options seront explorées : dans un premier temps, les calculs seront effectués avec un modèle très simple de type bloc glissant. Pour des pentes régulières sans singularités topographiques, les forces de courbures sont négligeables et l'équation du mouvement est alors rudimentaire : la variation de la vitesse du bloc v au point de pente locale ϕ correspond à la différence entre la composante horizontale de la force de gravité g et un terme de frottement noté génériquement $Frot$ et s'exprimant en $[m/s^2]$ (eq. 3.3).

Dans un second temps, un modèle fluide basé sur le formalisme de Barré de Saint-Venant sera utilisé. Celui-ci permet une représentation plus réaliste de l'avalanche en donnant accès à l'évolution de la hauteur h et de la vitesse v au cours de l'espace et du temps. Les phénomènes d'érosion/dépôt seront négligés, de sorte que la masse de l'avalanche sera conservée au cours de l'écoulement (eq. 3.4, haut). La variation de la quantité de mouvement est liée, tout comme pour le modèle de bloc glissant, à la différence entre gravité et frottement, mais l'intégration sur la verticale fait apparaître deux coefficients supplémentaires α_{sv} et k_{sv} . α_{sv} traduit la forme du profil de vitesse et k_{sv} l'anisotropie de la neige durant son écoulement (eq. 3.4, bas). Afin de ne pas alourdir l'inférence et conformément à la pratique en modélisation des avalanches, ces coefficients seront fixés.

Des observations et des simulations numériques ont montré que la densification de la neige s'effectue principalement immédiatement après le déclenchement et que la prise en compte de la densification au cours de l'écoulement modifie peu les distances d'arrêt et les profils de vitesse, sauf pour les très petites avalanches. Comme le montrent les équations de conservation retenues, la variation de la masse volumique au cours de l'écoulement sera donc négligée pour les descriptions autant « bloc glissant » que « fluide » bien que la neige soit un matériau compressible.

$$\frac{dv}{dt} = g \times \sin \phi - Frot \quad (\text{eq. 3.3})$$

$$\begin{cases} \frac{\partial h}{\partial t} + \frac{\partial(hv)}{\partial x} = 0 \\ \frac{\partial(hv)}{\partial t} + \frac{\partial}{\partial x} (\alpha_{sv} \times hv^2 + k_{sv} \times g \frac{h^2}{2}) = h \times [g \sin \phi - Frot] \end{cases} \quad (\text{eq. 3.4})$$

3.2.2.5. Loi de frottement

Le terme de frottement est, comme mentionné au chapitre 2, un sujet de recherche et d'âpres débats du fait de la connaissance limitée de la rhéologie de la neige. Son choix reste donc, au moins en partie, discutable. On se limite dans cette thèse à deux modèles simples : celui, « historique », de Voellmy (1955) noté F_{voel} et possédant deux coefficients de frottement μ_{voel} et ξ_{voel} (eq. 3.5) et celui, encore plus rudimentaire, de Coulomb noté F_{coul} et possédant un seul coefficient de frottement μ_{coul} (eq. 3.6). μ_{voel} et μ_{coul} sont adimensionnels tandis que ξ_{voel} s'exprime en $[m/s^2]$. Notons que le modèle de Coulomb correspond à la limite du modèle de Voellmy pour de très fortes valeurs de ξ_{voel} .

$$F_{voel} = \mu_{voel} \times g \times \cos \phi + \frac{g}{\xi_{voel} \times h} \times v^2 \quad (\text{eq. 3.5})$$

$$F_{coul} = \mu_{coul} \times g \times \cos \phi \quad (\text{eq. 3.6})$$

Puisque l'un des coefficients de frottement au moins varie d'une avalanche à l'autre du fait des variations de la qualité de la neige, plusieurs choix de modélisation sont possibles pour le modèle de

Voellmy : soit μ_{voel} est considéré comme latent et ξ_{voel} est considéré comme un paramètre, soit l'inverse, soit deux variables latentes sont considérées. Afin d'éviter le surparamétrage, l'option de deux coefficients de frottement latents a été écartée. En outre, afin de rester cohérent avec le modèle de Coulomb ainsi qu'avec l'interprétation de la directive suisse du modèle de Voellmy (Salm et al., 1990), il a été choisi de poser $x'' = \mu_{voel}$ et de considérer que ξ_{voel} reflète davantage des caractéristiques topographiques du couloir et se comporte de ce fait comme un paramètre.

Avec une loi de frottement coulombienne, au contraire, on n'a qu'un seul coefficient de frottement. μ_{coul} doit donc nécessairement être modélisé par une variable. Néanmoins, par construction, le modèle de Coulomb est inversible (on y reviendra au chapitre 4). μ_{coul} peut donc être considéré comme observé de sorte que la variable latente disparaît du modèle.

Pour résumer, avec le modèle de Coulomb $\theta_G = x'' = \emptyset$ et μ_{coul} se comporte comme l'une des variables observables x' , tandis qu'avec le modèle de Voellmy $x'' = \mu_{voel}$ et $\theta_G = \xi_{voel}$. La différence fondamentale entre les deux modèles de frottement est qu'avec le modèle de Coulomb, l'avalanche peut accélérer indéfiniment dès lors que la pente est supérieure à μ_{coul} , tandis qu'avec le modèle de Voellmy, elle n'accélère que jusqu'à une vitesse limite quand la pente est supérieure à μ_{voel} . En conséquence, les vitesses calculées avec le modèle de Coulomb sont souvent peu réalistes car beaucoup trop élevées alors qu'elles sont plus fiables mais parfois sous estimées avec le modèle de Voellmy.

3.2.2.6. Indépendance des événements

L'indépendance mutuelle des événements est une hypothèse de commodité mathématique très usuelle en statistique. La vraisemblance d'un jeu de données est alors simplement le produit des vraisemblances de chaque événement.

Cette hypothèse est facilement justifiable pour le modèle fréquentiel, l'indépendance du nombre d'événements se produisant durant deux hivers consécutifs étant relativement évidente (eq. 3.7). Pour le modèle d'intensité (eq. 3.8), elle est fortement liée à l'hypothèse d'indépendance entre intensité et fréquence. En effet, si toutes les avalanches se produisant au cours d'un même hiver sur un site donné sont séparées par des intervalles de temps longs, leurs caractéristiques quantitatives peuvent être considérées comme physiquement indépendantes car causées par des forçages météorologiques différents.

En pratique, on considère l'indépendance physique assurée entre deux avalanches lorsqu'elles sont séparées par plus d'un mois. De plus, pour des doublets d'événement séparés par moins de un mois, des tests statistiques standards permettent souvent d'établir l'indépendance statistique. Au final, l'hypothèse d'indépendance mutuelle des événements est donc presque toujours justifiée et les difficultés de modélisation qu'entraîne sa remise en cause ne paraissent pas justifiées. Elle sera donc toujours effectuée après étude préalable des dates d'occurrence des événements et mise en œuvre de tests standards (cf. Annexe et cas d'études).

$$l(a) = l(a_1, a_2, \dots, a_{T_{obs}}) = \prod_{t=1}^{T_{obs}} (l(a_t)) \quad (\text{eq. 3.7})$$

$$l(x, y) = l(x_1, y_1, x_2, y_2, \dots, x_N, y_N) = \prod_{i=1}^N (l(x_i, y_i)) \quad (\text{eq. 3.8})$$

3.2.2.7. Cadre paramétrique « simple »

On se restreint à un cadre paramétrique, de sorte que $\dim(\theta) < +\infty$. Ainsi, même si de la souplesse est perdue pour la modélisation, les difficultés d'inférence, déjà grandes du fait de l'emploi d'un modèle de propagation non linéaire, restent raisonnables.

De plus, on effectue l'hypothèse que des lois de probabilité relativement simples appartenant aux classes de lois courantes (familles exponentielles) sont susceptibles de bien représenter la variabilité des données. Cela n'est cependant pas incompatible avec une modélisation souple permettant de s'adapter à la variabilité des données grâce au recours à la modélisation hiérarchique et en particulier aux lois de mélange.

3.2.2.8. Indépendance a priori

La mise en œuvre de l'inférence bayésienne nécessite la spécification d'une loi *a priori* jointe pour l'ensemble des paramètres du modèle. Dans l'ensemble des applications, l'hypothèse d'indépendance *a priori* sera effectuée, de sorte que le *prior* joint est simplement le produit des lois *a priori* marginales (eq. 3.9). Ce choix, usuel en inférence bayésienne, n'est que très peu contraignant car la corrélation inter paramètres pouvant exister est prise en compte *a posteriori*. Dans le cas de *priors* non informatifs, elle est même absolument sans conséquence. Dans le cas de *priors* informatifs, elle facilite la démarche d'élicitation si des formes paramétriques simples sont choisies pour les loi *a priori* marginales en permettant l'expression de paris sur des grandeurs significantes (moyenne et/ou certains quantiles des lois *a priori* marginales).

$$p(\theta) = \prod_{k=1}^{\dim(\theta)} (p(\theta_k)) \quad (\text{eq. 3.9})$$

3.2.2.9. Période de retour pour la distance d'arrêt uniquement

On choisit de n'estimer des périodes de retour qu'avec la variable marginale distance d'arrêt x_{stop} . Du fait du caractère bidimensionnel de l'approche retenue, cette dernière s'identifie à l'abscisse d'arrêt, c'est-à-dire à la distance mesurée le long d'un axe horizontal dont le zéro est fixé à l'abscisse correspondant au sommet du couloir. $T_{x_{stop}}$ représente alors le temps moyen en années séparant deux avalanches dépassant l'abscisse x_{stop} (eq. 3.10). Ce choix, discuté au chapitre 8, est guidé par l'impossibilité théorique de considérer un aléa de référence unique pour un phénomène multivarié. D'un point de vue plus pratique il est également lié au caractère crucial de la variable distance d'arrêt pour le zonage et à la structure des données EPA. L'altitude d'arrêt constitue en effet le principal support d'apprentissage et par conséquent la sortie la plus fiable de n'importe quel modèle de propagation.

$$T_{x_{stop}} = \frac{1}{E[A|\theta_F] \times (1 - F(x_{stop}|\theta_M))} \quad (\text{eq. 3.10})$$

Toutefois, définir une avalanche de référence à partir de la période de retour associée à sa distance d'arrêt n'exclut pas d'examiner les autres sorties du modèle de propagation. On s'intéressera en particulier aux distributions des vitesses, hauteurs et nombre de Froude en un point quelconque de la zone d'arrêt de façon à étudier pour la prédétermination différents scenarii de référence correspondant à une même période de retour.

3.2.2.10. Modélisation de la pression

La pression Pr s'exprime de la façon la plus générale comme le produit entre l'énergie cinétique par unité de volume et un coefficient Cx dépendant des propriétés de l'écoulement (eq. 3.11). Le calcul de pression nécessite donc la connaissance de la masse volumique ρ_N de la neige en écoulement ainsi que la valeur du Cx . ρ_N est très difficile à modéliser. De plus, la détermination du Cx est délicate. La pression ne peut donc pas être obtenue à partir des modèles stochastiques proposés sans introduire de nouvelles sources d'incertitudes liées à l'évaluation de la masse volumique de la neige en écoulement et à la nature de la dépendance de la pression avec la vitesse.

$$Pr = Cx \times \frac{1}{2} \rho_N \times v^2 \quad (\text{eq. 3.11})$$

On se limitera à la pression théorique en écoulement libre, c'est-à-dire à la pression à Cx fixé et égal à 2, et aux pressions obtenues à partir soit d'une loi théorique, soit d'une loi expérimentale. La loi théorique a été obtenue par Naaïm (2007) par analogie avec un fluide visqueux. Son expression sera détaillée au chapitre 6.5 (eq. 6.6). La loi expérimentale a été calibrée sur différentes avalanches du site test suisse de la vallée de la Sionne (Sovilla et al., 2007). Elle traduit empiriquement la faible diminution de la pression avec la dynamique de l'écoulement sous la forme d'une fonction puissance du nombre de Froude Fr , avec un couple de paramètres ν_1 et ν_2 pour chaque événement (eq. 3.12).

$$Cx \propto \nu_1 \times Fr^{-\nu_2} \quad (\text{eq. 3.12})$$

Dans un premier temps, on considèrera la masse volumique de la neige et les paramètres de la loi expérimentale comme fixés. On calculera ainsi la distribution moyenne *a posteriori* des pressions d'impact par intégration sur la variabilité des propriétés de l'écoulement. Dans un second temps, on évaluera la sensibilité des pressions obtenues à la masse volumique en supposant connue la loi de probabilité $p(\rho_N)$. De ce fait, on effectuera le calcul de la loi prédictive de la pression par rapport à la variable ρ_N .

3.3. Inférence bayésienne du modèle proposé

3.3.1. Sélection et prétraitement des données pour la prédétermination locale

3.3.1.1. Principe

L'EPA fournit une information abondante mais en quantité variable selon les sites considérés et, comme détaillé au chapitre 2, très bruitée. De plus, la variabilité topographique inter site est très importante notamment en termes de profil de pente (de très régulier à très irrégulier) et de chenal d'écoulement (du couloir rectiligne au site dit « à branche » présentant plusieurs zones de départ et/ou d'arrivée potentielles). Avant de traiter les problèmes pratiques liés à la pénurie de données, à leur imperfection et à la complexité de la topographie, il convient de développer un modèle de prédétermination locale qui fonctionne correctement sur un site simple et bien documenté. Des couloirs relativement réguliers pour lesquels une information relativement fiable et de quantité suffisante était disponible ont donc été choisis. Les critères de sélection et les opérations de prétraitement effectuées ont été les suivantes :

- Les modèles proposés ne prenant pas en compte les erreurs d'observations, seuls des jeux de données supposés fiables ont été sélectionnés. Le prétraitement a été effectué manuellement et les considérations expertes utilisées sont détaillées dans l'Annexe qui présente les jeux de données retenus.
- L'emploi d'une méthode statistique nécessite qu'au moins un nombre minimal de données ait été collecté. Pour fixer les idées, disons qu'une vingtaine de données constitue une base de travail solide. Seuls des sites pour lesquels au moins vingt événements relativement fiables ont pu être identifiés ont donc été sélectionnés.
- Les hypothèses d'indépendance inter événements et intensité-fréquence doivent être étayées sur la base de considérations physiques et/ou statistiques. Les vérifications effectuées sur les jeux de données sélectionnés sont détaillées en Annexe.
- D'un point de vue de la topographie, l'existence d'une population sous-jacente unique ne peut être postulée sous hypothèse bidimensionnelle pour les sites EPA dits « à branche » où les avalanches empruntent des chenaux d'écoulement différents, de sorte que de tels sites ont été écartés.
- Enfin les modèles de propagation d'avalanche sont très sensibles à la pente locale. Ils sont de ce fait susceptibles de mal fonctionner lorsque des singularités topographiques existent. Par conséquent, les sites sélectionnés présentent une topographie relativement régulière avec des changements de concavités pas trop rapides.

L'ensemble de ces prérequis peut sembler assez lourd, mais, outre le fait qu'ils sont utilisés implicitement dans les approches statistiques-dynamiques existantes, ils ne filtrent qu'une partie relativement faible des sites EPA. Le nombre de sites avalancheux français sur lesquels les modèles développés sont applicables reste donc malgré tout assez important

3.3.1.2. Sites sélectionnés et transformation des données

Quatre sites EPA correspondant aux prérequis ont été sélectionnés pour illustrer les modèles développés dans le cadre de cette thèse. L'ensemble des informations a été examiné « manuellement » de façon à écarter les événements les plus douteux. Les sites choisis et les jeux de données correspondant seront brièvement introduits lorsqu'ils seront utilisés. Leur topographie ainsi que l'information historique disponible sont également présentées en détail en Annexe. En outre, sur l'un de ces sites, un jeu de données a été simulé afin de vérifier le bon fonctionnement des algorithmes d'inférence proposés au chapitre 4.

Les valeurs observées d'altitudes de départ et d'arrêt ont été converties en abscisses de départ x_{start} et d'arrivée x_{stop} grâce à la relation topographique $z=f(x)$. Les dimensions de dépôt ont été retrouvées dans les carnets forestiers lorsqu'elles ne figuraient pas dans la base EPA. Elles ont permis le calcul des volumes de dépôt v_{stop} selon la formule du « triangle camembert » présentée au chapitre 2. On travaille donc en considérant $x_{start}, x_{stop}, v_{stop}$ comme observés alors qu'ils ne le sont pas réellement.

3.3.2. Inférence

L'application du théorème de Bayes au modèle proposé permet l'obtention de la loi *a posteriori* des inconnues du problème. L'indépendance postulée entre intensité et fréquence autorise l'inférence séparée des paramètres caractérisant l'intensité et la fréquence des avalanches. Si l'on considère uniquement la prédétermination locale, c'est-à-dire la prédétermination sur un site EPA ou des données suffisantes sont disponibles, l'inférence du modèle fréquentiel est très simple puisque seul le paramètre θ_F est inconnu (eq. 3.13). Les conséquences sur l'inférence du modèle fréquentiel de l'introduction de niveaux hiérarchiques représentant des dimensions spatiales et spatio-temporelles supplémentaires seront toutefois discutées au chapitre 5. Pour le modèle d'intensité, les hypothèses simplificatrices effectuées permettent de restreindre les inconnues au coefficient latent x'' et au paramètre $\theta_M = (\theta_o, \theta_G)$ incluant les paramètres θ_o décrivant le vecteur d'entrées et le coefficient de frottement non latent $\theta_G = \xi_{voel}$ lorsque la loi de frottement de Voellmy est utilisée (eq. 3.14).

$$p(\theta_F | a) \propto p(\theta_F) \times l(a | \theta_F) \quad (\text{eq. 3.13})$$

$$p(\theta_M, x'' | x, y) \propto p(\theta_M) \times l(x, y | \theta_M, x'') \times p(x'' | \theta_M, x, y) \quad (\text{eq. 3.14})$$

Le calcul pratique des lois *a posteriori* suppose la mise en œuvre des algorithmes d'estimation mentionnés au chapitre 2. Pour chaque modèle, l'écriture du code de calcul *ad hoc* s'effectue en fonction de la structure du modèle, de sorte que les problèmes spécifiques liés à chaque inférence seront abordés lorsque les modèles seront présentés en détail. Signalons d'ores et déjà simplement qu'avec le modèle inversible de Coulomb, l'inférence peut être dissociée de la propagation, de sorte qu'elle ne présente pas de difficultés majeures et qu'elle peut être menée à bien à partir d'un logiciel spécialisé comme Winbugs (Spiegelhalter et al., 2000). En revanche, avec une loi de frottement de Voellmy, estimation et propagation ne peuvent être dissociées, de sorte que l'écriture d'un algorithme d'estimation est indispensable. Dans le cadre de cette thèse, l'environnement Matlab a été utilisé. Enfin, pour le modèle spatial et spatio temporel du chapitre 5, l'inférence est difficile étant donné les modèles considérés mais la résolution du problème ne faisant pas intervenir de code de mécanique des milieux continus, la plupart des calculs statistiques peuvent être menés sous Winbugs.

3.4. Prédiction par mise en œuvre des méthodes de Monte Carlo

3.4.1. Prédiction moyenne : moyenne *a posteriori* du modèle

Une fois l'inférence effectuée, des estimateurs $\left(\hat{\theta}_M, \hat{\theta}_F \right)$ peuvent être obtenus pour les paramètres du modèle. La moyenne *a posteriori* est le choix le plus traditionnel en tant que pari moyen minimisant une perte quadratique.

La moyenne *a posteriori* du modèle fréquentiel $p\left(A \mid \hat{\theta}_F \right)$ est très facile à obtenir puisqu'une formulation explicite du modèle est disponible. En revanche, pour le modèle d'intensité, la moyenne *a posteriori* des observables $p\left(x', y \mid \hat{\theta}_M \right)$ ne peut être obtenue que par simulation car on ne dispose pas de formulation explicite de la loi des sorties y conditionnellement à celle de x' . La difficulté est encore accentuée par la nécessité de tenir compte de la variable latente x'' . Le calcul se fait donc par

l'intermédiaire de la loi jointe $p\left(x', x'', y \middle| \hat{\theta}_M\right)$ en intégrant sur la variable latente « de nuisance » et en conditionnant de façon à faire apparaître les lois qui sont connues explicitement grâce à la déclaration conditionnelle du modèle : la loi des entrées observables $p\left(x' \middle| \hat{\theta}_M\right)$, celle du coefficient de frottement latent conditionnellement aux entrées observables $p\left(x'' \middle| \hat{\theta}_M, x'\right)$ et celle des sorties conditionnellement aux entrées $p\left(y \middle| \hat{\theta}_M, x', x''\right)$ (eq. 3.15). Numériquement, le calcul peut s'effectuer

à l'aide de tirages de Monte Carlo, ce qui apparente bien l'approche proposée aux méthodes statistiques dynamiques. Au passage, on obtient, en « laissant de côté » le conditionnement par les entrées observables, un échantillon de la moyenne *a posteriori* de la loi du coefficient latent

$$p\left(x'' \middle| \hat{\theta}_M\right).$$

$$p\left(x', y \middle| \hat{\theta}_M\right) = \int p\left(x', x'', y \middle| \hat{\theta}_M\right) \times dx'' = \int p\left(y \middle| \hat{\theta}_M, x', x''\right) \times p\left(x'' \middle| \hat{\theta}_M, x'\right) \times p\left(x' \middle| \hat{\theta}_M\right) \times dx'' \quad (\text{eq. 3.15})$$

L'estimation de la période de retour $T_{x_{stop}}$ associée à la distance d'arrêt x_{stop} s'obtient aisément (eq. 3.16) par combinaison de l'estimateur de la moyenne du modèle fréquentiel, c'est-à-dire du nombre moyen d'avalanches par an pour le couloir considéré, et de la moyenne *a posteriori* de la fonction de

répartition de la distance d'arrêt $\hat{F}(x_{stop}) = P\left(X_{stop} \leq x_{stop} \middle| \hat{\theta}_M\right)$. Cependant, en pratique, le problème

est souvent inverse dans la mesure où l'on recherche la distance d'arrêt correspondant à une probabilité de dépassement annuelle donnée (0.1 ou 0.01 par exemple). Il s'agit donc d'obtenir le quantile de distance d'arrêt x_{stop_T} correspondant à la période de retour T fixée et dont la probabilité de non dépassement est donnée par inversion de l'équation 3.16. Le calcul s'effectue à l'aide de la fonction de répartition inverse de la distance d'arrêt $F^{-1}_{x_{stop} \middle| \hat{\theta}_M}$ (eq. 3.17). De façon pratique, si

$(x_{stop_1}, x_{stop_2}, \dots, x_{stop_n})$ est l'échantillon ordonné obtenu par simulation, l'estimateur empirique de $\hat{F}(x_{stop_k})$ est simplement k/n où n est le nombre de tirages effectués. Le quantile empirique

correspondant à la période de retour T est quant à lui la $k\left(1 - 1/E\left[A \middle| \hat{\theta}_F\right] \times T\right)^{\text{ème}}$ valeur de l'échantillon ordonné.

$$T_{x_{stop}} = \frac{1}{E\left[A \middle| \hat{\theta}_F\right] \times \left(1 - \hat{F}(x_{stop})\right)} \quad (\text{eq. 3.16})$$

$$x_{stop_T} = F^{-1}_{x_{stop} \middle| \hat{\theta}_M} \left(1 - \frac{1}{E\left[A \middle| \hat{\theta}_F\right] \times T} \right) \quad (\text{eq. 3.17})$$

3.4.2. « Intervalles de confiance » associés aux valeurs de projet

La mise en œuvre d'une méthode de Monte Carlo pour le calcul des distances d'arrêt s'accompagne d'une erreur numérique pouvant être quantifiée. L'intervalle de confiance pour, par exemple, la

probabilité de dépassement $1 - \hat{F}(x_{stop})$ de la distance d'arrêt x_{stop} s'obtient en considérant le quantile de loi normale $q_{N_{\alpha_c}}$ correspondant au niveau de confiance α_c voulu (eq. 3.18). Sa largeur dépend du nombre de simulations effectuées et, comme on estime une probabilité (ou un quantile), elle dépend également de la probabilité de dépassement considérée : plus celle-ci est faible et plus le nombre de simulations requis pour obtenir un certain niveau de précision augmente. L'intervalle obtenu n'est cependant pas à proprement parler un intervalle de confiance dans la mesure où sa largeur ne dépend pas de l'information disponible pour effectuer l'inférence, c'est-à-dire du nombre de données. Il permet au statisticien de vérifier que son effort de calcul est suffisant. Cependant, il ne renseigne pas le décideur sur les erreurs associées aux valeurs de projet qui lui sont proposées et inhérentes à la méconnaissance locale du phénomène.

$$IC_{\alpha} = \hat{F}(x_{stop}) \pm q_{N_{\alpha_c}} \times \sqrt{\frac{\hat{F}(x_{stop}) \left(1 - \hat{F}(x_{stop})\right)}{n}} \quad (\text{eq. 3.18})$$

Notons au passage que la convergence des méthodes de Monte Carlo étant en $1/\sqrt{n}$, elle est relativement lente. Il existe bien sûr des méthodes telles la méthode FORM (Skaggs et Barry, 1997) ou l'Importance Sampling (Parent et Bernier, 2007, chapitre 14), permettant d'approcher plus rapidement les estimateurs recherchés. Néanmoins le problème du choix de la méthode de simulation la plus efficace n'a pas été abordé au cours de la thèse car il a été considéré que si l'inférence par MCMC était tolérable en termes de temps de calcul, la simulation par tirage de Monte Carlo devait l'être également. En pratique, il a effectivement toujours fallu bien moins d'itérations du modèle de propagation pour estimer les périodes de retour avec une précision suffisante que pour effectuer l'inférence des paramètres du modèle.

3.4.3. Lois prédictives

La prise en compte de la méconnaissance locale s'effectue au travers du calcul des lois prédictives combinant variabilité intrinsèque du phénomène et erreur d'estimation. La loi prédictive du modèle fréquentiel $p(a^{pred} | data_1)$ où $data_1 = (a_1, a_2, \dots, a_{T_{obs}})$ s'obtient aisément en moyennant le modèle par rapport à la loi *a posteriori* (eq. 3.19).

Pour le modèle d'intensité, le principe est le même avec $data_2 = (x_1, y_1, x_2, y_2, \dots, x_N, y_N)$, mais le calcul nécessite, comme pour la moyenne *a posteriori*, le passage par le coefficient de frottement latent et le conditionnement (eq. 3.20). Numériquement, le calcul s'effectue en deux étapes. La loi prédictive des variables d'entrée $p(x^{pred} | data_2)$ s'obtient aisément en même temps que l'inférence par MCMC en effectuant des tirages du modèle de vraisemblance conditionnellement aux valeurs des

inconnues à chaque itération. Ensuite, le calcul de $p(y^{pred} | data_2, x^{pred})$ s'effectue en transformant la loi prédictive des variables d'entrée à travers le modèle de propagation. Le calcul des lois prédictives s'effectue donc selon une approche de type Monte Carlo qui possède les mêmes avantages et limites en termes de facilité de mise en œuvre et de temps de calcul que celle proposée pour la moyenne *a posteriori*.

$$p(a^{pred} | data_1) = \int l(a^{pred} | \theta_F) \times p(\theta_F | data_1) \times d\theta_F \quad (\text{eq. 3.19})$$

$$p(x^{pred}, y^{pred} | data_2) = \iint p(y^{pred} | \theta_M, x', x'') \times p(x^{pred} | \theta_M, x', x'') \times p(x^{pred} | \theta_M) \times p(\theta_M | data_2) \times d\theta_M \times dx'' \quad (\text{eq. 3.20})$$

Des périodes de retour peuvent bien entendu être calculées à partir de la loi prédictive marginale des distances d'arrêt $p(x_{stop} | data_2)$. Celles-ci tiennent compte de la méconnaissance locale et constituent ainsi le meilleur pari sachant les observations. Du fait de la prise en compte d'une source de variabilité additionnelle liée à l'erreur d'estimation, elles sont plus pessimistes que les périodes de retour « classiques » calculées à partir de la moyenne *a posteriori*. Cependant, ces périodes de retour prédictives ne sont pratiquement jamais utilisées dans le domaine des risques naturels, et *a fortiori* dans celui des avalanches. Une raison acceptable est qu'il est compréhensible de privilégier un aléa de

référence moyen ne mélangeant pas les sources d'incertitude. Toutefois, plus vraisemblablement, la faible popularité des périodes de retour prédictives est liée à la large prédominance des méthodes classiques par rapport aux méthodes bayésiennes en ingénierie.

Pour le cas du calcul de pression, l'obtention des lois prédictives nécessite la prise en compte dans l'équations 3.11 de l'erreur d'estimation de la masse volumique de la neige en écoulement. L'erreur sur la loi expérimentale de l'eq. 3.12 peut éventuellement également être considérée. Les calculs seront menés au chapitre 6.

3.4.4. Intervalles de crédibilité sur les valeurs de projet

Si leur utilité pour le calcul de périodes de retour reste discutable, les lois prédictives présentent un intérêt certain pour le problème de la quantification de l'erreur sur les valeurs de projet en permettant l'obtention d'intervalles de crédibilité sur les quantiles estimés. Ainsi, si l'on note X_{stop_q} la variable aléatoire correspondant au quantile d'ordre q de distance d'arrêt, sa loi prédictive $p(x_{stop_q} | data_2)$ s'obtient en moyennant sur la loi *a posteriori* de l'inconnue θ_M la valeur en $q/100$ de la fonction de répartition inverse sachant l'inconnue θ_M de la distribution des distances d'arrêt $F_{x_{stop}|\theta_M}^{-1}$ (eq. 3.21).

Obtenir la loi prédictive $p(x_{stop_T} | data_1, data_2)$ de la variable aléatoire X_{stop_T} représentant le quantile annuel de distance d'arrêt nécessite de combiner les équations 3.16 et 3.21 de façon à appliquer la fonction de répartition inverse à la probabilité de dépassement pondérée par le nombre moyen d'avalanches par an et à moyenner à la fois sur les lois *a posteriori* des modèles d'intensité et de fréquence (eq 3.22). La loi prédictive obtenue inclut ainsi à la fois l'incertitude touchant l'intensité et la fréquence des avalanches sur le site étudié.

Enfin, il est également possible d'obtenir la loi prédictive $p(T_{x_{stop}} | data_1, data_2)$ de la variable $T_{x_{stop}}$ représentant la période de retour associée à l'abscisse x_{stop} en moyennant sur la loi *a posteriori* de θ_M la probabilité de dépassement $(1 - F(x_{stop} | \theta_M))$ plutôt que la fonction de répartition inverse tout en pondérant toujours par la fréquence avalancheuse moyenne (eq. 3.23). La dispersion des lois prédictives obtenues dépend de la variabilité de la loi *a posteriori* des paramètres (θ_M, θ_F) et donc de la quantité de données disponibles pour l'inférence. Asymptotiquement, elles convergent vers les vraies valeurs ponctuelles de x_{stop_q} , x_{stop_T} et $T_{x_{stop}}$.

$$p(x_{stop_q} | data_2) = \int F_{x_{stop}|\theta_M}^{-1}(q/100) \times p(\theta_M | data_2) \times d\theta_M \quad (\text{eq. 3.21})$$

$$p(x_{stop_T} | data_1, data_2) = \int F_{x_{stop}|\theta_M}^{-1} \left(1 - \frac{1}{E[A|\theta_M] \times T} \right) \times p(\theta_M | data_2) \times p(\theta_F | data_1) \times d\theta_M \times d\theta_F \quad (\text{eq. 3.22})$$

$$p(T_{x_{stop}} | data_1, data_2) = \int \frac{1}{E[A|\theta_M] \times (1 - F(x_{stop} | \theta_M))} \times p(\theta_M | data_2) \times p(\theta_F | data_1) \times d\theta_M \times d\theta_F \quad (\text{eq. 3.23})$$

Numériquement, le calcul des trois lois prédictives des équations 3.21 à 3.23 peut être mené "brutalement" par des simulations de Monte Carlo successives à partir d'un échantillon de taille n_1 de la loi *a posteriori* de (θ_M, θ_F) obtenue par MCMC. Pour chaque valeur $\theta_M^{(k)}$, $k \in [1, n_1]$, n_2 tirages dans la loi $p(x', y | \theta_M)$ doivent être effectués pour estimer le quantile $F_{x_{stop}|\theta_M}^{-1}(q/100)$ puis, en considérant également la valeur correspondante $\theta_F^{(k)}$, le quantile annuel $x_{stop_T} | \theta_M, \theta_F$ et la période de retour $T_{x_{stop}} | \theta_M, \theta_F$. Il doit cependant être noté que cette méthode est très calculatoire car $n_1 \times n_2$ itérations du modèle de propagation sont nécessaires, avec des valeurs de n_1 et n_2 suffisamment grandes pour assurer que l'erreur de Monte Carlo mesurée par l'équation 3.18 est faible par rapport à la variabilité des prédictives recherchées. De façon pratique, un tel travail n'est donc possible et utile que si la quantité de données disponibles pour l'inférence est relativement faible.

3.5. Remettre en cause les hypothèses

3.5.1. Validité et sensibilité du modèle

Une fois qu'un modèle est construit, il convient de tenter autant que possible de le valider, ou tout au moins de vérifier qu'aucun dysfonctionnement flagrant ne s'observe afin de permettre son utilisation pour des cas concrets. Il s'agit donc de vérifier si les hypothèses effectuées ne sont pas trop fortes pour traduire la réalité du phénomène modélisé et permettre l'obtention de résultats opérationnels raisonnables.

Pour ce faire, les hypothèses de modélisation doivent être examinées d'un œil critique, en particulier celles concernant les structures de modèles utilisées. Les critères de comparaison présentés en section 2, c'est-à-dire le facteur de Bayes calculé selon la méthode de Kaas et Raftery (1994) et le DIC (Spiegelhalter et al., 2002), seront donc utilisés pour sélectionner au sein d'une petite classe de modèles le plus à même de représenter le phénomène étudié (chapitre 5, 6.3.2 et 6.3.4).

D'un point de vue de la validation en prédiction, les modèles proposés seront également soumis à quelques examens usuels : tests d'adéquation entre données utilisées pour l'inférence et moyenne *a posteriori* ou loi prédictive du modèle, validations croisées avec séparation des données disponibles en échantillons de calage et validation, analyses de sensibilité aux différentes variables modélisées, etc. (chapitre 5 et 6).

Enfin, d'un point de vue pratique, se pose le problème des limites de la validité des modèles développés lorsque l'information disponible est faible, c'est-à-dire lorsque la quantité de données disponibles est peu importante et/ou l'information *a priori* très diffuse. Afin d'explorer le comportement des modèles proposés dans des situations moins favorables que dans celles où l'on s'est placé pour leur construction, des investigations concernant la sensibilité à l'information *a priori* (chapitre 6.3.1) et la robustesse du calcul des aléas de référence vis-à-vis de la taille de l'échantillon d'apprentissage (chapitre 6.3.4) seront donc conduites.

3.5.2. Indépendance inter-site

Certaines hypothèses peuvent également être reconsidérées afin de permettre une amélioration du modèle. Ainsi la prédétermination locale peut être affinée par la prise en compte de davantage d'information issue de sites voisins moyennant la définition d'une méthode de transfert d'information appropriée. Dans le formalisme proposé jusqu'ici, on s'est en effet volontairement limité, mis à part pour la loi *a priori*, à l'information locale, c'est-à-dire aux avalanches enregistrées sur le site étudié. Ce faisant, on fait l'hypothèse que tous les sites ont un fonctionnement indépendant puisqu'on considère que la connaissance des paramètres sur un site n'est pas susceptible d'améliorer la prédétermination sur un autre site. Bien que rigoureuse, cette hypothèse n'est ni fondée physiquement, ni efficace opérationnellement. En effet, il est bien connu que les couloirs avalancheux fonctionnent souvent en groupes sous l'effet de forçages climatiques homogènes à l'échelle de massifs entiers : on parle alors de crues avalancheuses. De plus, l'information locale est parfois très faible sur certains sites EPA et quasi inexistante sur les sites hors EPA, de sorte que l'utilité d'un modèle de prédétermination utilisant uniquement l'information locale reste limitée.

Pour prendre en compte de façon plus juste la physique du phénomène et permettre la prédétermination sur davantage de sites avalancheux, l'hypothèse d'indépendance inter site doit donc être reconsidérée, sans toutefois, comme c'est effectué dans les méthodes norvégiennes, mélanger les observations provenant de sites différents et les considérer malgré tout comme identiquement distribuées. La modélisation hiérarchique offre un cadre parfaitement adapté permettant d'appréhender à l'échelle régionale le fonctionnement de plusieurs sites similaires. Il n'a pas été possible pour cause de temps limité de hiérarchiser au cours de la thèse l'ensemble du cadre proposé et seul le volet fréquentiel a été traité.

On considère une région composée de M sous unités spatiales et l'on modélise à présent les nombres d'avalanches observés à l'échelle régionale $(a_{11}, a_{12}, \dots, a_{jt}, \dots, a_{MT_{obs}})$, $j \in [1, M]$, $t \in [1, T_{obs}]$ au cours des T_{obs} années d'observation. Dans l'unité spatiale j , le modèle fréquentiel reste inchangé par rapport au modèle local (eq. 3.24). Par contre, tous les paramètres des modèles fréquentiels locaux décrivant l'activité avalancheuse dans les différentes unités spatiales étudiées sont supposés issus d'une même loi « régionale » dont les hyper paramètres θ_{FR} caractérisent l'activité avalancheuse à

large échelle (eq. 3.25). Les θ_{F_j} deviennent ainsi des variables latentes. En outre, la structure hiérarchique régionale permet de rendre l'estimation locale plus robuste lorsque les données sont peu nombreuses en permettant le transfert d'information des unités spatiales bien documentées aux unités spatiales peu documentées.

$$(a_{j1}, a_{j2}, \dots, a_{jt}, \dots, a_{jT_{obs}}) \sim p(\theta_{F_j}) \quad (\text{eq. 3.24})$$

$$(\theta_{F_j}) \sim p(\theta_{FR}) \quad (\text{eq. 3.25})$$

Pour représenter la structure hiérarchique régionale, il est particulièrement intéressant d'utiliser un modèle spatial utilisant la coordonnée géographique comme une variable explicative de l'observation. Ainsi le partage d'information est d'autant plus fort que les observations sont proches. Dans le cadre de la thèse, on utilisera donc au chapitre 5 deux classes de modèles spatiaux : d'une part des modèles sur grille irrégulière (ou lattice) dérivées de l'épidémiologie spatiale (Elliott et al., 2000 ; Best et al., 2004) pour la modélisation des occurrences avalanches ; d'autre part des modèles de krigeage (Diggle et Ribeiro, 2007), c'est-à-dire d'interpolation d'un champ aléatoire continu, pour l'introduction au niveau communal d'une covariable nivo-météorologique.

3.5.3. Stationnarité du phénomène

De la même manière que l'on peut remettre en cause l'indépendance spatiale des occurrences avalanches, on peut remettre en cause leur indépendance vis-à-vis du temps, ce qui revient à considérer le phénomène modélisé comme non stationnaire. Le paramètre local devient ainsi également relatif à une année (eq. 3.26). De plus, les effets du temps et de l'espace peuvent ainsi être découplés sous la forme d'un modèle d'analyse de variance. Le paramètre $\theta_{F_{jt}}$ est ainsi considéré comme la somme d'une tendance spatiale $\theta_{F_{1j}}$, d'une tendance temporelle $\theta_{F_{2t}}$, et d'un terme croisé $\theta_{F_{3j}} \times \theta_{F_{4t}}$ (eq. 3.27). Par rapport au modèle spatial, trois couches latentes supplémentaires sont modélisées. On se retrouve donc avec quatre variables latentes à hiérarchiser (eq. 3.28). Parmi elles, la tendance temporelle est particulièrement intéressante d'un point de vue phénoménologique puisqu'elle peut s'interpréter comme l'effet annuel à l'échelle régionale, c'est-à-dire le reflet du forçage climatique moyen au cours de l'hiver. Le chapitre 5 s'attachera donc, après construction d'un premier modèle spatial de base, au développement d'un tel modèle spatio-temporel permettant d'identifier la structure spatiale et temporelle des occurrences avalanches.

$$a_{jt} \sim p(\theta_{F_{jt}}) \quad (\text{eq. 3.26})$$

$$\theta_{F_{jt}} = \theta_{F_{1j}} + \theta_{F_{2t}} + \theta_{F_{3j}} \times \theta_{F_{4t}} \quad (\text{eq. 3.27})$$

$$\begin{cases} \theta_{F_{1j}} \sim p(\theta_{FR1}) \\ \theta_{F_{2t}} \sim p(\theta_{FT1}) \\ \theta_{F_{3j}} \sim p(\theta_{FR2}) \\ \theta_{F_{4t}} \sim p(\theta_{FT2}) \end{cases} \quad (\text{eq. 3.28})$$

3.6. Construire un cadre décisionnel

3.6.1. Motivations et difficultés

Jusqu'à présent, on s'est limité à la modélisation de l'aléa dans une optique de quantification des aléas de référence en termes de période de retour. Toutefois, une telle approche pose deux difficultés majeures : d'une part l'inadéquation entre des aléas de référence calculés sur des variables marginales et un phénomène multivarié déjà évoquée au chapitre 2 et qui sera illustrée par un cas concret au chapitre 7.1., d'autre part la dissociation de l'inférence et de la décision pour le choix d'aménagement discutée au chapitre 2.3.2.5.

On cherche donc à inscrire le modèle d'aléa local développé jusqu'ici dans un cadre permettant d'optimiser l'utilité d'une décision. Plus précisément, on s'intéresse au problème du dimensionnement d'un ouvrage de protection passive sous la forme d'une digue verticale dont on suppose la position dans la zone d'arrêt connue mais dont on souhaiterait optimiser la hauteur.

Comme indiqué au chapitre 2, un formalisme décisionnel est utilisé de façon assez routinière en hydraulique classique pour le dimensionnement de retenues de stockage ou de digues déviatrices de crue (Bernier, 2003). Les applications au domaine des avalanches sont au contraire quasi inexistantes. Pour dimensionner un ouvrage paravalanche à l'aide d'une approche économique, les difficultés principales sont que : (i) l'information disponible est pauvre, à la fois en termes de données locales et de rhéologie du matériau. (ii) l'influence de l'ouvrage sur l'écoulement est difficile à appréhender, en particulier si l'on essaie de modéliser son effet dans la zone d'arrêt au niveau des enjeux (iii) les modèles numériques d'avalanches les plus complexes demandent des temps de calculs trop importants pour être intégrés dans des calculs numériques de risque. Le problème (i) a déjà été traité grâce au recours à l'inférence bayésienne d'un modèle hiérarchique sur variables latentes et l'on supposera à présent connue la loi *a posteriori* $p(\theta_M, \theta_F | data)$. Le problème (ii) va être traité en intégrant dans le modèle d'aléa les effets de l'ouvrage de protection sur l'écoulement sous la forme de formules semi empiriques issues d'expérimentations sur modèles réduits et en propageant la perturbation locale dans la zone d'arrêt. Le problème (iii) va être surmonté en effectuant des hypothèses simplificatrices sur la modélisation des coûts. Celles-ci, pendant naturel des hypothèses simplificatrices effectuées sur le modèle d'aléa, permettent aux calculs de risque d'être eux aussi effectués par simulations de Monte Carlo dans un temps raisonnable, voire, si des hypothèses extrêmement fortes sont effectuées, de traiter le problème de façon entièrement analytique.

3.6.2. Notations additionnelles

On envisage la construction d'une digue verticale à l'abscisse x_d et le problème est de choisir la hauteur de digue h_d optimale qui minimise la perte moyenne attendue. Le nombre moyen d'avalanches excédant l'abscisse de la digue $E_d[a | \theta_F, \theta_M]$ est obtenu en pondérant la fréquence avalancheuse moyenne par la probabilité de dépassement (eq. 3.29). Les avalanches ayant une distance d'arrêt supérieure à l'abscisse de la digue sont alors caractérisées par un modèle POT multivarié modélisant les variables de sorties (distance d'arrêt, vitesse, hauteur, pression, etc.), ces dernières étant toujours notées génériquement y . Une digue est un ouvrage de protection passive. Elle ne modifie donc que les variables caractérisant l'avalanche dans la zone d'arrêt, c'est-à-dire la distance d'arrêt et les champs de hauteurs, vitesses et pressions au-delà de la digue. Les variables d'entrées x restent en particulier inchangées.

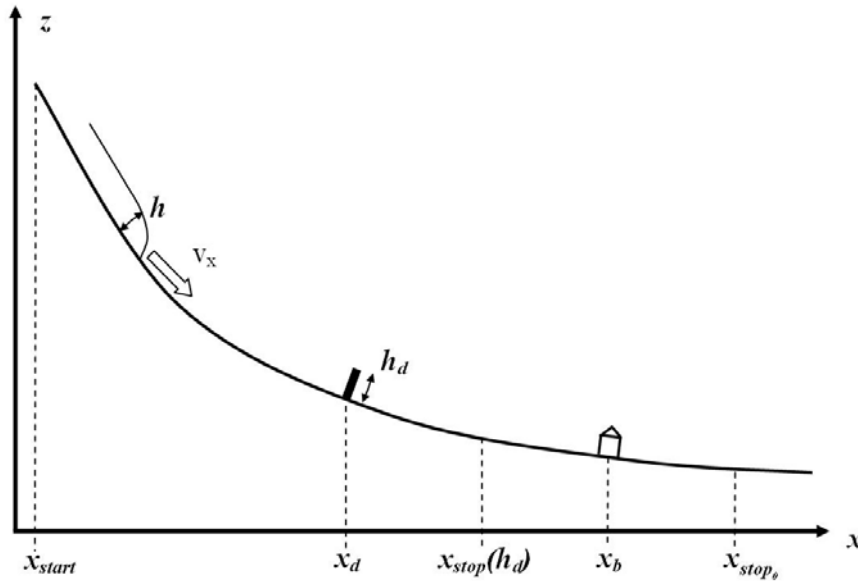
On introduit les notations $x_{stop_o} = x_{stop} | x_{stop} \geq x_d, h_d = 0$ pour la distance d'arrêt dite « de référence » d'une avalanche sans obstacle dépassant l'abscisse de la digue, $y_o = y | x_{stop} \geq x_d, h_d = 0$ pour la réalisation multidimensionnelle correspondante de l'ensemble des variables de sortie, $x_{stop}(h_d) = x_{stop} | x_{stop} \geq x_d, h_d$ pour la distance d'arrêt associée au même événement mais conditionnellement à l'existence d'une digue de hauteur h_d à l'abscisse x_d et $y(h_d) = y | x_{stop} \geq x_d, h_d$ pour la réalisation correspondante de l'ensemble des variables de sortie. On note également v_{x_o} la vitesse de l'avalanche sans digue à l'abscisse x et $v_x(h_d)$ la vitesse correspondante avec une digue de hauteur h_d .

$$E_d[a | \theta_F, \theta_M] = E[a | \theta_F] \times P(x_{stop} > x_d | \theta_M) \quad (\text{eq. 3.29})$$

Enfin, comme le problème de l'évaluation du coût n'est pas réellement traité dans le cadre de cette thèse, on se limitera à une approche économique simple où les enjeux sont restreints à un seul bâtiment vide situé à l'abscisse ponctuelle x_b de la zone d'arrêt. Si différents bâtiments sont situés approximativement à la même abscisse, le même formalisme générique peut bien entendu être utilisé. Il en va de même si le choix est fait de donner une valeur économique à la vie humaine. Certaines de

ces notations sont reportées sur la Figure 3.3 présentant une vision synthétique de l'effet d'une digue verticale sur une avalanche.

Figure 3.3 : Effet d'une digue verticale sur un écoulement avalancheux



L'avalanche se déclenche à l'abscisse x_{start} . Elle se propage avec une hauteur h supposée constante et une vitesse v_x à l'abscisse x . x_{stop_0} est la distance d'arrêt dite « de référence » d'une avalanche dépassant l'abscisse x_d de la digue sans obstacle. $x_{stop}(h_d)$ est la distance d'arrêt associée au même événement mais conditionnellement à l'existence d'une digue de hauteur h_d . Un bâtiment est situé à l'abscisse x_b de la zone d'arrêt.

3.6.3. Modéliser les effets d'un ouvrage de protection sur l'écoulement

3.6.3.1. Dissipation d'énergie locale

Dans le cadre d'un partenariat européen (projet européen CADZIE, 5^{ème} PCRD, 1999-2003), de nombreuses données ont pu être acquises sur cinq dispositifs expérimentaux mis en place par différents laboratoires (Cemagref de Grenoble, université de Bologne, université de Bristol, Office Météorologique Islandais à Reykjavik, Institut Fédéral pour l'Etude de la Neige et des Avalanches à Davos en Suisse). Elles concernent des avalanches de matériaux granulaires secs impactant un obstacle. Elles ont été analysées en termes de réduction de l'extension maximale de l'avalanche. L'ensemble des informations a permis, malgré des différences entre les caractéristiques de chaque expérience (matériau granulaire utilisé, rugosité de fond, géométrie des zones d'arrêt), d'établir une loi d'échelle linéaire valide au premier ordre pour l'ensemble des données et expliquant la réduction de la distance d'arrêt maximale par la hauteur de l'obstacle h_d normalisée par une hauteur de référence correspondant à la hauteur h de l'avalanche incidente non perturbée par l'obstacle (Faug et al., 2003 ; Faug, 2004). La Figure 3.4 montre ainsi une bonne adéquation entre le modèle linéaire et certaines des données expérimentales, la quasi-totalité des observations tombant dans l'intervalle de confiance à 95%. Un seul coefficient de proportionnalité α quantifie donc le pouvoir dissipatif d'une digue verticale (eq. 3.30). L'analyse de données obtenues sur sites réels a également montré une relation linéaire entre la distance d'arrêt avec obstacle $x_{stop}(h_d)$ et la hauteur de l'obstacle h_d (Gauer et Kristensen, 2005).

$$\frac{x_{stop}(h_d) - x_d}{x_{stop_0} - x_d} = 1 - \alpha \times \frac{h_d}{h} \quad (\text{eq. 3.30})$$

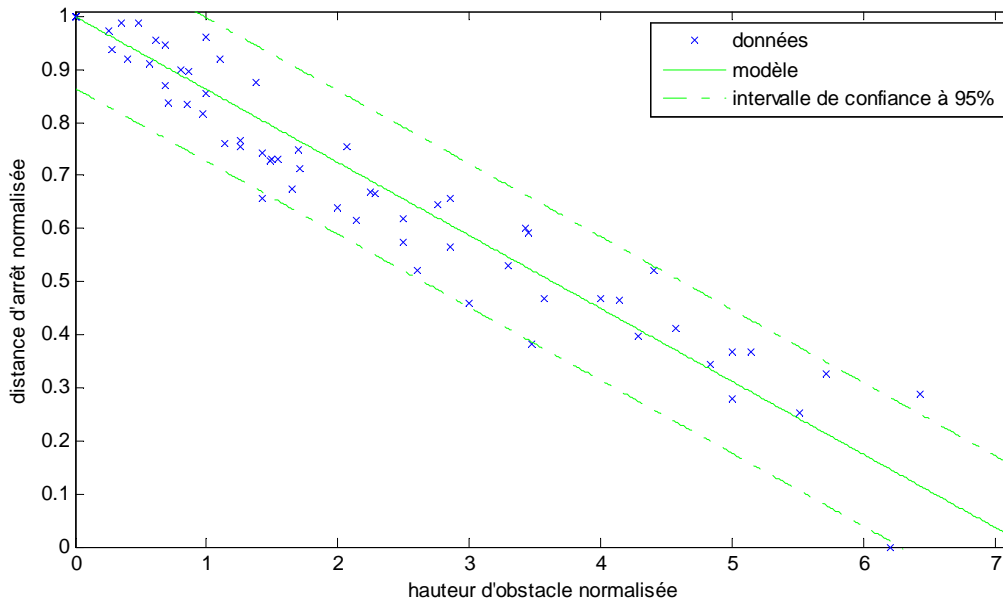
L'approche expérimentale est compatible avec une approche plus théorique : si l'effet de la digue est modélisé par une boîte noire ponctuelle, un développement au premier ordre de l'énergie dissipée par la digue permet de retrouver la relation expérimentale (Faug et al., 2004). De plus, si l'on considère

une zone d'arrêt de pente constante, la perturbation locale du profil de vitesse longitudinale peut être quantifiée (eq. 3.31).

$$\frac{v_{x_d}^2(h_d)}{v_{x_o}^2} = 1 - \frac{\alpha h_d}{2h} \quad (\text{eq. 3.31})$$

Les eqs. 3.30 et 3.31 ne sont évidemment valides que pour $h_d < \frac{h}{\alpha}$, car, pour des hauteurs de digue plus importantes, l'écoulement est totalement stoppé et $x_{stop}(h_d) - x_d = 0$. Les limites à l'utilisation des formules semi-empiriques 3.30 et 3.31 en termes de régime d'écoulement seront discutées au chapitre 7. Comme pour les calculs de pression, on supposera dans un premier temps α connu avec une précision suffisante pour négliger l'erreur d'estimation sur ce paramètre, de sorte que l'on fixera $\alpha = \hat{\alpha} = 0.1376$, c'est-à-dire égal à sa valeur expérimentale. Cette hypothèse se justifie par l'abondance et la concordance des données expérimentales, et, de façon plus pratique, par l'impossibilité de faire l'inférence sur α pour un projet d'ingénierie. Néanmoins, dans un second temps, on vérifiera la sensibilité des résultats obtenus au choix de la valeur de α , notamment en intégrant la fonction de risque par rapport à la densité de α considéré cette fois-ci comme un paramètre du modèle

Figure 3.4 : Adéquation entre loi expérimentale et données



Ajustement de la relation linéaire sur les données expérimentales issues du travail de différentes équipes européennes. $\alpha = 0.1376 \pm 0.006$

Les données brutes concernant les dispositifs expérimentaux de Davos, Bristol et Reykjavik sont en particulier disponibles dans Hakonardottir et al. (2001) et celles concernant les dispositifs de Grenoble et Bologne dans Faug et al. (2003).

3.6.3.2. Propagation de la perturbation dans la zone d'arrêt

Pour le calcul des dommages en zone d'arrêt, la dissipation d'énergie au niveau de l'obstacle doit être propagée jusqu'au niveau de l'enjeu. Avec les modèles d'aléa utilisés, propager la dissipation locale s'effectue simplement en recalculant un nouveau profil longitudinal de vitesse à partir de $v_{x_d}(h_d)$ et des caractéristiques de l'écoulement en termes de quantité de neige en mouvement et de coefficients de frottement.

De façon à limiter les temps de calcul, on choisit d'effectuer ce calcul de propagation de manière totalement analytique, ce qui au final revient à avoir une formulation analytique complète de l'effet de l'obstacle sur l'écoulement. On fait ainsi l'hypothèse que la zone d'arrêt entre la digue et le bâtiment

situé au point de coordonnées (x_b, z_b) est constituée d'un tronçon de pente constante $\tan(\phi_b) = \frac{z_d - z_b}{x_b - x_d}$. Avec un modèle de propagation de type bloc glissant et une loi de frottement de

Voellmy, une solution analytique de l'équation du mouvement est alors disponible donnant la vitesse au niveau de l'enjeu $v_{x_b}(h_d)$ en fonction de la vitesse au niveau de la digue $v_{x_d}(h_d)$, du coefficient de frottement latent μ_{voel} et de $\kappa = \frac{g}{\xi_{voel} h}$ (eq. 3.32). Bien que critiquable, cette hypothèse de pente

constante est de toute façon nécessaire pour assurer la validité de l'équation 3.31. Il convient de noter que si la distance entre la digue et l'enjeu est suffisamment grande, la vitesse limite autorisée par la loi de frottement de Voellmy est alors atteinte à l'abscisse de l'enjeu pour les avalanches caractérisées par un coefficient de frottement μ_{voel} inférieur à la pente moyenne $\tan(\phi_b)$.

$$v_{x_b}^2(h_d) = \frac{\exp[-2\kappa(x_b - x_d)\cos(\phi_b)]}{\kappa} \left(\kappa v_{x_d}^2(h_d) + g \cos(\phi_b) [\tan(\phi_b) - \mu_{voel}] [\exp\{-2\kappa(x_b - x_d)\cos(\phi_b)\} - 1] \right) \quad (\text{eq. 3.32})$$

3.6.4. Modéliser les coûts

3.6.4.1. Coûts de construction et dommages

Plutôt que d'utiliser le formalisme des analyses de risque dans le domaine des avalanches présenté au chapitre 2, on adopte ici le formalisme classique de la théorie de la décision qui est d'usage courant pour le dimensionnement d'ouvrages hydrauliques. La fonction de coût $C(h_d, y, a)$ quantifie le coût de chaque couple y, a pour chaque hauteur de la variable de décision h_d . En ingénierie hydraulique, elle s'exprime classiquement comme la somme d'un terme de construction $C_o(h_d)$ dépendant de la hauteur de la digue h_d et d'un terme quantifiant les dommages totaux infligés aux enjeux par l'aléa, soit ici les avalanches successives se produisant à partir de la date de construction (eq. 3.33). $C_1(h_d, y_{tk})$ est le coût des dommages infligé par l'avalanche $k \in [1, a_t]$ de l'année $t \in [1, +\infty[$ sachant que la hauteur de la digue est h_d et que a_t avalanches se sont produites durant l'hiver t . Ce coût doit être actualisé en supposant connu le taux d'intérêt i_t de l'année t . Notons qu'une telle fonction de coût ne représente pas explicitement le coût des dommages infligés à la digue par les avalanches successives, ceux-ci pouvant être pris en compte sous la forme d'un coût de construction actualisé ramené à la taille de l'ouvrage.

Tous les y_{tk} étant identiquement distribués sous hypothèse de stationnarité et l'intensité des avalanches étant supposée indépendante des nombres annuels d'avalanches a_t , le coût total ne dépend en fréquence que de la fréquence moyenne de dépassement $E_d[a | \theta_F, \theta_M]$. Si l'on fait en outre l'hypothèse d'un coût de construction linéaire avec la hauteur de l'obstacle, l'expression de la fonction de coût peut être considérablement simplifiée (eq. 3.34), avec une actualisation totale A_{ct} dépendant uniquement des taux d'intérêts annuels choisis (eq. 3.35).

$$C(h_d, y, a) = C_o(h_d) + \sum_{t=1}^{+\infty} \sum_{k=1}^{a_t} \frac{1}{(1+i_t)^t} \times C_1(h_d, y_{tk}) \quad (\text{eq. 3.33})$$

$$C(h_d, y, a) = C_o h_d + A_{ct} \times E_d[a | \theta_F, \theta_M] \times C_1(h_d, y) \quad (\text{eq. 3.34})$$

$$A_{ct} = A_{ct}(i_t) = \sum_{t=1}^{+\infty} \frac{1}{(1+i_t)^t} \quad (\text{eq. 3.35})$$

Pour une quantification facile de la pertinence de toute décision, un état de référence quantifiant le coût de ne rien faire est introduit, ce qui définit une fonction d'utilité linéaire $u(h_d, y, a)$ (eq. 3.36). Celle-ci quantifie la perte lorsque l'événement y, a se produit si la décision de construire une digue de

hauteur h_d a été prise plutôt que de ne rien faire. Plus simplement, l'utilité de la digue de hauteur h_d est alors simplement l'opposé du bénéfice procuré par sa construction. A noter que le choix est ici fait de ne pas pondérer les coûts par le comportement du décideur face au risque (Von Neumann et Morgenstern, 1953 ; Munier et Rousseau, 1995).

$$u(h_d, y, a) = C_o h_d + A_{ct} \times E_d \left[a | \theta_F, \theta_M \right] \times (C_1(h_d, y) - C_1(0, y)) \quad (\text{eq. 3.36})$$

3.6.4.2. Lien entre intensité des avalanches et des dommages

Comme le lien entre pression d'impact et dommages reste mal connu et que la pression elle-même est difficile à modéliser, on s'en tiendra, conformément aux approches existantes (Jonasson et al., 1999; Barbolini et al., 2004b), à une formulation empirique simple du coût dépendant de la vitesse de l'avalanche plutôt que de sa pression.

Toutefois, les relations proposées dans la littérature et présentées au chapitre 2 concernent principalement des vies humaines et quantifient une vulnérabilité plutôt qu'un coût total (i.e. leur valeur maximum est 1), de sorte qu'elles ne peuvent être utilisées directement. On se propose donc plutôt d'employer une relation simple dérivée de la pratique courante en ingénierie du zonage. Les zones d'aléa moyen et fort sont généralement délimitées par des lignes d'iso pression de 3 et $30 \text{ kN} / \text{m}^2$ dérivées de l'utilisation du modèle analytique de Voellmy proposé dans la directive suisse (Salm et al., 1990). Ces pressions correspondent, sous hypothèses d'une pression en écoulement libre calculée avec un $C_x=2$ et d'une masse volumique moyenne de la neige de $300 \text{ kg} / \text{m}^3$, à des vitesses $v_m = 3.3 \text{ m} / \text{s}$ et $v_M = 10 \text{ m} / \text{s}$ respectivement. Deux vitesses seuil v_m et v_M sont de ce fait introduites, de sorte que pour une vitesse au niveau du bâtiment $v_b < v_m$, le bâtiment n'est pas du tout endommagé tandis qu'il est entièrement détruit pour $v_b > v_M$. Entre ces deux seuils, le dommage est supposé croître linéairement avec v_b . Au final, la fonction de coût proposée dépend en intensité du phénomène des variables x_{stop}, v_b et bien sûr en fréquence du nombre moyen de dépassements $E_d \left[a | \theta_F, \theta_M \right]$. Elle se formalise aisément à l'aide de fonctions indicatrices traduisant l'appartenance à un intervalle, par exemple $I_{\{x_{stop}(h_d) \geq x_b\}} = 1$ si $x_{stop}(h_d) \geq x_b$ et $I_{\{x_{stop}(h_d) \geq x_b\}} = 0$ si $x_{stop}(h_d) < x_b$ (eq. 3.37).

$$C(h_d, y, a) = C_o h_d + A_{ct} C_1 \times E_d \left[a | \theta_F, \theta_M \right] \times I_{\{x_{stop}(h_d) \geq x_b\}} \times \left(\frac{v_b(h_d) - v_m}{v_M - v_m} I_{\{v_m \leq v_b(h_d) \leq v_M\}} + I_{\{v_b(h_d) > v_M\}} \right) \quad (\text{eq. 3.37})$$

Au chapitre 7.2, pour le développement d'un modèle analytique simplifié, une formulation plus rudimentaire encore sera utilisée. On supposera que les dommages causés au bâtiment considéré sont maximaux dès lors que celui-ci est atteint et évidemment nuls si l'avalanche s'arrête avant de l'atteindre. La fonction de coût ne dépend plus alors en intensité de l'avalanche que de la distance d'arrêt et quantifie de ce fait la perte maximale possible (eq. 3.38).

$$C(h_d, y, a) = C_o h_d + A_{ct} \times C_1 \times E_d \left[a | \theta_F, \theta_M \right] \times I_{\{x_{stop}(h_d) \geq x_b\}} \quad (\text{eq. 3.38})$$

3.6.5. Calculer le risque et optimiser l'ouvrage de protection

3.6.5.1. Principe

Comme cela a été dit au chapitre 2.3.2.5, le risque au sens classique est l'espérance de l'utilité par rapport au phénomène étudié (eq. 3.39). Il quantifie la perte moyenne attendue si la décision h_d est prise plutôt que de ne rien faire, ce qui, plus simplement, correspond à l'opposé du bénéfice moyen de la construction. Avec le modèle d'utilité choisi, seule l'intégration sur la variabilité de l'intensité des avalanches est nécessaire (eq. 3.40). Au sens de la théorie de la décision classique, la décision optimale h_c^* maximise ce bénéfice sachant les paramètres θ_F, θ_M du modèle d'aléa. La hauteur optimale classique est donc obtenue en résolvant l'équation 3.41, c'est-à-dire en assimilant les estimateurs ponctuels $\hat{\theta}_M, \hat{\theta}_F$ aux vraies valeurs des paramètres.

$$R_C(h_d, \theta_M, \theta_F) = E_{y,a} [u(h_d, y, a)] \quad (\text{eq. 3.39})$$

$$R_C(h_d, \theta_M, \theta_F) = C_o h_d + A_{ct} \times E_d [a | \theta_F] \times \int (C_1(h_d, y) - C_1(0, y)) \times l(y | \theta_M) \times dy \quad (\text{eq. 3.40})$$

$$\frac{\partial R_C(h_d, \hat{\theta}_M, \hat{\theta}_F)}{\partial h_d} = 0 \quad (\text{eq. 3.41})$$

Dans la pratique, les paramètres θ_F, θ_M du modèle d'aléa ne sont pas connus très précisément lorsque l'inférence a été effectuée avec une quantité faible de données. Le risque bayésien $R_B(h_d)$ prend en compte pour la décision l'incertitude liée à la méconnaissance locale en moyennant le risque classique sur toute la loi *a posteriori* des paramètres θ_M, θ_F (eq. 3.42). La détermination de la décision bayésienne optimale h_B^* est alors facile (eq. 3.43). En contrepartie, l'optimum bayésien peut présenter une sensibilité à la loi *a priori* $p(\theta_M, \theta_F)$.

$$R_B(h_d) = E_{\theta_M, \theta_F} [R_C(h_d, \theta_M, \theta_F)] = \int R_C(h_d, \theta_M, \theta_F) \times p(\theta_M, \theta_F | \text{data}) \times d\theta_M \times d\theta_F \quad (\text{eq. 3.42})$$

$$h_B^* = \text{Arg min}_{h_d} (R_B(h_d)) \quad (\text{eq. 3.43})$$

3.6.5.2. Calcul numérique par application des méthodes de Monte Carlo

- *Principe*

Tout comme le calcul des aléas de référence, les calculs d'optimisation peuvent être menés numériquement par application des méthodes de Monte Carlo. En pratique, il faut bien-sûr que les temps de calculs de propagation, d'influence de l'obstacle sur l'écoulement et de coût soient supportables.

Le calcul nécessite tout d'abord une discrétisation de la hauteur de digue h_d . Ensuite, les fonctions de risque classique et bayésiens sont estimées numériquement pour chaque hauteur h_d . Enfin, les hauteurs de digue qui minimisent les deux fonctions sont recherchées.

Le calcul de la fonction de risque classique est relativement simple. A partir de l'estimateur ponctuel $\hat{\theta}_M, \hat{\theta}_F$, on simule la distribution moyenne des dépassements $y_o | \hat{\theta}_M$ et l'on estime la fréquence

moyenne de dépassement $E_d [a | \hat{\theta}_F, \hat{\theta}_M] = E [a | \hat{\theta}_F] \times P(x_{stop} > x_d | \hat{\theta}_M)$. Pour chacune des

avalanches dépassant la digue, la réduction de distance d'arrêt et la dissipation d'énergie locale sont calculées pour chaque hauteur de digue et propagées le long du profil topographique. L'utilité associée à chaque couple (hauteur de digue, intensité du dépassement) est ensuite calculée. Enfin le risque classique correspondant à chaque hauteur de digue est obtenu en moyennant l'utilité associée aux différents dépassements.

Pour le risque bayésien, un effort de calcul supplémentaire est nécessaire. Pour chaque valeur $\theta_M^{(k)}$ de la loi *a posteriori* du modèle d'intensité, on simule un échantillon de variables de sortie $y | \theta_M^{(k)}$ et l'on retient les dépassements $y_o | \theta_M^{(k)}$. La fréquence de dépassement $E_d [a | \theta_F^{(k)}, \theta_M^{(k)}]$ est obtenue en pondérant $E [a | \theta_F^{(k)}]$ par la probabilité de dépassement $P(x_{stop} > x_d | \theta_M^{(k)})$. La valeur de risque correspondent au couple $(\theta_M^{(k)}, \theta_F^{(k)})$ est ensuite calculée de façon analogue au calcul du risque classique. Enfin, pour chaque hauteur de digue, le risque bayésien est obtenu en moyennant sur la loi *a posteriori*.

- *Temps de calcul*

Pour le risque classique, le temps de calcul numérique dépend de la longueur du couloir, du pas de discrétisation, du nombre d'avalanches simulées et du nombre de hauteurs de digue testées. Un nombre élevé d'avalanches est nécessaire pour éviter d'obtenir une estimation de la fonction de risque irrégulière à cause de l'erreur de Monte Carlo.

Pour la fonction de risque bayésienne, le nombre de valeurs de la loi *a posteriori* des paramètres doit également être choisi. Pour chaque valeur de paramètre, un nombre bien moins élevé d'avalanches que pour le risque classique peut être simulé du fait de l'effet lissant de la moyenne sur la loi *a posteriori*. Par contre, le nombre de valeurs de paramètres doit être suffisant pour que l'évaluation numérique de l'éq. 3.42 soit précise.

Pour les deux fonctions de risque, le nombre de hauteurs de digue doit être suffisant pour permettre une évaluation précise de la hauteur optimale. Néanmoins, un nombre trop important de hauteurs de digue augmente les temps de calcul sans présenter beaucoup d'intérêt opérationnel. Au final, le calcul numérique des fonctions de risque est donc assez lourd, mais avec les choix de modélisation effectués, notamment le calcul analytique des effets de l'obstacle sur l'aléa, les temps de calcul restent raisonnables même sur un PC standard.

3.7. Conclusion du chapitre 3

Ce chapitre s'est inspiré de la synthèse bibliographique proposée au chapitre 2 pour dégager un cadre formel permettant la prédétermination locale. L'intensité des avalanches a été modélisée indépendamment de leur fréquence. Le principe de l'inférence grâce au théorème de Bayes d'un modèle fréquentiel et d'un modèle de propagation d'avalanche quelconque sur les données d'un site EPA a été exposé. De même, en prédiction, le détail de l'ensemble des calculs allant jusqu'aux lois prédictives des aléas de référence a été exposé. Enfin, les jalons pour les extensions spatiale et décisionnelle du modèle de prédétermination locale ont été posés. L'ensemble de ces considérations génériques vont à présent être déclinées dans les quatre chapitres suivants sur des cas concrets. Pour ce faire, des hypothèses supplémentaires concernant notamment les modèles d'aléa vont être nécessaires.

4. Construction et estimation du modèle d'intensité

Préambule

Ce chapitre est consacré à la construction et à l'inférence de différents modèles d'intensité pour la prédétermination locale. A partir du cadre générique défini dans le chapitre 3, trois jeux d'hypothèses de modélisation correspondant à différents compromis entre disponibilité des données et description de la propagation de l'avalanche sont proposés. Des algorithmes d'inférence de type MCMC adaptés aux trois modèles sont mis en œuvre. Les problèmes techniques posés et l'adéquation entre modèle et données d'entrée sont discutés.

4.1. Un modèle simplifié inversible

4.1.1. Modèle stochastique

4.1.1.1. Propagation

Dans un premier temps, un modèle de propagation de type bloc glissant avec une loi de frottement de Coulomb est utilisé. Dans son article de 2005 mentionné au chapitre 2, Ancy justifie ce choix à partir d'observations de terrain (Gubler et al., 1986 ; Meunier et al., 2003) et de la théorie des écoulements granulaires (Savage et Hutter, 1991 ; Hutter et al., 1995). De tels arguments physiques confortent l'option retenue, mais le choix d'un modèle aussi rudimentaire est effectué avant tout pour des raisons de commodité, quitte à n'y voir comme Meunier et al. (2004) qu'un opérateur conceptuel reliant des entrées à des sorties.

L'équation du mouvement est extrêmement simple puisque la variation de la vitesse du bloc correspond à la différence entre la pente locale et le coefficient de frottement unique μ_{coul} (eq. 4.1). Elle s'intègre très facilement le long de la topographie entre la position initiale (x_{start}, z_{start}) et n'importe quel point (x, z) du profil topographique moyennant la spécification d'une vitesse initiale v_{zero} . L'intégration est en particulier possible sur l'ensemble du parcours de l'avalanche, donnant ainsi accès à la valeur du coefficient de frottement associé à chacune des avalanches observées dont les positions de départ et d'arrêt (x_{stop}, z_{stop}) ont été enregistrées (eq. 4.2). On peut donc considérer que, si l'on fixe v_{zero} , on dispose des réalisations des coefficients de frottement coulombiens $\mu_{coul_i}, i \in [1, N]$ associées à chacune des avalanches enregistrées. Ainsi, il n'y a pas de variable latente x'' et le coefficient de frottement devient l'une des variables observables x' .

Il doit cependant être noté que les profils de vitesse calculés à l'aide de ce modèle peuvent être totalement irréalistes, de sorte que son utilisation doit être limitée au calcul de distances d'arrêt. On pose donc $y = (x_{stop})$ pour ce modèle.

$$\frac{dv}{dt} = g \cos \phi (\tan \phi - \mu_{coul}) \quad (\text{eq. 4.1})$$

$$\mu_{coul} = \frac{z_{start} - z_{stop}}{x_{stop} - x_{start}} + \frac{v_{zero}^2}{2 \times g \times (x_{stop} - x_{start})} \quad (\text{eq. 4.2})$$

4.1.1.2. Variables d'entrée

Dans l'EPA, l'abscisse de départ est la seule variable d'entrée quantitative presque toujours disponible. C'est également la seule nécessaire au modèle de propagation de Coulomb. Pour ce premier modèle, seule l'abscisse de départ est donc utilisée comme variable d'entrée quantitative en plus du coefficient de frottement, soit $x_i = x_i' = (x_{start_i}, \mu_{coul_i}), i \in [1, N]$.

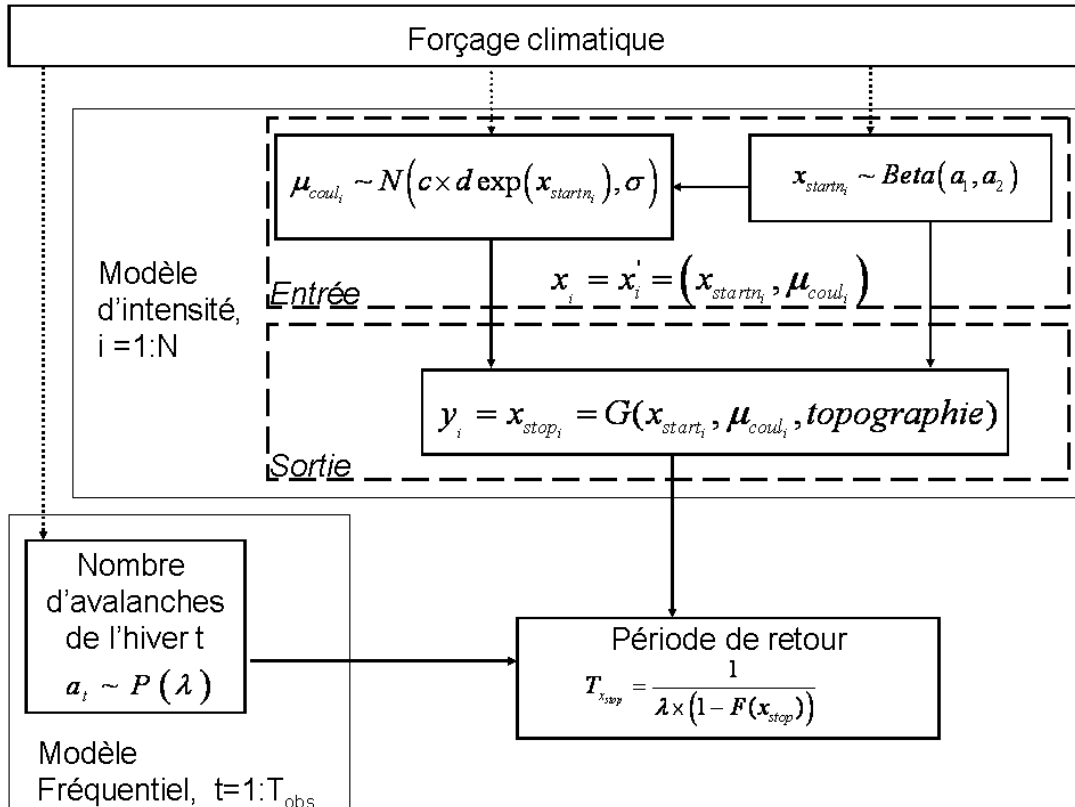
Avec seulement deux variables observables, l'écriture conditionnelle du modèle est très simple puisque $l(x_{start}, \mu_{coul}) = l(x_{start}) \times l(\mu_{coul} | x_{start})$. Cette écriture est cohérente avec la succession naturelle des phénomènes, la variable de propagation étant conditionnée par la variable de

déclenchement. Ainsi on évite soigneusement de considérer x_{start} et μ_{coul} indépendants de façon à pouvoir prendre en compte l'influence de l'altitude sur le frottement.

4.1.1.3. Modèle de vraisemblance

Les caractéristiques physiques du site avalancheux étudié permettent de déterminer des valeurs maximales et minimales x_{max} et x_{min} de l'abscisse de départ. Celles-ci seront toujours supposées connues par la suite. D'après les travaux de Meunier et al., (2001), un modèle probabiliste de type bêta est acceptable pour l'abscisse de départ normalisée notée $x_{startn} = \frac{x_{start} - x_{min}}{x_{max} - x_{min}}$. On introduit ainsi deux paramètres (a_1, a_2) à estimer (eq. 4.3).

Figure 4.1 : DAG du modèle de Coulomb



Le modèle fréquentiel est le modèle poissonien présenté au chapitre 5.

Pour le modèle d'intensité, la distance d'arrêt est la seule variable de sortie. Un seul coefficient de frottement μ_{coul} est considéré. De plus, le modèle étant inversible, le coefficient de frottement est observable. Les variables d'entrée sont limitées à l'abscisse de départ et au coefficient de frottement. Soit $x' = (x_{start}, \mu_{coul})$,

$$x'' = \{\emptyset\}, \quad y = (x_{stop}) \quad \text{et} \quad \theta_G = (\emptyset).$$

Les paramètres des modèles de fréquence et d'intensité sont $\theta_F = \lambda$, et $\theta_M = \theta_o = (a_1, a_2, c, d, \sigma)$ respectivement.

La façon dont se distribue le ou les coefficient(s) de frottement, c'est-à-dire μ_{coul} avec le modèle choisi, constitue un sujet de discussion classique pour les spécialistes des avalanches. Comme le frottement traduit à la fois la rugosité du couloir, la qualité et le volume de neige en écoulement et que les variabilités des différents effets se combinant sont supposées être d'ordres de grandeur similaires, une ou plusieurs distributions gaussiennes sont souvent utilisées dans les approches combinant un opérateur de propagation et des simulations de type Monte Carlo (Maggioni, 2005). Ancy (2005) a néanmoins montré qu'avec un modèle de Coulomb, le frottement dépend fortement du volume de l'avalanche, de sorte que toutes les valeurs d'un jeu de données ne peuvent être considérées comme

identiquement distribuées. Il propose de ce fait un modèle de mélange normal dont la moyenne décroît avec le volume observé.

On adapte le formalisme d'Ancey aux variables topographiques disponibles sous la forme d'un modèle de mélange normal dont la moyenne dépend de l'abscisse de départ selon

$E(\mu_{coul_i}) = c \times \exp(d \times x_{start_i})$ et une variance σ^2 constante (eq. 4.4). Le paramètre c quantifie la force de la corrélation entre l'abscisse de départ et le frottement coulombien tandis que le paramètre d est simplement la moyenne du coefficient de frottement en absence de corrélation. Pour $d=0$, toutes les observations sont identiquement distribuées et on se ramène à une approche statistique-dynamique standard avec un frottement basal gaussien comme dans Maggioni (2005). Le lien exponentiel choisi est dérivé des travaux d'Ancey. Il permet de garantir la positivité du terme en x_{start} de façon à pouvoir interpréter facilement les variations du paramètre c .

Pour résumer, le modèle stochastique de Coulomb se traduit par une absence de paramètre de propagation et 5 paramètres caractérisant la distribution des variables d'entrée, soit $\theta_G = (\emptyset)$ et $\theta_M = \theta_o = (a_1, a_2, c, d, \sigma)$. L'ensemble des hypothèses de modélisation est résumé en Figure 4.1. sous la forme d'un DAG. Les flèches pointillées y rappellent que le forçage climatique n'apparaît pas explicitement dans le modèle si ce n'est sous la forme d'opérateurs stochastiques purement aléatoires décrivant la variabilité du vecteur d'entrées.

$$l(x_{start_i} | a_1, a_2, x_{max}, x_{min}) = \frac{\Gamma(a_1) \times \Gamma(a_2)}{\Gamma(a_1 + a_2)} \times \left(\frac{x_{start_i} - x_{max}}{x_{max} - x_{min}} \right)^{a_1 - 1} \times \left(1 - \left(\frac{x_{start_i} - x_{max}}{x_{max} - x_{min}} \right) \right)^{a_2 - 1} \quad (\text{eq. 4.3})$$

$$l(\mu_{coul_i} | x_{start_i}, c, d, \sigma) = \frac{1}{\sigma \times \sqrt{2\pi}} \times \exp \left(-\frac{1}{2 \times \sigma^2} \times \left(\mu_{coul} - c \times \exp \left(d \times \left(\frac{x_{start_i} - x_{max}}{x_{max} - x_{min}} \right) \right) \right)^2 \right) \quad (\text{eq. 4.4})$$

4.1.2. Cas d'étude choisi

L'exemple choisi est le couloir EPA numéro 5 de la commune française de Bessans dans le département de la Savoie. La zone de départ potentielle débute par une forte cassure de pente. Elle est située sur un versant assez large à la pente moyenne et régulière. Le couloir est d'une longueur de 1200 m du point le plus haut au point le plus bas. Il est très peu chenalisé et presque rectiligne avec un dénivelé total d'environ 850 m. La pente d'abord forte se stabilise rapidement autour de 30 degrés. A partir de $x=850$ m, elle décroît rapidement et le couloir débouche dans la vallée du torrent d'Avéroles entre le hameau de la Goulaz et l'oratoire Sainte Catherine. La zone de départ potentielle est donc relativement longue et les valeurs $x_{max}=800$ m et $x_{min}=0$ ont été choisies. La route secondaire Bessans-Avéroles est directement menacée (Figure 4.2).

D'après l'EPA et les Carnets Forestiers, 69 avalanches ont été observées en 101 hivers. Cependant, seulement $N=34$ d'entre elles correspondant à la période 1960–2003 ont été retenues pour l'estimation du modèle d'intensité. Elles sont supposées donner un aperçu suffisant de l'activité avalancheuse sur le site. Le nombre maximal d'avalanches par hiver sur la période d'observation est 3 et le temps séparant deux événements consécutifs est toujours raisonnablement long, de sorte que les hypothèses d'indépendance entre intensité et fréquence ainsi que d'indépendance mutuelle des événements semblent justifiables. Davantage d'informations concernant ce cas d'étude peuvent être trouvées en Annexe A1.

Les abscisses de départ observées s'échelonnent entre $x_{start}=7.4$ m et $x_{start}=621.6$ m, avec une moyenne autour de 214 m (Tableau 4.1). La majeure partie des déclenchements se sont donc produits sur les fortes pentes en début de couloir mais que quelques déclenchements à plus basses altitudes ont également été observés. Quant aux abscisses d'arrivée, comprises entre $x_{stop}=801.5$ m et $x_{stop}=1186.8$ m avec une moyenne de 1047.1 m, elles se répartissent sur les pentes faibles au bas du couloir. A noter que 10 avalanches ont traversé la route et atteint le torrent d'Avéroles.

La transformation des données sous l'hypothèse naturelle $v_{zero}=0$ permet d'obtenir des valeurs de coefficient de frottement relativement hautes par rapport aux valeurs classiques proposées par Cassassa et al. (1989) et présentant une dispersion faible (Tableau 4.1). La corrélation empirique entre l'abscisse de départ et le coefficient de frottement est fortement négative avec un coefficient de corrélation linéaire de -0.83, ce dernier étant bien évidemment significativement non nul au seuil des

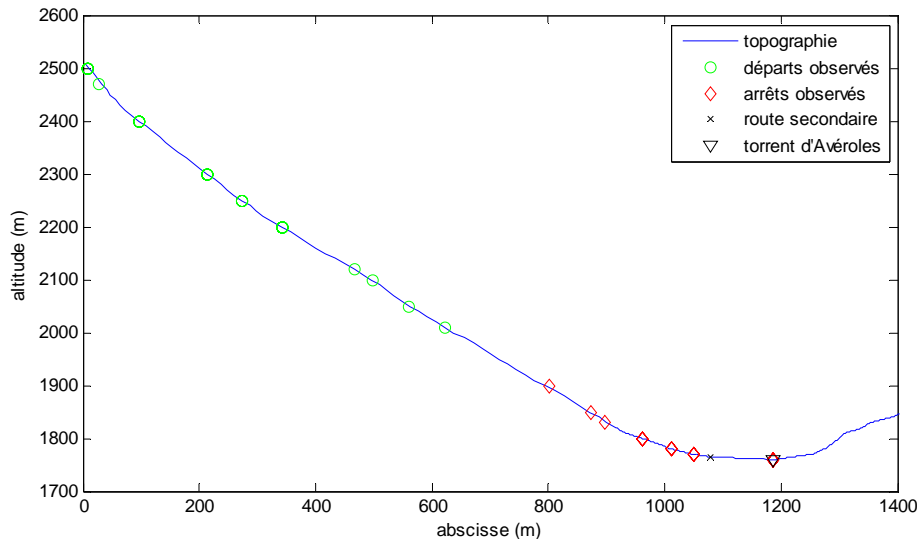
95%. Les avalanches de faible abscisse de départ sont donc caractérisées par les coefficients de frottement les plus forts. Cela s'explique si l'on considère que les fortes pentes sont localisées au sommet du couloir, de sorte que les avalanches qui partent de très haut atteignent des vitesses très élevées et ont de ce fait « besoin » d'un frottement fort pour être stoppées. Néanmoins, un tel comportement n'est pas très physique dans la mesure où l'on attendrait plutôt un frottement plus faible pour les avalanches dont la zone de départ est située très en altitude et fournit de ce fait une neige en écoulement plus froide et plus sèche. Ce paradoxe semble bien confirmer que le modèle proposé est davantage un « opérateur empirique » reliant des entrées à des sorties qu'une représentation réaliste du phénomène.

Les distributions empiriques des abscisses de départ et d'arrêt sont discrètes puisque plusieurs abscisses correspondent à différents événements du fait des arrondis dus à l'observation. De ce fait, l'abscisse de la rivière correspond à des périodes de retour empiriques allant de 3 à 30 ans (Figure 6.2) et plusieurs événements présentent des coefficients de frottement « observés » identiques.

Tableau 4.1 : Statistiques descriptives des données d'entrée pour le modèle de Coulomb

	Moy.	SD	q _{2.5}	q _{97.5}
x_{start} (m)	213.9	174.7	7.4	600.3
μ_{coul}	0.631	0.083	0.45	0.754

Figure 4.2 : Cas d'étude pour le modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

4.1.3. Construction de la loi *a priori*

La construction de la loi *a priori* a été effectuée à l'aide d'un couloir d'apprentissage. Dans la base EPA, un couloir analogue au cas d'étude choisi en termes de longueur, altitudes maximale et minimale, profil de pente, chenalisation et sinuosité a été sélectionné. Le choix a été effectué de façon experte plutôt que sur la base d'une procédure automatisée de type plus proche voisin faisant intervenir la minimisation d'une distance. L'inférence du modèle a ensuite été effectuée sur ce couloir d'apprentissage à l'aide des données locales disponibles et les *priors* marginaux peu informatifs du Tableau 4.2 (gauche). Seul le *prior* sur σ a été pris conjugué, de sorte qu'un algorithme MCMC a du être employé pour obtenir la loi *a posteriori* de $\theta_M = (a_1, a_2, c, d, \sigma)$ comme détaillé au chapitre 4.1.4.

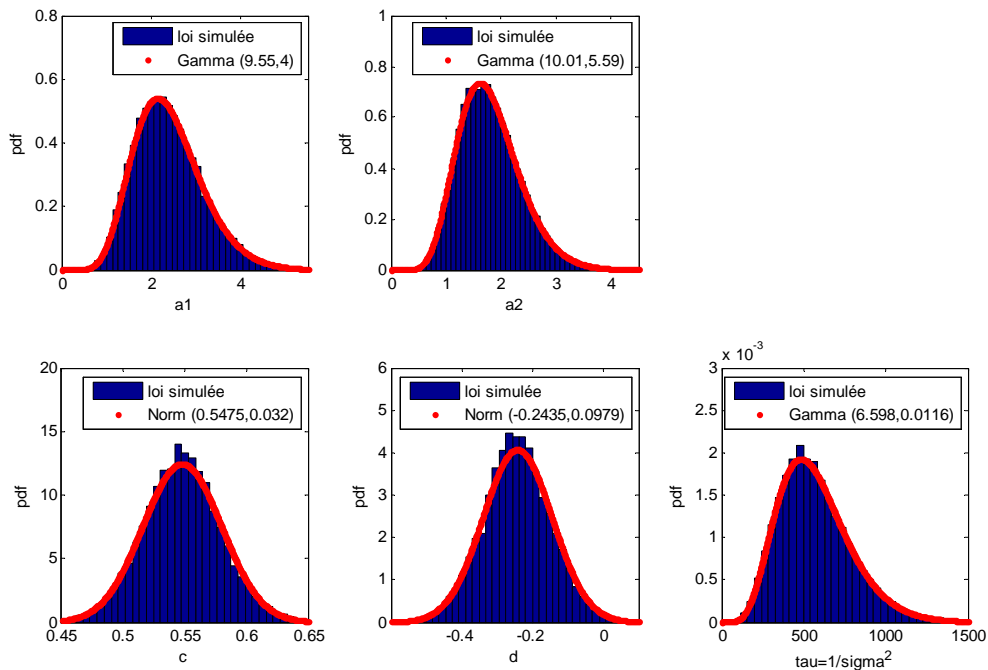
Comme les distributions marginales obtenues par simulation étaient proches de distributions standard, des modèles paramétriques simples ont pu être ajustés (Figure 4.3). Ceux-ci sont de type Gamma pour les paramètres de la distribution de l'abscisse de départ ainsi que pour la précision du modèle de mélange, et de type gaussien pour les paramètres de moyenne du coefficient de frottement. Les distributions ajustées sont synthétisées dans le Tableau 4.2 (droite). Leur produit constitue un *prior* informatif permettant le transfert d'information du couloir d'apprentissage vers le cas d'étude. Ce

faisant, conformément à l'hypothèse faite au chapitre 3, la corrélation inter paramètre *a priori* est négligée.

Tableau 4.2 : Lois a priori marginales pour le modèle de Coulomb

	Priors non informatifs pour le couloir d'apprentissage	Priors informatifs pour le cas d'étude
a_1	$a_1 \sim U(0,100)$	$a_1 \sim \text{Gamma}(9.55,4)$
a_2	$a_2 \sim U(0,100)$	$a_2 \sim \text{Gamma}(10.01,5.59)$
c	$c \sim U(0,100)$	$c \sim N(0.5475,0.032)$
d	$d \sim U(-100,100)$	$d \sim N(-0.2435,0.0979)$
σ	$\tau = \frac{1}{\sigma^2} \sim \text{Gamma}(0.01,0.001)$	$\tau = \frac{1}{\sigma^2} \sim \text{Gamma}(6.598,0.0116)$

Figure 4.3 : Construction de la loi a priori sur le couloir d'apprentissage



Le modèle de Coulomb est calibré sur le couloir d'apprentissage avec un prior non informatif. Des lois paramétriques simples sont ajustées sur les posteriors marginaux obtenus par simulation.

4.1.4. Inférence

4.1.4.1. Loi a posteriori des paramètres

Le théorème de Bayes (eq. 4.5) permet le calcul de la loi *a posteriori* jointe en combinant l'information apportée par le *prior* et les données $data_2 = (x_{start_1}, \mu_{coul_1}, \dots, x_{start_N}, \mu_{coul_N})$. La vraisemblance des données s'exprime simplement sous hypothèse d'indépendance des événements comme le produit des vraisemblances de chaque événement. En outre, la vraisemblance de chaque événement se décompose en deux termes faisant intervenir la loi conditionnelle du coefficient de frottement. Bien que certains des *priors* marginaux aient été conjugués, l'inférence a été effectuée par simulation MCMC pour tous les paramètres. Ainsi un échantillon de la loi *a posteriori* était directement disponible pour le calcul des lois prédictives des quantiles élevés de distance d'arrêt détaillé au chapitre 6.

Pour le modèle de mélange gaussien, les trois lois conditionnelles complètes sont facilement accessibles, de sorte qu'un algorithme de Gibbs a été employé. Pour les deux paramètres (a_1, a_2) , en revanche, une étape de Métropolis-Hastings a été nécessaire. 510 000 itérations ont été effectuées et

les 10 000 premières ont été laissées de côté (période de chauffe). La convergence a été vérifiée par comparaison de différentes chaînes démarrant de différents points de l'espace des paramètres (les tests effectués seront détaillés aux sections 4.2 et 4.3). Une fois le régime de croisière atteint, le nombre d'itérations ne sert qu'à améliorer la précision numérique de l'estimation au prix du temps calcul. Avec un échantillon si large, on est sûr d'évaluer même les quantiles les plus rares de la loi *a posteriori* avec une excellente précision. En outre, comme le modèle reste très simple et « purement » statistique (i.e. pas de propagation numérique du fait de l'inversibilité du modèle de Coulomb), le calcul n'a pris qu'environ 10 minutes avec un PC standard récent.

Les lois *a posteriori* marginales obtenues sont présentées en Figure 4.4. Comme on pouvait s'y attendre, elles ont toute une variance plus faible que les lois *a priori*. Cela indique que l'information *a priori* a bien été complétée par l'information apportée par les données. De plus, tous les *posteriors* obtenus présentent un mode bien marqué, ce qui facilite la détermination d'un estimateur ponctuel raisonnable. On constate également que pour trois paramètres, i.e. (a_2, d, σ) , le *posterior* est relativement proche du *prior*. En revanche, une forte divergence est observée entre les *priors* et *posteriors* marginaux des deux autres paramètres. Ceci indique que le *prior* choisi n'était pas forcément tout à fait adapté, c'est-à-dire que, même si le couloir d'apprentissage était "proche" du cas d'étude, il en différait par certains aspects. La sensibilité au *prior* sera de ce fait discutée au chapitre 6.

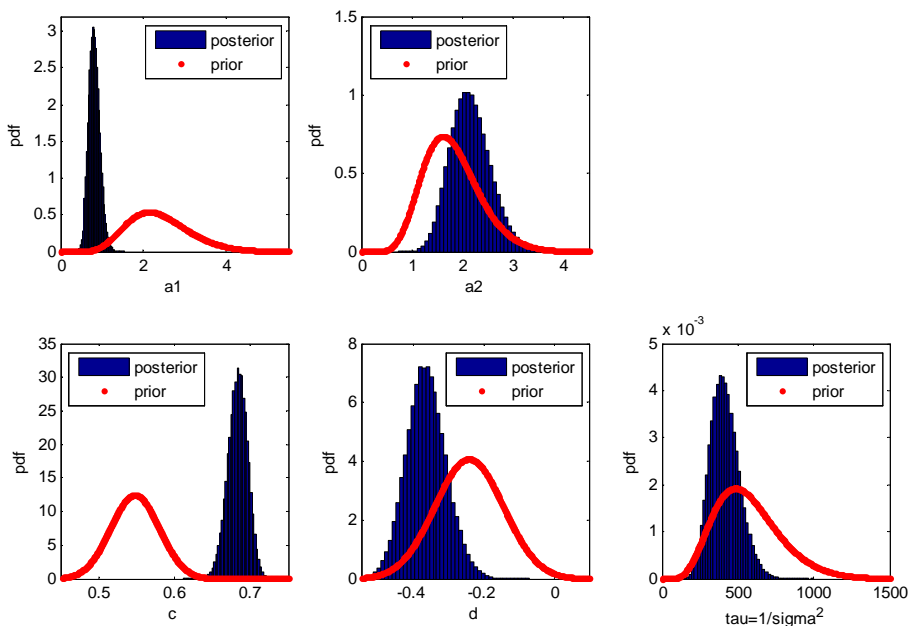
$$p(a_1, a_2, c, d, \sigma | data_2) \propto \prod_{i=1}^N \left(l(x_{start_i} | a_1, a_2) \times l(\mu_{coul_i} | x_{start_i}, c, d, \sigma) \right) \times p(a_1, a_2, c, d, \sigma) \text{ (eq 4.5)}$$

Les principales caractéristiques des lois *a posteriori* marginales obtenues sont résumées dans le Tableau 4.3. Il apparaît que $a_1 < a_2$, ce qui indique que la distribution des abscisses de départ présente une forte asymétrie à droite. De plus, d est fortement négatif, ce qui traduit la corrélation négative importante observée dans les données entre l'abscisse de départ et le coefficient de frottement coulombien.

Tableau 4.3 : Statistiques descriptives des lois *a posteriori* marginales pour le modèle de Coulomb

	Moy.	SD	q _{2.5}	q _{97.5}
a_1	0.82	0.14	0.57	1.1
a_2	2.15	0.4	1.44	3.01
c	0.683	0.013	0.656	0.708
d	-0.36	0.06	-0.47	-0.25
σ	0.05	0.006	0.04	0.06

Figure 4.4 : Prior et posterior pour le modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

4.1.4.2. Moyenne a posteriori des variables d'entrée

La moyenne *a posteriori* des variables d'entrée du modèle $p\left(x_{start}, \mu_{coul} \mid \hat{\theta}_M\right) = p\left(x_{start} \mid \hat{a}_1, \hat{a}_2\right) \times p\left(\mu_{coul} \mid x_{start}, \hat{c}, \hat{d}, \hat{\sigma}\right)$ quantifie la variabilité du couple

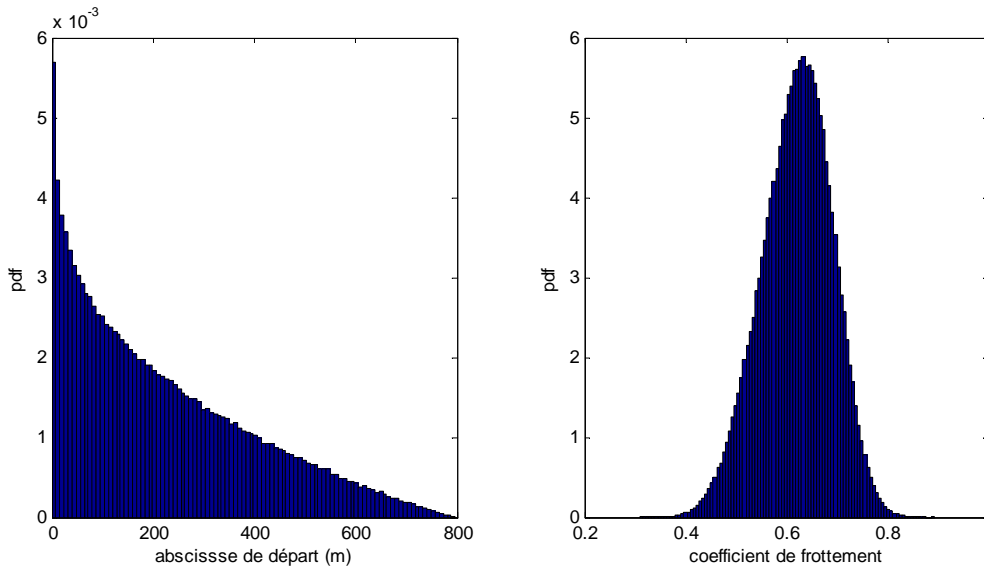
x_{start}, μ_{coul} sachant $\theta_M = \hat{\theta}_M$. La loi marginale de l'abscisse de départ s'obtient simplement par plug-in de l'estimateur ponctuel. Pour le coefficient de frottement, une intégration numérique sur la distribution de l'abscisse de départ est nécessaire. Les deux lois marginales sont résumées dans le Tableau 4.4 La moyenne *a posteriori* des distances d'arrêt nécessite quant à elle la mise en œuvre de calculs de propagation qui seront discutés au chapitre 6.

Tableau 4.4 : Distributions moyennes a posteriori des variables du modèle de Coulomb

	Moy.	SD	q _{2.5}	q _{97.5}
μ_{coul}	0.62	0.07	0.476	0.748
x_{start} (m)	220.3	179.3	3.9	636
x_{stop} (m)	1027.6	205.1	401.3	1240.2

La Figure 4.5 confirme la forte asymétrie de la distribution des abscisses de départ. Cela indique que la majorité des déclenchements se produisent sur les pentes sommitales du couloir. La distribution correspondante du coefficient de frottement présente globalement une forme en cloche mais également une légère asymétrie à gauche. Cette dernière caractéristique souligne l'intérêt du modèle proposé. Les valeurs de coefficient de frottement les plus faibles correspondent en effet généralement aux avalanches présentant les distances d'arrêt les plus longues. Il est de ce fait très important de choisir un modèle prenant en compte les valeurs les plus extrêmes de μ_{coul} , ce que permet un modèle de mélange mais pas un modèle gaussien symétrique.

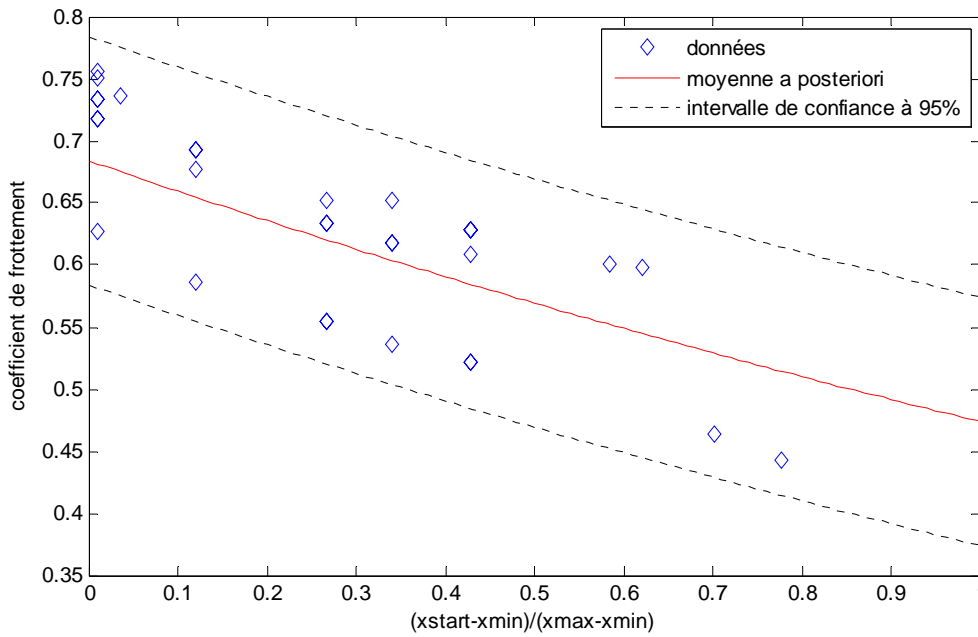
Figure 4.5 : Distributions moyennes a posteriori des variables d'entrée du modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

La Figure 4.6 propose une comparaison graphique entre données et modèle. Empiriquement, le coefficient de frottement décroît bien avec l'abscisse de départ tandis que la dispersion de la distribution ne semble pas dépendre de l'abscisse de départ. Le modèle proposé capture bien ces effets avec un paramètre de corrélation négatif et une variance du modèle de mélange constante, de sorte que toutes les observations tombent dans l'intervalle de confiance à 95% du modèle. Celui-ci est construit en considérant la moyenne *a posteriori* ± 1.96 écart types. En outre, le modèle décroît quasi linéairement avec l'abscisse de départ. Cela suggère qu'un lien linéaire entre coefficient de frottement et abscisse de départ pourrait tout aussi bien fonctionner que le lien exponentiel proposé dont le choix a été effectué plus pour des raisons de commodité qu'à cause d'arguments physiques.

Figure 4.6 : Adéquation entre données et moyenne a posteriori du modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

4.1.5. Prise en compte de l'irrégularité topographique

4.1.5.1. Un second cas d'étude

On s'intéresse maintenant au couloir EPA numéro 13 de la commune de Bessans avec $N=41$ événements retenus sur la même période que pour le cas d'étude précédent, à savoir 1960–2003. Par rapport à ce dernier, ce couloir se caractérise par sa taille plus importante : 2300 m de long pour un dénivelé total de 1500 m. La chenalisation est réduite et le couloir se termine dans la vallée de l'Arc, à cet endroit large et plate. En termes d'enjeux, seule la chapelle Saint Sébastien est située dans l'axe du couloir, mais suffisamment loin pour sembler hors d'atteinte même d'avalanches exceptionnelles (Figure 4.7).

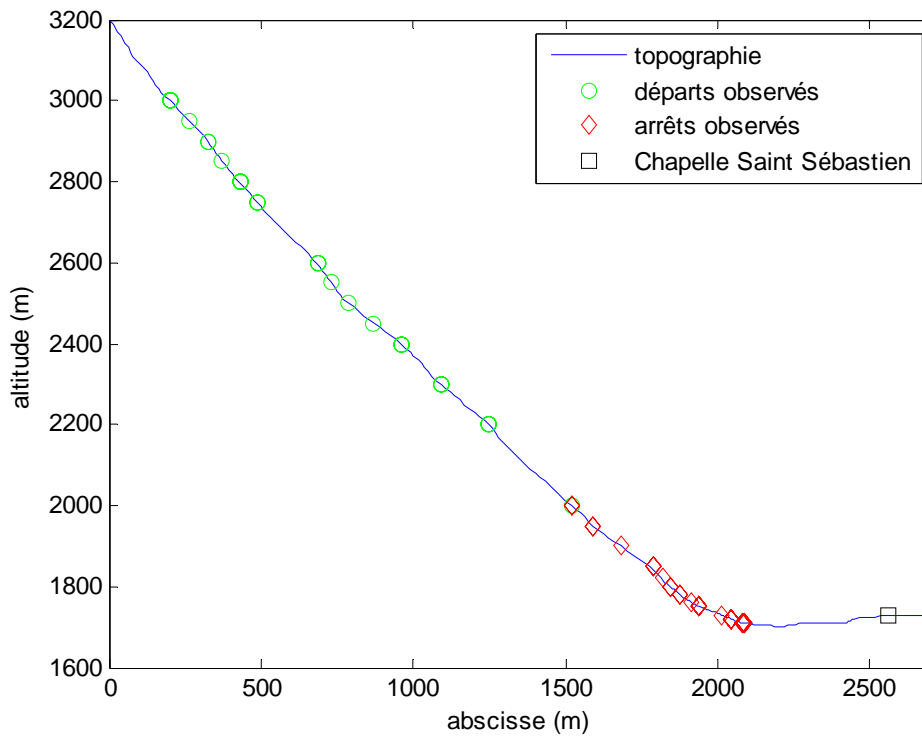
La pente moyenne est forte (34.2°) et reste presque toujours supérieure à 30° pendant les 1900 premiers mètres du couloir, de sorte que l'on fixe $x_{max}=1900$ m et $x_{min}=0$. La surface de départ potentielle est donc très longue. Une analyse cartographique permet toutefois d'établir que son extension latérale est réduite. Par rapport au cas d'étude précédent, une différence notable est la moindre régularité du profil de pente. En effet, plusieurs changements de concavité modérés mais significatifs existent, notamment deux zones de pente plus faibles dans la zone de départ et au contraire un léger ressaut en zone d'arrêt (Figure 4.8).

Les abscisses de départ observées présentent une forte dispersion, de $x_{start}=203$ m à $x_{start}=1519$ m avec une moyenne de 592 m et un écart type de 373.3 m (Tableau 4.5). Les abscisses d'arrêt observées varient quant à elles de $x_{stop}=1519$ m à $x_{stop}=2090$ m avec une moyenne de 1909.3 m et un écart-type plus réduit de 159 m (Tableau 4.5). La plupart des avalanches s'arrêtent donc sur les pentes modérées précédant l'arrivée dans la vallée de l'Arc entre $x=1900$ m et $x=2090$ m. Cependant, quelques avalanches s'arrêtent plus tôt, c'est-à-dire sur la zone de pente moyenne précédant le dernier ressaut entre $x=1519$ m et $x=1850$ m. On peut même remarquer que le déclenchement le plus bas correspond à la distance d'arrêt la plus faible. Davantage d'informations concernant ce cas d'étude peuvent être trouvées en Annexe A2.

Tableau 4.5 : Données d'entrée pour l'application du modèle de Coulomb bruité

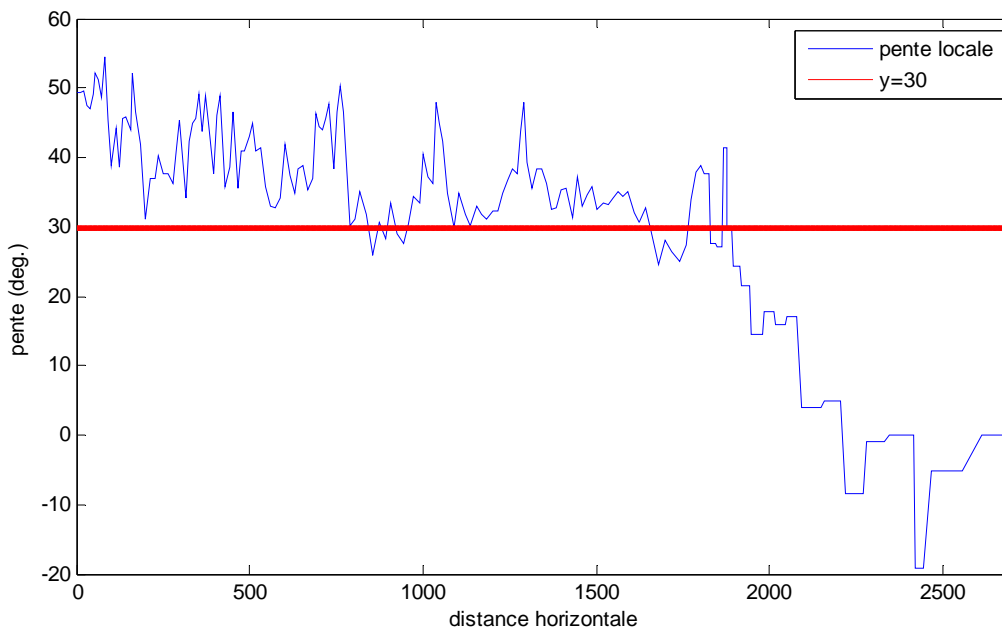
	Moy.	SD	q _{2.5}	q _{97.5}
x_{start} (m)	592	373.3	203.3	1382.6
μ_{coul}	0.71	0.031	0.641	0.784
ε (m)	11.2	50.3	-137.6	104.6

Figure 4.7 : Cas d'étude pour le modèle de Coulomb bruité



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Figure 4.8 : Profil de pente du cas d'étude pour le modèle de Coulomb bruité



Couloir EPA 13 de la commune de Bessans, Savoie, France.

4.1.5.2. Introduction d'une erreur de propagation et d'une vitesse initiale dans le modèle

Le modèle de propagation de Coulomb peut être utilisé pour recalculer numériquement sachant v_{zero} les distances d'arrêt correspondant à chacune des observations selon $x_{stop_{cal}} = G(x_{start}, \mu_{coul}, topographie, v_{zero})$ en définissant l'abscisse d'arrêt comme le point du profil où la vitesse devient nulle. Comme les erreurs d'observation ne sont pas prises en compte et que l'inversion de l'équation du mouvement donnant accès au coefficient de frottement est bijective, les

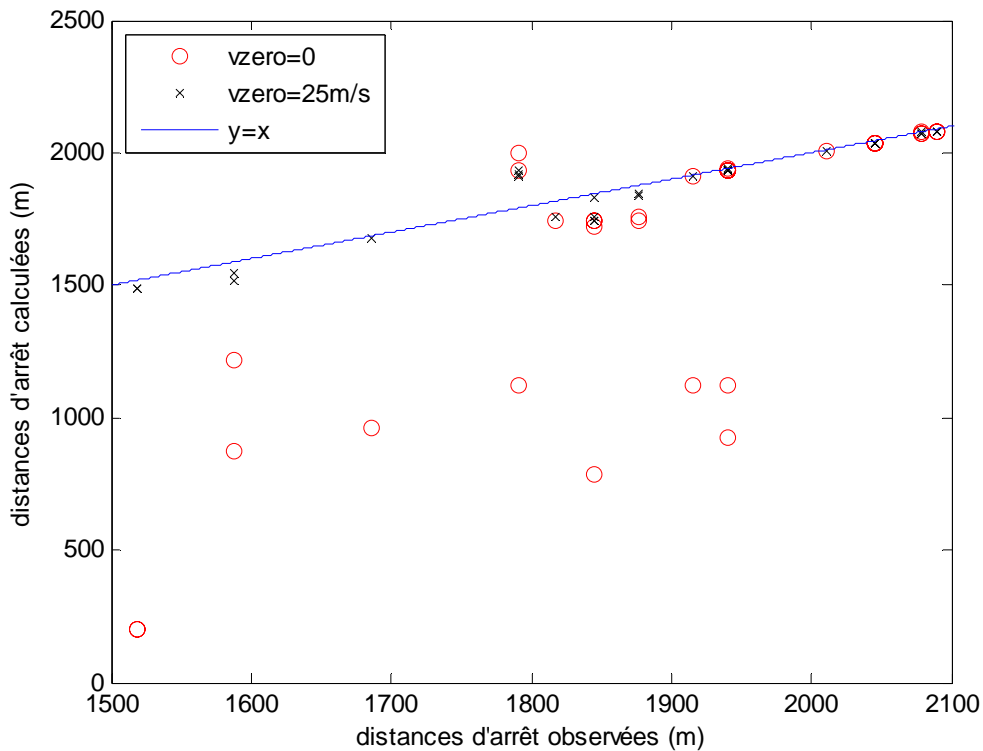
différences entre distances d'arrêt observées et recalculées devraient toujours être nulles aux éventuelles erreurs de discrétisation numérique près. Néanmoins cela n'est pas toujours le cas et des différences très importantes sont parfois observées. On introduit ainsi l'erreur de propagation ε (eq. 4.6).

$$\varepsilon = x_{stop} - x_{stop,cal} \text{ (eq. 4.6)}$$

Ce point n'a pas du tout été discuté pour le premier cas d'étude car sa topographie était suffisamment régulière pour que, pour tous les événements enregistrés, l'erreur de propagation ε soit négligeable. C'est par contre loin d'être le cas pour le couloir 13 de Bessans pour lequel le modèle de Coulomb précédemment décrit est inefficace. La Figure 4.9 montre en effet qu'avec une vitesse initiale nulle, environ un tiers des avalanches observées présentent des distances d'arrêt recalculées extrêmement courtes. Cela indique qu'elles sont arrêtées tout de suite après leur départ, voire qu'elles ne partent pas du tout lorsque la pente au niveau de l'abscisse de départ est trop faible pour permettre l'accélération. Simuler des distances d'arrêt avec une vitesse initiale nulle et ne pas tenir compte de l'erreur de propagation conduit donc à très mal appréhender le fonctionnement avalancheux de ce couloir du fait de sa relative complexité topographique.

Une façon simple d'y remédier est de travailler avec une vitesse initiale fixée différente de 0 permettant de recalculer de façon assez satisfaisante toutes les avalanches observées. Ainsi, pour le cas d'étude choisi, une vitesse initiale de 25 m/s est suffisante pour s'approcher de toutes les distances d'arrêt observées (Figure 4.9).

Figure 4.9 : Distances d'arrêt observées et calculées pour différentes vitesses initiales



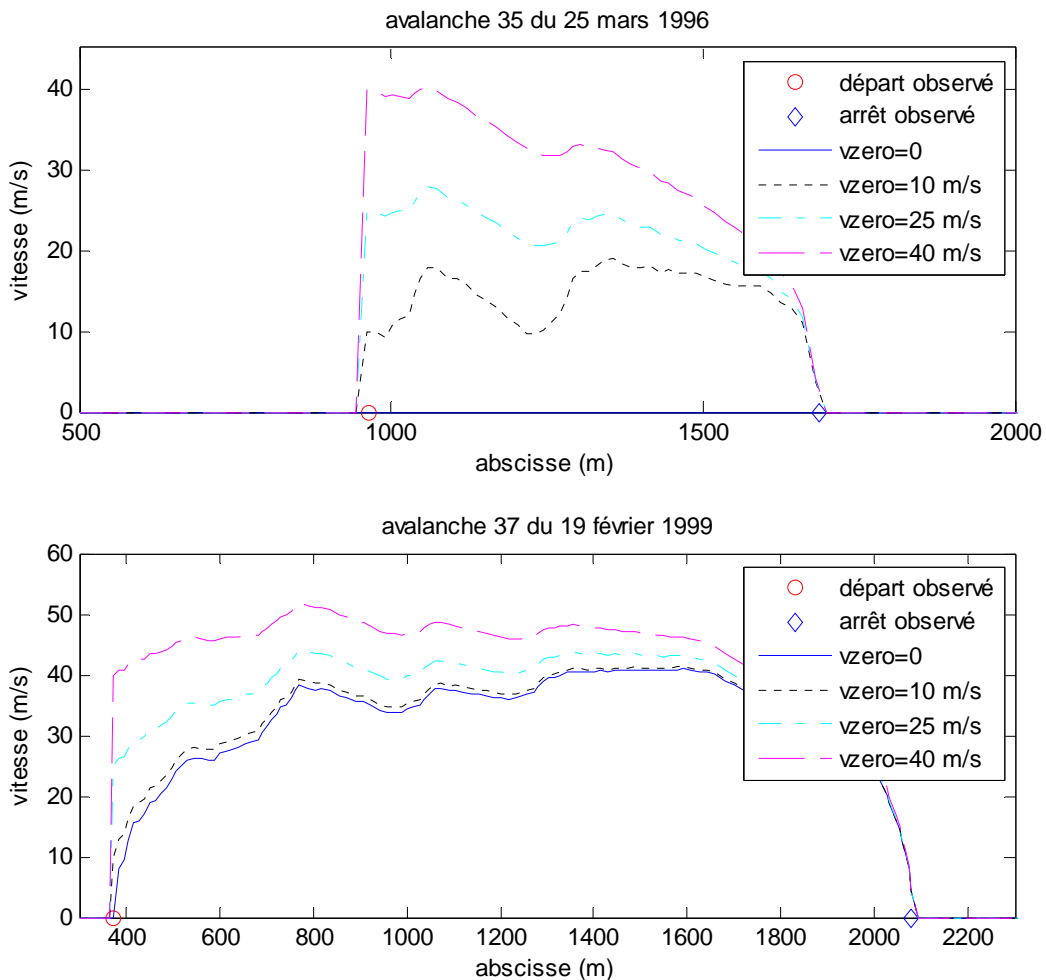
Couloir EPA 13 de la commune de Bessans, Savoie, France. Distances d'arrêt calculées avec le modèle de propagation de Coulomb.

Bien sûr, l'introduction d'une vitesse initiale non nulle est discutable. Des arguments physiques peuvent toutefois être apportés. En effet, après la fracture, la neige mise en mouvement se concentre au sommet du couloir, de sorte que la formation d'un bloc avalancheux tel qu'il est modélisé et si tant est qu'il existe n'est pas instantanée. De plus, des mesures de terrain montrent qu'au cours des premières secondes suivant le déclenchement, l'avalanche accélère extrêmement rapidement (Ancey et Meunier, 2004). Imposer une vitesse initiale au bloc censé modéliser l'avalanche n'est donc pas une idée totalement dépourvue de sens.

L'argument principal en faveur de la méthode proposée est cependant d'ordre pratique. Pour chacune des avalanches enregistrées, il existe une vitesse seuil dépendant de la pente locale au niveau de l'abscisse de départ. Pour des vitesses initiales inférieures à ce seuil, l'avalanche recalculée ne démarre pas ou alors s'arrête immédiatement. Pour des vitesses plus élevées que la vitesse seuil, l'avalanche est correctement recalculée à une erreur de propagation raisonnable près. Choisir des vitesses initiales différentes supérieures au seuil modifie le coefficient de frottement et le profil de vitesse mais pas l'abscisse d'arrêt comme illustré en Figure 4.10. Avec une vitesse initiale nulle, l'avalanche du 25 mars 96 ne démarre pas, tandis qu'avec des vitesses initiales de 10 à 40 m/s, voire supérieures, la même distance d'arrêt est recalculée. Pour l'avalanche du 19 février 99, la pente au niveau de l'abscisse de départ est suffisamment forte pour que, même avec une vitesse initiale nulle, l'avalanche puisse être recalculée de façon satisfaisante. La différence entre distance d'arrêt observée et recalculée, c'est-à-dire l'erreur de propagation, est pour ces deux avalanches de l'ordre de quelques mètres seulement.

Par contre, dès lors qu'une vitesse initiale est suffisante pour faire démarrer toutes les avalanches, faire varier la vitesse initiale ne permet pas de minimiser les erreurs de propagation. Il ne semble de ce fait ni nécessaire ni même possible de considérer la vitesse initiale comme un paramètre aléatoire supplémentaire à calibrer pour prendre en compte la variabilité des distances d'arrêt. On propose donc de choisir la vitesse initiale minimale qui permet de recalculer tous les événements observés sur le site étudié, soit 25 m/s pour le couloir 13 de Bessans, et d'utiliser cette valeur fixée pour l'ensemble des calculs prédictifs.

Figure 4.10 : Effet de la vitesse initiale sur les profils de vitesses avec le modèle de Coulomb

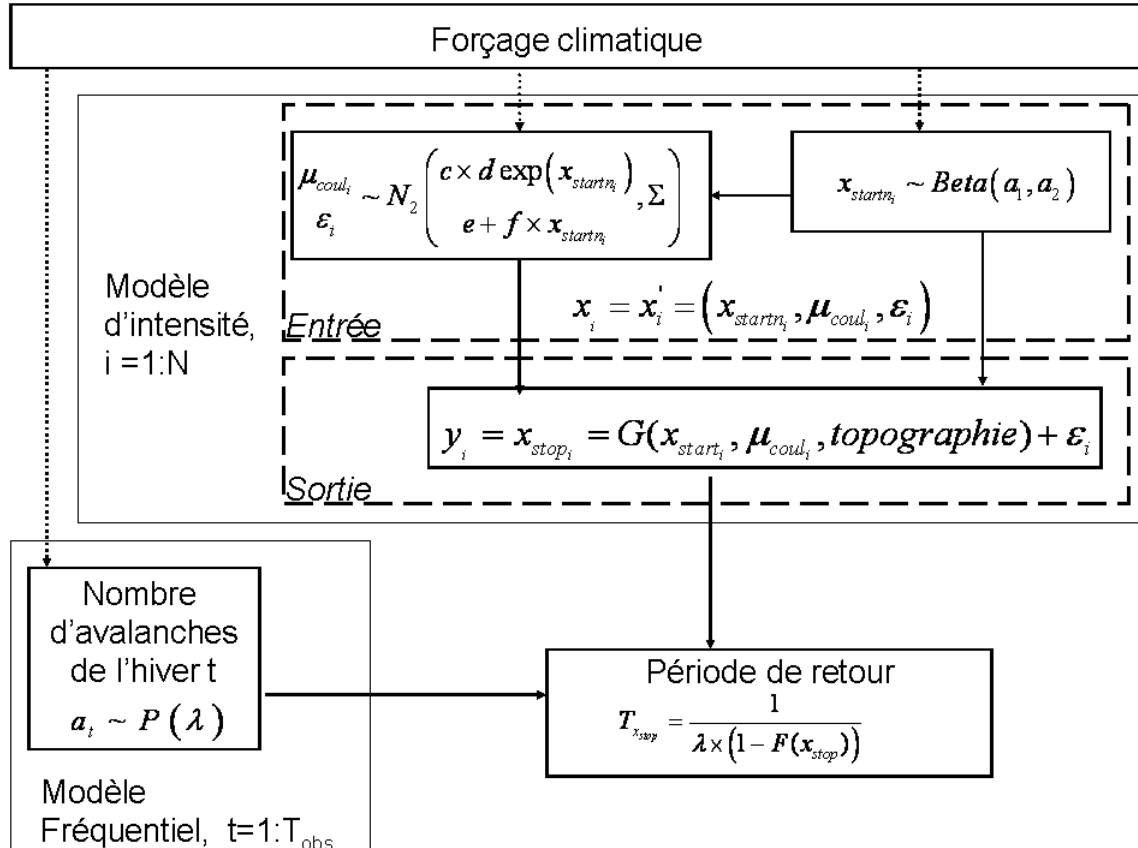


Couloir EPA 13 de la commune de Bessans, Savoie, France. Profils de vitesse calculés avec le modèle de propagation de Coulomb.

4.1.5.3. Adaptation du modèle stochastique

Le modèle de Coulomb substitue des couples d'observables (x_{start}, μ_{coul}) aux « vraies » observations (x_{start}, x_{stop}) . Si l'on considère en plus la distance d'arrêt recalculée $x_{stop,cal}$ que l'on transforme en une erreur de propagation ε , on se ramène à un triplet de variables d'entrée $x = x' = (x_{start}, \mu, \varepsilon)$ dont on dispose des réalisations $data_{2_i} = (x_{start_i}, \mu_i, \varepsilon_i), i \in [1, N]$. Les profils de vitesses calculés avec une vitesse initiale non nulle étant au moins aussi questionnables que ceux obtenus avec une vitesse initiale nulle, on limitera bien sûr encore une fois les sorties opérationnelles à la distance d'arrêt, soit $y = (x_{stop})$.

Figure 4.11 : DAG du modèle de Coulomb bruité



Le modèle fréquentiel est le modèle poissonien présenté au chapitre 5.

Pour le modèle d'intensité, la variable de sortie est la distance d'arrêt uniquement. Un seul coefficient de frottement μ_{coul} est considéré. De plus, le modèle étant inversible, le coefficient de frottement est observable.

Une erreur de propagation ε représente les différences entre distances d'arrêt calculées et observées. Les autres variables d'entrée sont limitées à l'abscisse de départ et au coefficient de frottement. Soit

$x' = (x_{start}, \mu, \varepsilon)$, $x'' = \{\emptyset\}$, $y = x_{stop}$ et $\theta_G = \{\emptyset\}$. Les paramètres des modèles de fréquence et

d'intensité sont $\theta_F = \lambda$, et $\theta_M = \theta_o = (a_1, a_2, c, d, e, f, \Sigma)$ respectivement.

Le modèle de Coulomb précédemment proposé doit donc être adapté au triplet. De plus, on désire conserver la possibilité de prendre en compte toutes les corrélations intervariables. Un modèle de mélange gaussien bivarié est donc proposé pour le couple μ_{coul}, ε avec une moyenne dépendant de l'abscisse de départ normalisée x_{start} et une matrice de variance-covariance Σ constante (eq. 4.7). Pour la moyenne de la loi marginale de ε , un lien linéaire est postulé, avec deux paramètres f et e à estimer quantifiant respectivement la corrélation linéaire avec x_{start} et la moyenne de la loi marginale de ε en l'absence de corrélation. Pour μ_{coul} , le même lien exponentiel avec x_{start} que pour le modèle

non bruité est conservé de façon à ce que les deux modèles soient identiques lorsque l'erreur de propagation tend vers zéro. Σ_{11} représente la variance de la loi marginale du coefficient de frottement, Σ_{22} la variance de la loi marginale de l'erreur de propagation et Σ_{12} le terme de covariance quantifiant la corrélation entre coefficient de frottement et erreur de propagation. Enfin, le modèle marginal pour l'abscisse de départ x_{start} est bien évidemment conservé, de sorte que, pour ce modèle, $\theta_M = \theta_o = (a_1, a_2, c, d, e, f, \Sigma)$, soit 9 paramètres au total à inférer. Bien entendu on a toujours $\theta_G = (\emptyset)$. On qualifiera ce modèle de « modèle de Coulomb bruité » par opposition au modèle de Coulomb non bruité précédemment décrit. L'ensemble des notations est résumé en Figure 4.11 sous la forme d'un DAG.

$$\left(\mu_{coul_i}, \varepsilon_i \mid x_{start_i}, c, d, e, f, \Sigma \right) \sim N_2 \left(\begin{pmatrix} c \times \exp(d \times x_{start_i}) \\ e + f \times x_{start_i} \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix} \right) \quad (\text{eq. 4.7})$$

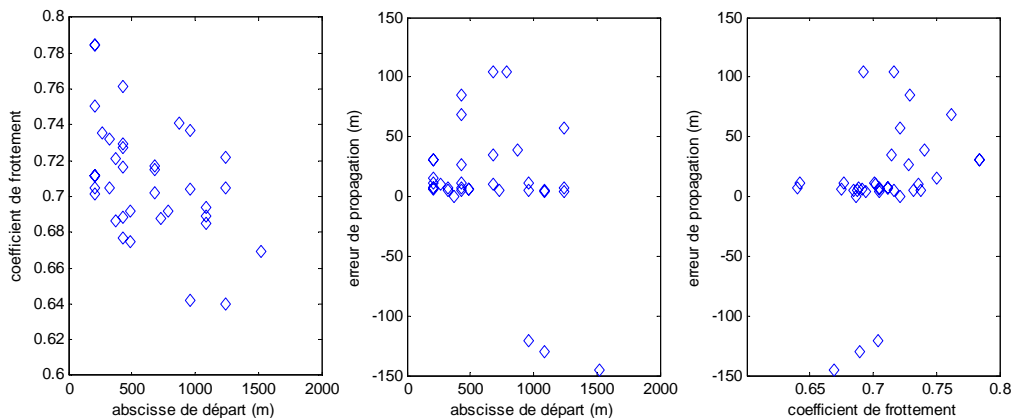
La vraisemblance des données s'exprime, comme pour le modèle de Coulomb non bruité, comme le produit pour chaque événement de la probabilité de la variable de déclenchement x_{start} et de la probabilité conditionnelle du couple $(\mu_{coul}, \varepsilon)$ représentant la propagation (eq. 4.8).

$$l(data_2 \mid \theta_M) = \prod_{i=1}^N \left(l(x_{start_i} \mid a_1, a_2) \times l(\mu_{coul_i}, \varepsilon_i \mid x_{start_i}, c, d, e, f, \Sigma) \right) \quad (\text{eq. 4.8})$$

4.1.5.4. Application au cas d'étude

Les valeurs du triplet d'entrées obtenues pour le cas d'étude avec $v_{zero} = 25$ m/s sont représentées en Figure 4.12 sous la forme de scatter plots. Leurs distributions sont résumées dans le Tableau 4.5. Les valeurs de coefficient de frottement sont très élevées par rapport aux valeurs standard proposées par Casassa et al. (1989) ainsi que par rapport aux valeurs obtenues pour le couloir 5 de Bessans. Cela est dû à la pente moyenne forte du couloir mais aussi au frottement additionnel qui compense la vitesse initiale dans l'équation 4.2. La distribution des coefficients de frottement est également caractérisée par sa faible dispersion : toutes les valeurs sont comprises entre 0.64 et 0.784 avec un écart-type de seulement 0.03, soit deux fois moins que pour le couloir numéro 5 de Bessans. Cela indique que, pour ce couloir, les parcours empruntés par les différents événements ne sont pas très différents les uns des autres.

Figure 4.12 : Scatter plot des données d'entrée pour le modèle de Coulomb avec terme d'erreur



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Les erreurs de propagation sont en moyenne positives, ce qui indique que le modèle de propagation sous-estime systématiquement les distances d'arrêt. L'écart type empirique est assez élevé (50.3 m) : si beaucoup d'avalanches sont bien recalculées par le modèle de propagation de Coulomb, quelques unes d'entre elles le sont donc néanmoins assez mal avec des erreurs de propagation de l'ordre de ± 100 m. Ces erreurs importantes correspondent aux avalanches dont les abscisses d'arrêt observées se trouvent au niveau du ressaut en zone d'arrêt. Elles indiquent que le modèle de propagation très

simplifié de Coulomb ne permet pas, malgré l'introduction d'une vitesse initiale, de bien décrire une topographie complexe. Elles justifient de ce fait l'introduction explicite d'une erreur de propagation dans le modèle stochastique.

Enfin, même si la corrélation empirique entre les données d'entrées n'est pas très marquée sur les « scatter-plots » de la Figure 4.12, les coefficients de corrélations linéaires sont tous trois supérieurs à 0.3 en valeur absolue (Table 4.6, gauche) et significativement non nuls au seuil des 95%. Comme pour le cas d'étude du modèle de Coulomb non bruité, le coefficient de frottement et l'abscisse de départ sont fortement négativement corrélés, avec une explication vraisemblablement similaire. L'erreur de propagation est, quant à elle, corrélée de façon moins forte avec les deux autres variables. La corrélation négative avec l'abscisse de départ est assez naturelle dans la mesure où l'on peut penser qu'en moyenne plus la distance parcourue par l'avalanche est faible et plus le modèle de propagation devrait la reproduire efficacement. La corrélation positive avec le coefficient de frottement est plus difficile à expliquer. Elle est cohérente avec les deux autres corrélations négatives mais il ne s'agit pas uniquement d'un artefact de corrélation croisée puisque les corrélations partielles, bien que moins fortes, restent toutes les trois non nulles. (Tableau 4.6, droite).

Tableau 4.6 : Corrélations totales et partielles entre variables du modèle de Coulomb bruité

	μ		ϵ	
x_{start}	-0.46	-0.39	-0.34	-0.21
μ_{coul}	1	1	0.34	0.22

4.1.5.5. Inférence

Pour ce cas d'étude, la quantité de données a été supposée suffisante pour ne pas recourir à l'information *a priori*, de sorte que les distributions marginales très peu informatives du Tableau 4.7 ont été utilisées. La loi de Wishart choisie comme loi *a priori* marginale de la matrice de précision, c'est-à-dire l'inverse de la matrice de variance-covariance, est simplement une généralisation de la loi Gamma à l'espace des matrices (Parent et Bernier, 2007, p 344). Elle possède les mêmes propriétés de conjugaison facilitant l'inférence bayésienne vis-à-vis d'un vecteur gaussien qu'une loi Gamma vis-à-vis d'une variable gaussienne monodimensionnelle. Elle est d'autant moins informative que son nombre de degrés de liberté, ici 2, est faible. De façon analogue au modèle de Coulomb non bruité, l'inférence a été effectuée à l'aide d'un algorithme de Gibbs. 30 000 itérations de différentes chaînes démarrant à différents points de l'espace des paramètres se sont révélées suffisantes pour garantir la convergence, les 10 000 premières constituant la période de chauffe.

Tableau 4.7 : Lois a priori pour le modèle de Coulomb bruité

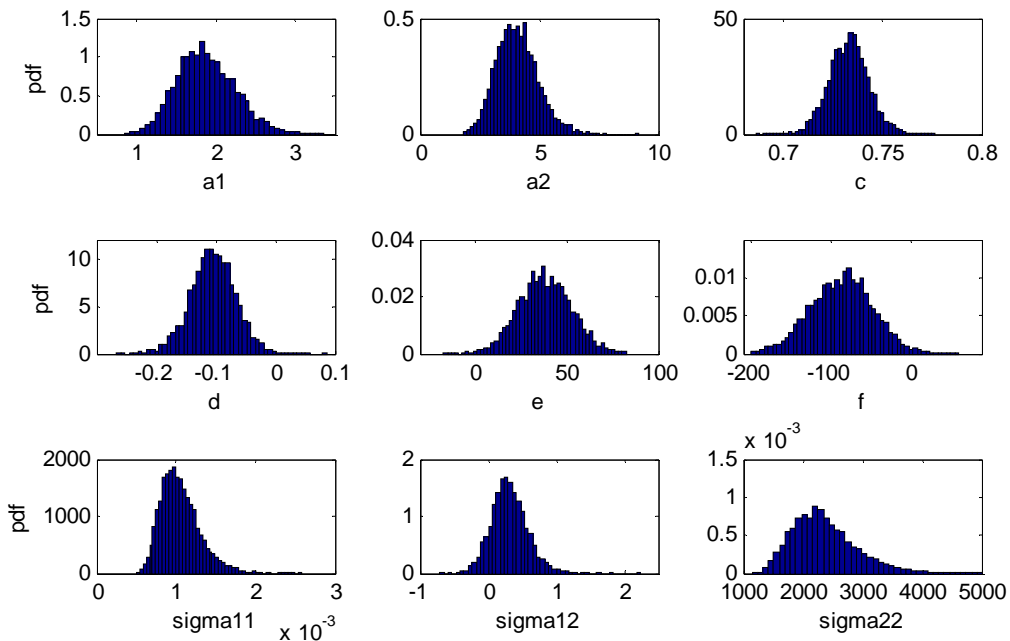
a_1	$a_1 \sim U(0,100)$ i.e. $p(a_1) = \frac{1}{100}$ pour $a_1 \in [0, 100]$
a_2	$a_2 \sim U(0,100)$ i.e. $p(a_2) = \frac{1}{100}$ pour $a_2 \in [0, 100]$
c	$\log(c) \sim U(-10,10)$ i.e. $p(\log(c)) = \frac{1}{20}$ pour $\log(c) \in [-10, 10]$
d	$d \sim U(-100,100)$ i.e. $p(d) = \frac{1}{200}$ pour $d \in [-100, 100]$
e	$e \sim U(-100,100)$ i.e. $p(e) = \frac{1}{200}$ pour $e \in [-100, 100]$
f	$f \sim U(-200,200)$ i.e. $p(f) = \frac{1}{400}$ pour $f \in [-200, 200]$
Σ	$\Sigma^{-1} \sim W_2 \left(\begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}, 2 \right)$ i.e. $p(\Sigma^{-1}) = \frac{1}{10000} \times \Sigma^{-1} ^{-\frac{1}{2}} \times \exp \left(\frac{-(\Sigma_{11}^{-1} + \Sigma_{22}^{-1})}{200} \right)$ pour $\Sigma^{-1} \in M_{2 \times 2}$ sym. def. pos.

Les *posteriors* marginaux sont présentés en Figure 4.13 et résumés dans le Tableau 4.8. La Figure 4.13 montre que toutes les lois *a posteriori* présentent des modes bien marqués, ce qui les distingue nettement des lois *a priori* peu informatives utilisées. L'information contenue dans les données a donc bien été capturée par le modèle. De façon plus détaillée, on observe que, de même que pour le couloir numéro 5 de Bessans, a_1 semble inférieur à a_2 , mais à un degré moindre. Les paramètres d et f sont négatifs au seuil des 95%, indiquant une corrélation négative significative entre les deux variables de propagation μ_{coul}, ε et l'abscisse de départ x_{start} . Le terme non diagonal de la matrice de covariance Σ_{12} n'est par contre pas significativement différent de 0. Ainsi, même si le coefficient de frottement et l'erreur de propagation semblent positivement corrélés, il ne peut être exclu que cette corrélation ne soit pas significative.

Tableau 4.8 : Statistiques descriptives des lois *a posteriori* pour le modèle de Coulomb bruité

	Moy.	SD	q _{2.5}	q _{97.5}
a_1	1.9	0.38	1.23	2.69
a_2	4.1	0.88	2.53	5.97
c	0.733	0.01	0.713	0.754
d	-0.105	0.039	-0.182	-0.03
e	38.1	14.6	9.9	66.8
f	-85.9	40	-165	-7.8
Σ_{11}	0.0011	0.0003	0.0007	0.0017
Σ_{12}	0.31	0.27	-0.19	0.89
Σ_{22}	2368	567	1501	3768
ε (m)	10.9	49.5	-86.2	108.1
μ_{coul}	0.71	0.035	0.64	0.78
x_{start} (m)	601	336	84	1344

Figure 4.13 : Loi *a posteriori* des paramètres du modèle de Coulomb bruité



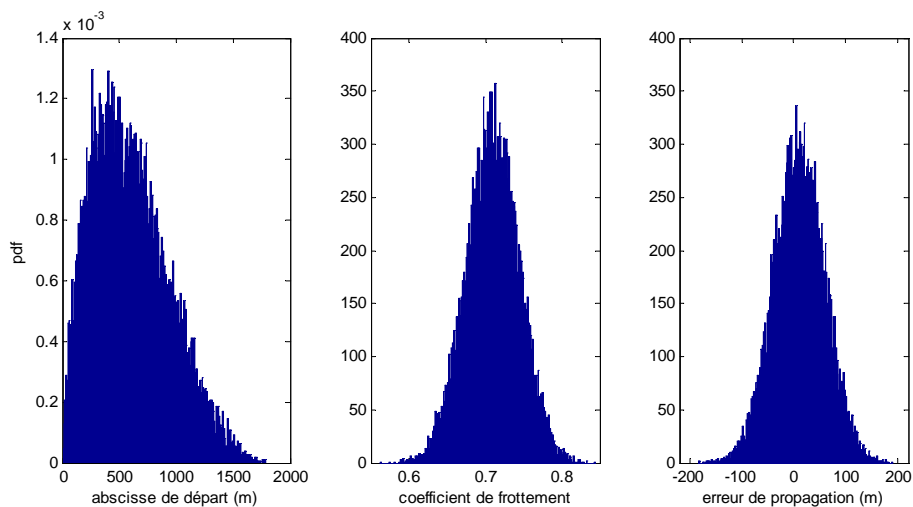
Couloir EPA 13 de la commune de Bessans, Savoie, France.

4.1.5.6. Prédictions moyennes *a posteriori*

Les distributions moyennes *a posteriori* sont calculées comme pour le modèle non bruité, c'est-à-dire par simple « plug-in » pour l'abscisse de départ et par intégration sur la variabilité de x_{start} pour le

douplet de propagation $(\mu_{coul}, \varepsilon)$. Les lois obtenues pour le cas d'étude sont présentées en Figure 4.14 et synthétisées dans le Tableau 4.8. La distribution des abscisses de départ est bornée et asymétrique. Pour ce couloir, la majorité des déclenchements se produit donc, comme pour le couloir 5 de Bessans, dans les pentes supérieures du couloir. L'asymétrie est toutefois moins marquée pour ce couloir car le mode se situe autour de $x=400$ m alors qu'il était proche de zéro pour le couloir de 5 de Bessans. Comme sur le couloir 5 de Bessans, la distribution du coefficient de frottement présente une forme globalement en cloche avec toutefois une légère asymétrie à gauche du fait de la corrélation négative avec l'abscisse de départ. On remarque de plus que, conformément aux données et de façon encore plus spectaculaire que pour le couloir 5 de Bessans, la distribution du coefficient de frottement présente une dispersion très faible par rapport à sa variabilité habituelle dans les approches statistiques dynamiques, l'intervalle de crédibilité à 95% étant limité à $[0.64, 0.78]$. Effectuer comme Maggioni (2005) des simulations de Monte Carlo sur ce couloir avec $(\mu_{coul} | \mu_o) \sim N(\mu_o, 0.2 \times \mu_o)$ en choisissant μ_o d'après la directive Suisse (Salm et al., 1990) conduirait donc à considérablement surestimer la variabilité du coefficient de frottement sur ce site et, de ce fait, à surdispenser les sorties du modèle de propagation en prédiction. Cela illustre la nécessité d'utiliser les données locales pour l'inférence plutôt qu'une distribution moyenne supposée extrapolable d'un site à l'autre. La distribution des erreurs de propagation est également en forme de cloche avec une légère asymétrie à gauche traduisant la corrélation négative avec x_{start} . Les valeurs obtenues sont relativement importantes avec un intervalle de crédibilité à 95% de $[-86, 108]$. Elles sont en particulier supérieures aux erreurs d'observation que l'on estime de l'ordre de 20 mètres. Cela indique que la géométrie du couloir est effectivement suffisamment trop complexe pour permettre un calcul de propagation exacte de l'ensemble des avalanches avec un modèle aussi rudimentaire que celui de Coulomb, ce qui justifie la prise en compte explicite de l'erreur de propagation.

Figure 4.14 : Distributions moyennes des variables du modèle de Coulomb bruité

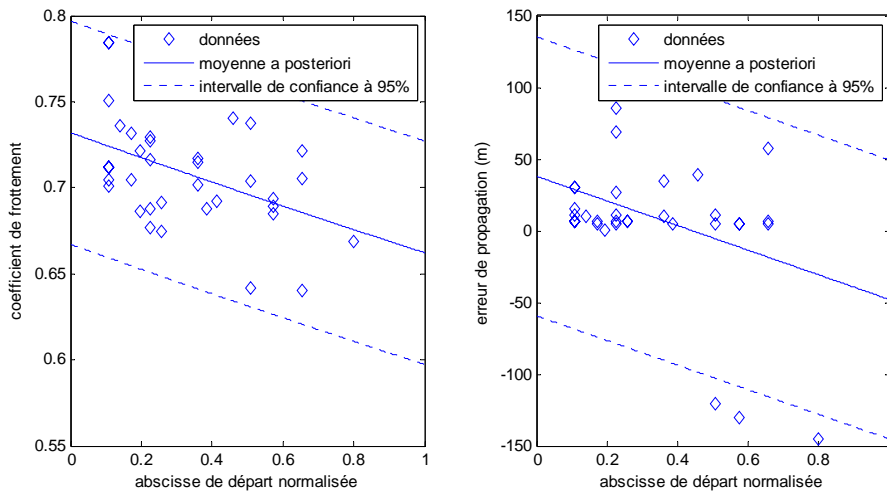


Couloir EPA 13 de la commune de Bessans, Savoie, France.

L'adéquation entre les données et le modèle de mélange est présentée en Figure 4.15. De façon analogue au modèle de Coulomb non bruité, l'ajustement est très satisfaisant pour le coefficient de frottement. En effet, toutes les données se situent dans l'intervalle de confiance à 95% et les caractéristiques des données, c'est-à-dire corrélation négative et variance à peu près constante quelle que soit l'abscisse de départ, sont bien reproduites.

L'agrément est nettement moins bon pour l'erreur de propagation puisque la tendance négative dans les données est davantage due à 3 avalanches mal reproduites qu'à une vraie diminution de l'erreur de propagation avec l'abscisse de départ. Ces trois événements sortent d'ailleurs de l'intervalle de confiance à 95%. Ainsi le modèle de mélange bivarié proposé semble robuste pour le coefficient de frottement mais moins adapté à la structure de l'erreur de propagation. La raison en est que cette dernière n'est sans doute pas normalement distribuée du fait de sa dépendance par rapport au profil de pente.

Figure 4.15 : Adéquation entre données et moyenne a posteriori du modèle de Coulomb bruité



Couloir EPA 13 de la commune de Bessans, Savoie, France.

4.1.6. Discussion : avantages et limites du modèle inversible proposé

- *Un modèle statistique souple et adapté aux données mais imparfait*

Le modèle de Coulomb et son avatar bruité possèdent le gros avantage de pouvoir être mis en œuvre très facilement à partir des données de l'EPA. Les données nécessaires sont disponibles et, du fait de l'inversibilité de l'équation du mouvement, le modèle se ramène à un modèle de mélange gaussien dont l'inférence ne pose pas de problèmes particuliers.

Conformément aux objectifs définis au chapitre 3.1, une structure suffisamment souple pour prendre en compte toutes les corrélations intervariables a été construite. Etant donné le rôle primordial en prédiction du coefficient de frottement sur lequel on reviendra au chapitre 6, c'est cette variable qui a focalisé l'essentiel des efforts de modélisation et le modèle proposé semble s'ajuster correctement aux données. Concernant l'abscisse de départ et l'erreur de propagation quand cette dernière est utilisée, les modèles proposés sont plus rudimentaires. Une distribution beta est sans doute trop simpliste pour rendre compte de la variabilité véritable des abscisses de départ, notamment pour les couloirs tels le site EPA 13 de Bessans où différents ressauts sont observés en zone de départ. Le développement d'un modèle de mélange pour les abscisses de départ pourrait donc constituer un prolongement intéressant, mais vraisemblablement peu utile en termes de prédétermination des variables de sortie du fait de la capacité d'un écoulement avalancheux à « oublier » ses conditions initiales. Quant à l'erreur de propagation, sa distribution non gaussienne dépend fortement de la topographie. Vouloir la modéliser de façon plus réaliste revient donc à renoncer à l'idée de séparer inférence statistique et calculs de propagation, ce qui réduit quasiment à zéro l'intérêt d'un modèle de propagation inversible.

- *Prendre en compte davantage de variables?*

Même si l'une des principales forces du modèle de Coulomb (éventuellement bruité) est de ne nécessiter que les abscisses de départ et d'arrivée, d'autres variables quantitatives comme la surface de départ ou la hauteur de la fracture peuvent être incorporées dans le modèle si elles sont disponibles. Toutefois, étant donné qu'elles ne sont pas nécessaires à la propagation, elles ne peuvent intervenir que sous la forme de covariables permettant d'offrir une représentation plus réaliste des variations de x_{start} , μ_{coul} et éventuellement ε . Du fait du rôle crucial du coefficient de frottement et des données EPA disponibles, il conviendrait de tester en priorité l'introduction dans le modèle proposé de la dépendance du coefficient de frottement avec le volume de dépôt v_{stop} établie par Ancy (2005). Celle-ci peut s'effectuer naturellement par l'introduction d'un paramètre supplémentaire d' dans le modèle de mélange utilisé jusqu'ici (eq. 4.9).

$$\mu_{coul} \mid x_{start}, v_{stop}, c, d, d', \sigma \sim N\left(c \times \exp\left(d \times x_{start} + d' \times v_{stop}\right), \sigma\right) \text{ (eq 4.9)}$$

- *Un modèle de propagation insuffisant*

L'utilisation d'un modèle de propagation aussi rudimentaire que celui de Coulomb est fortement discutable. Pour des couloirs à la pente régulière comme le couloir EPA numéro 5 de Bessans, le coefficient de frottement peut être obtenu par inversion de l'équation du mouvement avec une vitesse initiale nulle. La valeur obtenue possède ainsi un sens physique bien connu dans d'autres disciplines telle l'étude des glissements de terrain. Malheureusement, de tels couloirs sont rares, de sorte que le modèle de Coulomb non bruité est davantage une curiosité qu'un outil opérationnel utile. La plupart des couloirs réels présentent en effet, comme le couloir 13 de Bessans, des irrégularités dans le profil de pente rendant impossible l'utilisation du modèle de Coulomb non bruité et nécessitant l'introduction d'une vitesse initiale non nulle et d'une erreur de propagation. Les modifications ainsi introduites font perdre au modèle de Coulomb originel une bonne part de son sens physique en augmentant artificiellement les valeurs de frottement obtenues. Le modèle de propagation devient ainsi un opérateur complètement empirique reliant des entrées et des sorties. De plus, son efficacité, même sous forme non bruitée, reste limitée aux distances d'arrêt.

Il convient donc maintenant de voir comment le même formalisme probabiliste peut être associé à un modèle de propagation plus riche permettant de mieux décrire la physique de l'écoulement tout en donnant accès à la distribution de davantage de variables présentant un intérêt pour le zonage de l'aléa.

4.2. Un modèle intermédiaire non inversible

4.2.1. Modélisation

Le modèle de propagation que l'on utilise à présent reste de type bloc glissant mais avec une loi de frottement de Voellmy (eq. 3.5) avec deux coefficients de frottement μ_{voel} et ξ_{voel} . Pour des valeurs raisonnables du paramètre ξ_{voel} , on a ainsi $\mu_{voel} \gg \mu_{coul}$ pour un même couple (x_{start}, x_{stop}) , de sorte que le problème des avalanches ne démarrant pas avec une vitesse initiale nulle ne se pose pratiquement plus. A partir de maintenant, on adoptera donc systématiquement $v_{zero} = 0$, ce qui est tout de même physiquement plus satisfaisant. De plus, la loi de frottement à deux coefficients de frottement offre davantage de souplesse. Ainsi, d'une part les profils de vitesse obtenus sont plus réalistes et peuvent de ce fait être exploités, et, d'autre part, le modèle « sait » s'adapter à des topographies plus complexes, de sorte que l'erreur de propagation devient plus faible.

En contrepartie, le modèle de propagation n'est plus inversible. μ_{voel} et $x_{stop_{cal}}$ sont donc cette fois des variables latentes et les calculs d'inférence sont plus délicats car ils ne peuvent être dissociés de la propagation. La résolution de l'équation du mouvement s'effectue numériquement sur le profil topographique discrétisé. La distance d'arrêt s'obtient toutefois de façon exacte en calculant à l'aide de l'eq. 3.32 l'abscisse où la vitesse devient nulle.

Le modèle de type Beta pour l'abscisse de départ normalisée est conservé (eq. 4.3). De même, on conserve un modèle de mélange gaussien pour le coefficient de frottement avec trois paramètres c, d, e quantifiant les corrélations avec x_{start} et h ainsi que la moyenne en absence de corrélation et toujours une variance σ^2 constante. Puisque la forme fonctionnelle du lien entre variables d'entrée est apparue peu importante avec le modèle de Coulomb, on se contente d'un modèle linéaire (eq. 4.10).

$$\mu_{voel} | c, d, e, \sigma, x_{start}, h \sim N(c + dx_{start} + eh, \sigma) \quad (\text{eq. 4.10})$$

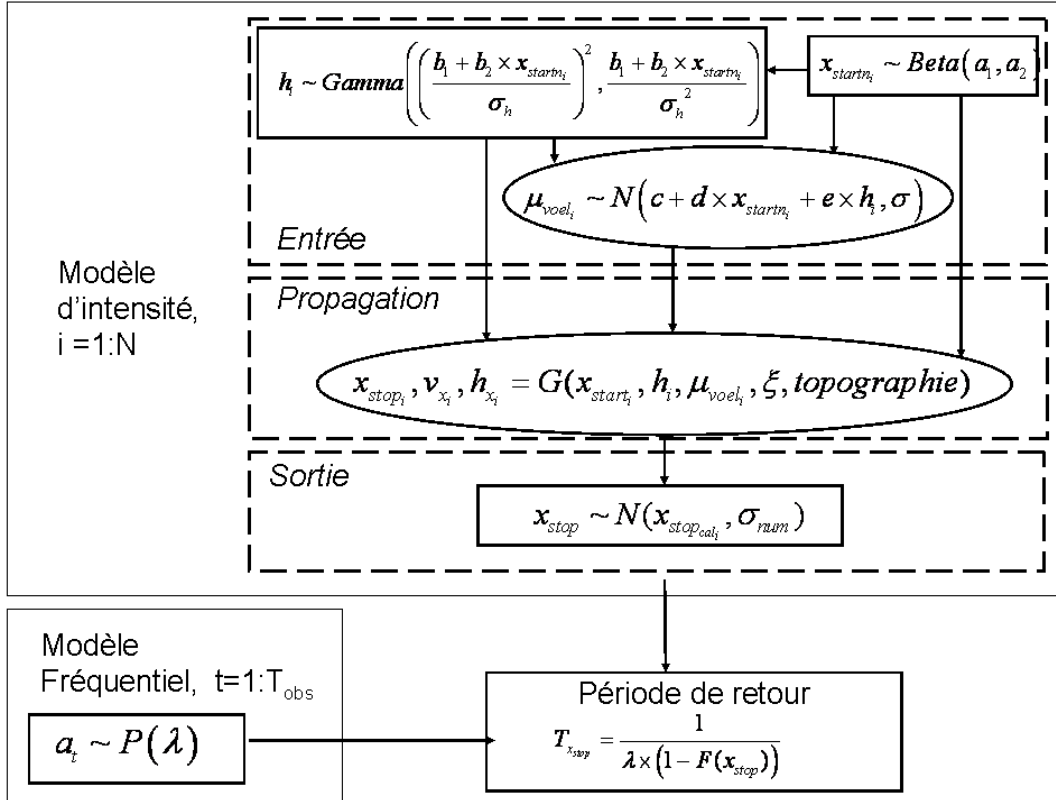
Une variable d'entrée supplémentaire, la hauteur du bloc en écoulement h , doit être modélisée. Pour elle, un modèle de mélange Gamma est proposé. Sa justification première est que, globalement, la hauteur en écoulement est une variable asymétrique au même titre qu'une hauteur de précipitation ou qu'un débit de crue et que pour de telles variables « hydrologiques », un tel modèle est souvent choisi. On tient toutefois à garder la possibilité de prendre en compte toutes les corrélations intervariables, de sorte que l'on modélise la dépendance avec l'abscisse de départ grâce aux paramètres b_1 et b_2 représentant, comme pour le coefficient de frottement, la corrélation linéaire avec l'abscisse de départ normalisée et la moyenne en absence de corrélation avec une variance σ_h^2 constante. Les deux paramètres d'échelle et de forme de la loi Gamma s'expriment facilement en fonction des trois paramètres b_1, b_2, σ_h et de l'abscisse de départ normalisée x_{start} (eq. 4.11).

$$h|b_1, b_2, \sigma_h, x_{start} \sim \text{Gamma}\left(\frac{1}{\sigma_h^2}(b_1 + b_2 x_{start})^2, \frac{1}{\sigma_h}(b_1 + b_2 x_{start})\right) \quad (\text{eq. 4.11})$$

L'erreur de propagation traduisant la différence entre valeurs observées et recalculées est supposée très faible. On néglige donc sa corrélation avec les autres variables contrairement à ce qui avait été fait pour le modèle de Coulomb bruité. En revanche, on ne peut pas négliger complètement l'erreur de propagation pour des raisons pratiques de mise en œuvre de l'algorithme d'estimation. Du fait de l'imperfection du modèle de propagation, on n'arrive en effet jamais à reproduire parfaitement certaines des distances d'arrêt réellement observées, de sorte que l'on est obligé de conserver « un peu de mou » pour retenir les candidats acceptables lors de l'étape d'acceptation/rejet de l'algorithme MCMC. A la différence du modèle de Coulomb bruité, on interprète donc l'erreur de propagation avec ce modèle dit « Voellmy bloc glissant » comme une erreur numérique que l'on modélise par un bruit blanc. On note σ_{num} sa variance que l'on fixe afin de rendre le modèle identifiable (eq. 4.12).

$$x_{stop} | \sigma_{num}, x_{stop_{cal}} \sim N(x_{stop_{cal}}, \sigma_{num}) \quad (\text{eq. 4.12})$$

Figure 4.16 : DAG des modèles de Voellmy bloc glissant et Saint-Venant



Le modèle fréquentiel est le modèle poissonien présenté au chapitre 5. Pour le modèle d'intensité, les variables de sortie sont la distance d'arrêt et les profils longitudinaux de vitesse et de hauteur. Deux coefficients de frottement μ_{voel} et ξ_{voel} sont considérés, le premier étant latent et le second un paramètre du modèle de propagation. Les variables d'entrée sont limitées à l'abscisse de départ, à la hauteur de départ et au coefficient de frottement latent. Soit $x' = (x_{start}, h)$, $x'' = \mu_{voel}$, $y = (x_{stop_{cal}}, h_x, v_x)$ et $\theta_G = \xi_{voel}$. Les paramètres des modèles de fréquence et d'intensité sont $\theta_F = \lambda$, et $\theta_M = (\theta_o, \theta_G) = (a_1, a_2, b_1, b_2, \sigma_h, c, d, e, \sigma, \xi)$ respectivement.

Avec le modèle de Voellmy bloc glissant, la hauteur de l'avalanche reste constante au cours de son écoulement, soit $h_x = h$. Pour le modèle de Saint-Venant $h = h_{plaque}$, c'est-à-dire la hauteur de la couche de neige mise en mouvement dans la zone de départ. La hauteur h_x varie ensuite au cours de l'écoulement en fonction des caractéristiques du terrain et de la neige.

Le modèle proposé est résumé sous la forme d'un DAG en Figure 4.16. Il comprend 10 paramètres, puisque $\theta_M = (a_1, a_2, b_1, b_2, \sigma_h, c, d, e, \sigma, \xi_{voel})$. Selon les notations introduites au chapitre 3, seul ξ_{voel} correspond à θ_G , c'est-à-dire au paramètre du modèle déterministe de propagation, tandis que les 9 autres paramètres correspondent à θ_o , c'est-à-dire aux paramètres des distributions des variables d'entrée : deux pour l'abscisse de départ, trois pour la hauteur de bloc et quatre pour le coefficient de frottement μ_{voel} . Toutefois, les inconnues du problème comprennent également les deux variables latentes $\mu_{voel}, x_{stop_{cal}}$. La première est une variable d'entrée, soit $x'' = \mu_{voel}$. La seconde est une variable de sortie au même titre que le profil de vitesse, soit $y = (x_{stop_{cal}}, v_x)$.

4.2.2. Inférence bayésienne

L'inférence des paramètres $\theta_M = (a_1, a_2, b_1, b_2, \sigma_h, c, d, e, \sigma, \xi_{voel})$ et des variables latentes $(\mu, x_{stop_{cal}}) = (\mu_1, x_{stop_{cal_1}}, \dots, \mu_N, x_{stop_{cal_N}})$ s'effectue au moyen du théorème de Bayes (eq. 4.13). Le *posterior* $p(\theta_M, \mu_{voel}, x_{stop_{cal}} | data_2, \sigma_{num})$ dépend des données, de l'erreur numérique tolérée et du *prior*.

$$p(\theta_M, \mu_{voel}, x_{stop_{cal}} | data_2, \sigma_{num}) \quad (eq. 4.13)$$

$$\propto p(\theta_M) \times \prod_{i=1}^N \left(l(x_{start_i}, h_i, x_{stop_i} | \theta_M, \mu_{voel_i}, x_{stop_{cal_i}}, \sigma_{num}) \times p(\mu_{voel_i}, x_{stop_{cal_i}} | \theta_M, x_{start_i}, h_i, x_{stop_i}, \sigma_{num}) \right)$$

La vraisemblance de chaque événement sachant les paramètres et les variables latentes $l(x_{start_i}, h_i, x_{stop_i} | \theta_M, \mu_{voel_i}, x_{stop_{cal_i}}, \sigma_{obs})$ s'exprime grâce aux lois conditionnelles spécifiant le modèle proposé (eq. 4.14). La probabilité d'occurrence des variables latentes $p(\mu_{voel_i}, x_{stop_{cal_i}} | \theta_M, x_{start_i}, h_i, x_{stop_i}, \sigma_{num})$ se décompose en deux termes. Le premier correspond au modèle de mélange pour le coefficient de frottement et le second à la propagation déterministe avec le modèle G, δ étant la distribution de Dirac (eq 4.15).

$$l(x_{start_i}, h_i, x_{stop_i} | \theta_M, \mu_{voel_i}, x_{stop_{cal_i}}, \sigma_{num}) = l(x_{start_i} | a_1, a_2) \times l(h_i | b_1, b_2, \sigma_h, x_{start_i}) \times l(x_{stop_i} | \sigma_{num}, x_{stop_{cal_i}}) \quad (eq. 4.14)$$

$$p(\mu_{voel_i}, x_{stop_{cal_i}} | \theta_M, x_{start_i}, h_i, x_{stop_i}, \sigma_{num}) = p(\mu_{voel_i} | c, d, e, \sigma, x_{start_i}, h_i) \times \delta(G(x_{start_i}, h_i, \mu_{voel_i}, \xi_{voel})) \quad (eq. 4.15)$$

Pour l'inférence de toutes les inconnues du modèle, un algorithme d'estimation de type Metropolis-Hastings a été construit. Conformément aux remarques faites au chapitre 2, une structure séquentielle a été privilégiée afin de pouvoir, dans la mesure du possible, ajuster l'exploration de façon unidirectionnelle et limiter ainsi le nombre d'itérations nécessaires à la convergence. Pour l'abscisse de départ et la hauteur de bloc, cela ne pose pas de problèmes particuliers. Pour le coefficient de frottement latent, en revanche, l'itération séquentielle n'est pas possible car la propagation de tous les événements dépend d'une même valeur de ξ_{voel} . Les deux variables latentes et le paramètre ξ_{voel} doivent donc être calibrés au cours d'une même étape d'acceptation/rejet.

Les fonctions d'exploration unidimensionnelles employées pour l'algorithme sont présentées dans le Tableau 4.9. Elles sont de type normales centrées sur la valeur à l'itération k de l'algorithme pour les paramètres pouvant être positifs et négatifs et de type log-normal centrées sur la valeur à l'itération k de l'algorithme pour les paramètres strictement positifs, c'est-à-dire les paramètres de variance σ et σ_h ainsi que ξ_{voel} . La variable latente μ_{voel} est bien entendu également strictement positive mais, du fait de la faible variance d'exploration nécessaire au bon fonctionnement de l'algorithme, une exploration normale peut être employée sans risque de génération de valeurs négatives.

Le contrôle de la convergence s'effectue classiquement par comparaison de deux chaînes démarrant à des points différents de l'espace des paramètres. Celles-ci permettent le calcul de la variable de test de Gelman et Rubin présentée au chapitre 2, c'est-à-dire le rapport des variances intra et inter-chaînes, celui-ci tendant vers 1 lorsque la convergence est atteinte.

L'adaptation de l'algorithme d'estimation au cas d'étude requiert le réglage des variances d'exploration et le choix de valeurs initiales. Les variances d'exploration doivent être ajustées afin d'obtenir des taux d'acceptation conformes à ceux recommandés par la littérature pour obtenir une convergence rapide, c'est-à-dire environ 25% pour ξ_{voel} et les deux variables latentes car l'acceptation-rejet concerne alors plus de 3 inconnues simultanément et environ 35% pour tous les autres paramètres pour lesquels l'acceptation-rejet se fait marginalement. Quant au choix des valeurs initiales pour les différentes chaînes, il doit à la fois permettre à l'algorithme de démarrer et à la convergence d'être testée. Il faut donc spécifier des valeurs à la fois « raisonnables » et suffisamment distantes les uns des autres dans l'espace des paramètres, le compromis pouvant être difficile à trouver suivant le cas d'étude traité.

4.2.3. Inférence sur des données simulées

4.2.3.1. Simulation des données

Pour vérifier le bon fonctionnement de l'algorithme construit et en particulier tester sa capacité à inférer les lois cibles recherchées, un jeu de données simulées a été créé. La topographie très simple du couloir numéro 5 de Bessans a été reprise et on y a simulé 50 événements, c'est-à-dire un jeu de 50 variables d'entrée/sortie, à l'aide des valeurs de paramètres du Tableau 4.13. Ce faisant, les 50 valeurs du coefficient de frottement latent correspondantes ont bien sûr également été simulées. Notons que l'on a volontairement choisi des paramètres de corrélation (b_2, d, e) non nuls de façon à vérifier la capacité de l'algorithme à les inférer. De plus, on n'a pas bruité les sorties du modèle de propagation car on suppose que les sorties déterministes $x_{stop_{cal}}$ sont les vraies abscisses d'arrêt des avalanches et que σ_{num} , que l'on fixe ici à son minimum de 5 m qui correspond au pas de discrétisation des calculs numériques de propagation, n'est qu'un moyen de permettre à l'algorithme d'estimation de tourner. Les 50 événements simulés sont présentés en Annexe A5.

4.2.3.2. Construction de l'algorithme

Les variances d'exploration présentées dans le tableau 4.9 ont été ajustées afin d'obtenir des taux d'acceptation conformes à ceux recommandés par la littérature pour obtenir une convergence rapide. Le Tableau 4.10 montre que ces taux cibles ont effectivement pu être atteints, avec un taux d'acceptation de 28% pour le saut multidimensionnel associé aux coefficients de frottement et des taux compris entre 32 et 40% pour les autres paramètres suivant le paramètre et la chaîne considérés.

Tableau 4.9 : Fonctions d'exploration pour les données simulées

$a_1(k) \sim N(a_1(k-1), 0.4)$	$c(k) \sim N(c(k-1), 0.02)$
$a_2(k) \sim N(a_2(k-1), 0.45)$	$d(k) \sim N(d(k-1), 0.05)$
$b_1(k) \sim N(a_1(k-1), 0.05)$	$e(k) \sim N(e(k-1), 0.0095)$
$b_2(k) \sim N(b_2(k-1), 0.3)$	$\sigma(k) \sim \log N(\log(\sigma(k-1)), 0.35)$
$\sigma_h^2(k) \sim \log N(\log(\sigma_h^2(k-1)), 0.5)$	$\xi_{voel}(k) \sim \log N(\log(\xi_{voel}(k-1)), 0.01)$
	$\mu_{voel_i}(k) \sim N(\mu_{voel_i}(k-1), 0.0028)$

Tableau 4.10 : Taux d'acceptation pour les données simulées

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel} et μ_{voel_i}
Chaîne 1	0.36	0.36	0.32	0.35	0.38	0.39	0.39	0.39	0.33	0.28
Chaîne 2	0.36	0.35	0.32	0.36	0.38	0.40	0.39	0.39	0.33	0.28

4.2.3.3. Loi a priori

Dans un premier temps, des lois *a priori* très peu informatives ont été testées. Malheureusement, il n'a alors pas été possible de faire converger l'algorithme car les différentes chaînes allaient se bloquer dans des modes locaux de la loi *a posteriori*. Des lois *a priori* plus informatives sont donc ici

nécessaires. Elles ont été choisies sur la base de considérations expertes de façon à restreindre l'espace des valeurs possibles à un intervalle raisonnable pour que l'algorithme ne se bloque pas tout en n'étant tout de même pas trop informatives pour laisser les données s'exprimer.

Des lois uniformes ont pu être utilisées pour les paramètres de la distribution des abscisses de départ dont l'inférence ne pose pas de difficulté. Des *priors* assez informatifs ont par contre été choisis pour la moyenne du coefficient de frottement latent c , le paramètre de variance σ et surtout le coefficient de frottement turbulent ξ_{voel} . Pour ces paramètres, la connaissance experte disponible est d'ailleurs suffisamment abondante pour permettre l'élicitation, notamment au travers des valeurs tabulées de coefficient de frottement proposées par la directive suisse (Salm et al., 1990). Enfin, les paramètres de corrélation ont été *a priori* centrés sur zéro de façon à laisser les données trancher le sens de la corrélation éventuelle entre les différentes variables du modèle (Tableau 4.11).

Tableau 4.11 : Lois a priori pour les données simulées

$a_1 \sim U(0,10)$	$c \sim N(0.5,10)$
$a_2 \sim U(0,10)$	$d \sim N(0,10)$
$b_1 \sim N(3,1)$	$e \sim N(0,10)$
$b_2 \sim N(0,1)$	$\sigma \sim \text{Gamma}(1,10)$
$\sigma_h^2 \sim \text{Gamma}(10,100)$	$\xi_{voel} \sim N(700,50)$

4.2.3.4. Fonctionnement de l'algorithme et contrôle de la convergence

A chaque itération de l'algorithme, 50 itérations du modèle de propagation sont effectuées, i.e. une par avalanche. Du fait de la longueur limitée du couloir et de la simplicité du modèle de bloc glissant utilisé, le temps de calcul reste toutefois raisonnable puisque l'ordre de grandeur pour effectuer 80 000 itérations avec deux chaînes est celui du jour. Les 30 000 premières ont été oubliées (période de chauffe) et les 50 000 suivantes ont été utilisées pour contrôler la convergence et estimer les inconnues.

La Figure 4.17 montre la trajectoire sur les 50 000 itérations retenues de deux chaînes pour deux paramètres : a_1 , l'un des paramètres de la loi des abscisses de départ, et le coefficient de frottement ξ_{voel} . Pour a_1 , le domaine de variation est bien stabilisé, les variations sont fortes d'une itération à l'autre et les deux chaînes se mélangent bien, le tout indiquant une convergence sûre et facile. Pour ξ_{voel} , les deux chaînes se mélangent également mais moins facilement avec une forte autocorrélation des chaînes. Cela indique que, même si la convergence est atteinte, beaucoup d'itérations sont nécessaires pour simuler la loi *a posteriori* avec une précision satisfaisante.

Le calcul de la variable de test de Gelman et Rubin permet de confirmer l'impression visuelle et de systématiser le contrôle de la convergence pour toutes les inconnues. Le Tableau 4.12 montre, comme le suggère la Figure 4.17, que, pour tous les paramètres à l'exception de ξ_{voel} , la convergence ne pose pas de problèmes. Pour ξ_{voel} , en revanche, 50 000 itérations au delà de la période de chauffe sont nécessaires pour obtenir un rapport de Gelman-Rubin de 1.02, ce qui est certes largement suffisant pour, en pratique, accepter l'hypothèse que la convergence est atteinte mais indique tout de même qu'un nombre important d'itérations de l'algorithme doit être effectué.

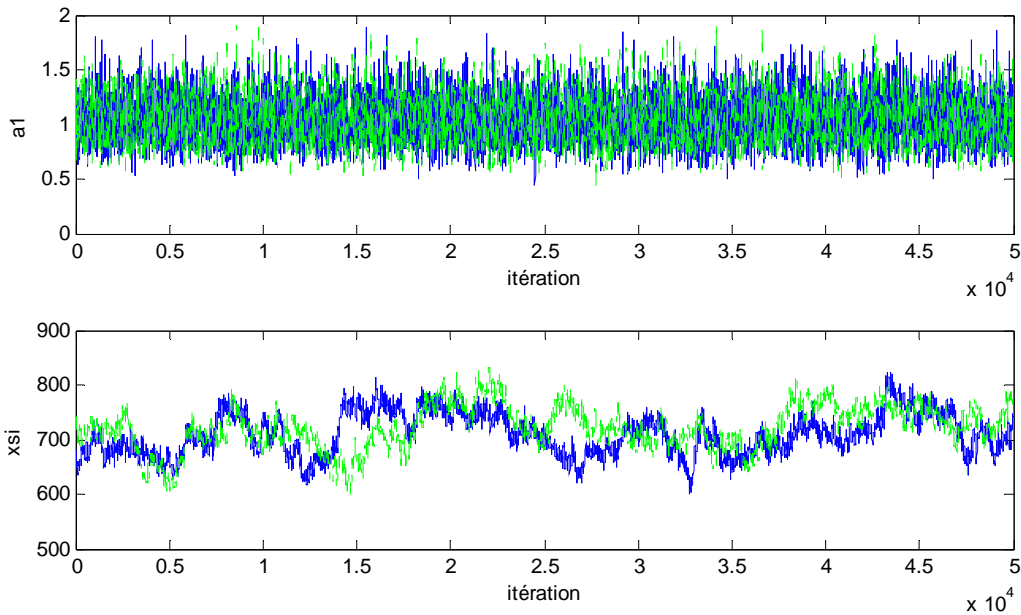
Tableau 4.12 : Test de convergence de Gelman et Rubin avec les données simulées

a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}
1	1	1	1	1	1	1	1	1	1.02

Le calcul de la variable de test peut également s'effectuer sur les variables latentes. La Figure 4.18 propose un histogramme des 50 valeurs obtenues pour l'estimation des coefficients de frottement latents μ_{voel_i} . Elle montre que la convergence est là aussi atteinte avec toutes les valeurs de la variable de test inférieures à 1.04. Néanmoins, toutes les valeurs à l'exception de 2 étant très proches de 1, la Figure 4.18 montre également que la convergence est très bonne pour la plupart des événements mais

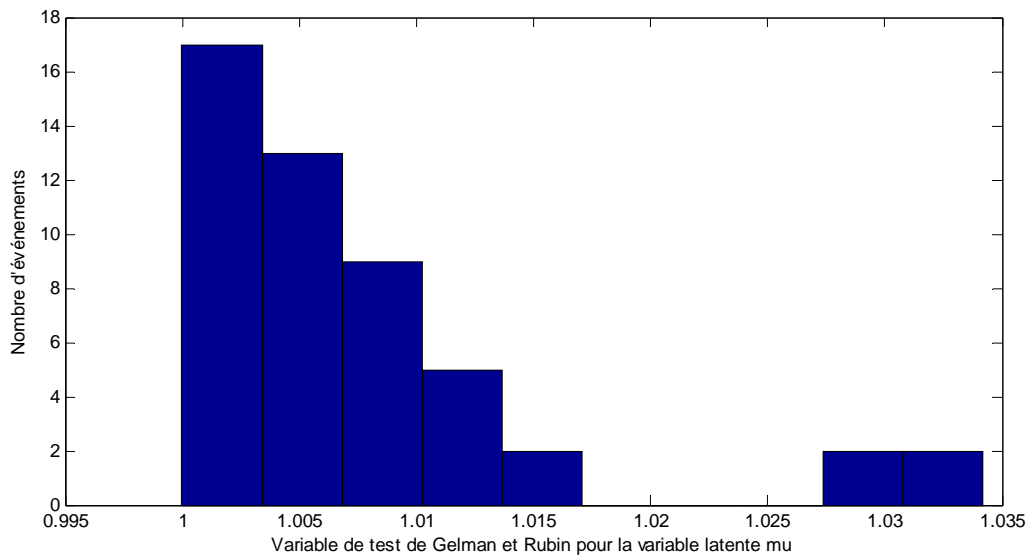
plus problématique pour certains événements particuliers pour lesquels la calibration des coefficients de frottement est difficile.

Figure 4.17 : Valeurs successives de deux chaînes avec les données simulées



Topographie du couloir EPA 5 de la commune de Bessans, Savoie, France. 50 événements fictifs.

Figure 4.18 : Convergence des variables latentes avec les données simulées



Topographie du couloir EPA 5 de la commune de Bessans, Savoie, France. 50 événements fictifs.

4.2.3.5. Résultats

- Paramètres

En termes d'inférence des paramètres, les résultats de l'algorithme sont très satisfaisants. Pour tous sauf ξ_{voel} , la « vraie » valeur utilisée en simulation se retrouve dans l'intervalle de crédibilité à 95%.

De plus, l'erreur sur l'estimation ponctuelle est faible, tout au plus 0.33 écarts-types (Tableau 4.13). Ce résultat est d'autant plus remarquable que les priors utilisés étaient pour plusieurs paramètres relativement informatifs, de sorte que des estimateurs ponctuels assez différents de ceux utilisés pour la simulation des données auraient pu être obtenus même avec un algorithme qui fonctionne correctement.

Pour ξ_{voel} , la valeur estimée est substantiellement supérieure à la valeur utilisée en simulation avec une erreur de près de trois écart-types. Plutôt qu'un réel dysfonctionnement de l'algorithme, ce résultat traduit la force du *prior* utilisé pour l'inférence, c'est-à-dire une loi normale centrée sur 700 avec une faible variance. Il est à relier au problème discuté au chapitre 4.2.5 de l'existence d'une infinité de couples de coefficients de frottement pouvant correspondre à une même distance d'arrêt, le modèle de Coulomb présenté au chapitre 4.1 correspondant au cas limite peu réaliste de ξ_{voel} tendant vers l'infini.

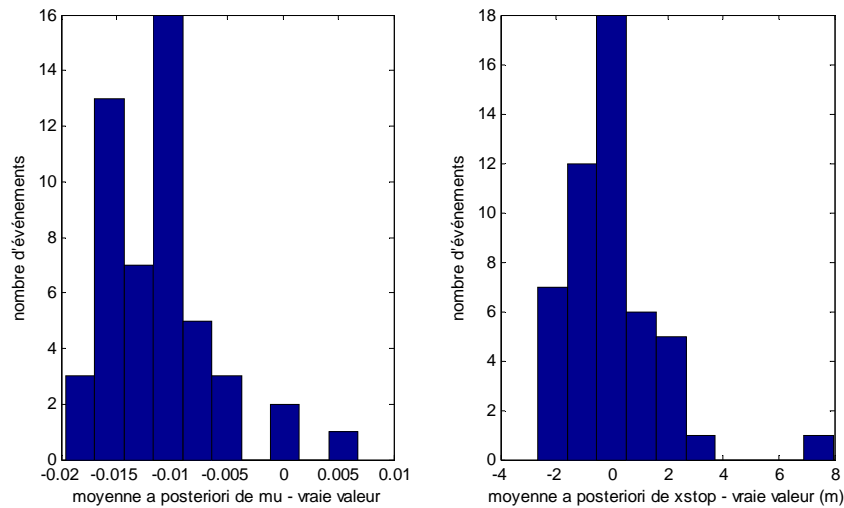
Tableau 4.13 : Lois a posteriori et vraies valeurs pour les données simulées

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}
Moy.	1.04	2.13	2.98	-2.96	0.31	0.59	-0.28	-0.053	0.051	715.7
SD	0.19	0.43	0.07	0.17	0.028	0.067	0.072	0.022	0.006	39.5
q_{2.5}	0.71	1.38	2.83	-3.27	0.26	0.45	-0.42	-0.095	0.041	638.2
q_{97.5}	1.45	3.05	3.12	-2.63	0.37	0.72	-0.13	-0.008	0.063	788.3
vraie valeur	1	2	3	-3	0.3	0.6	-0.3	-0.05	0.05	600
erreur	0.23	0.31	0.33	0.23	0.28	0.17	0.33	0.12	0.1	2.9

- *Variables latentes*

Pour ce qui est des variables latentes, les résultats sont également satisfaisants (Figure 4.19). Les moyennes *a posteriori* des distances d'arrêt simulées sont extrêmement proches des « vraies » valeurs. Les erreurs de propagation sont en effet toutes inférieures à l'erreur σ_{num} fixée ici à 5 m sauf pour un événement. La calibration des distances d'arrêt semble donc s'effectuer de manière très satisfaisante et non biaisée. Les valeurs de coefficients de frottement obtenues sont également très proches des valeurs utilisées en simulation mais systématiquement sous-estimées, ce biais compensant celui observé pour ξ_{voel} , quant à lui surestimé.

Figure 4.19 : Différence entre moyenne a posteriori des variables latentes et vraies valeurs

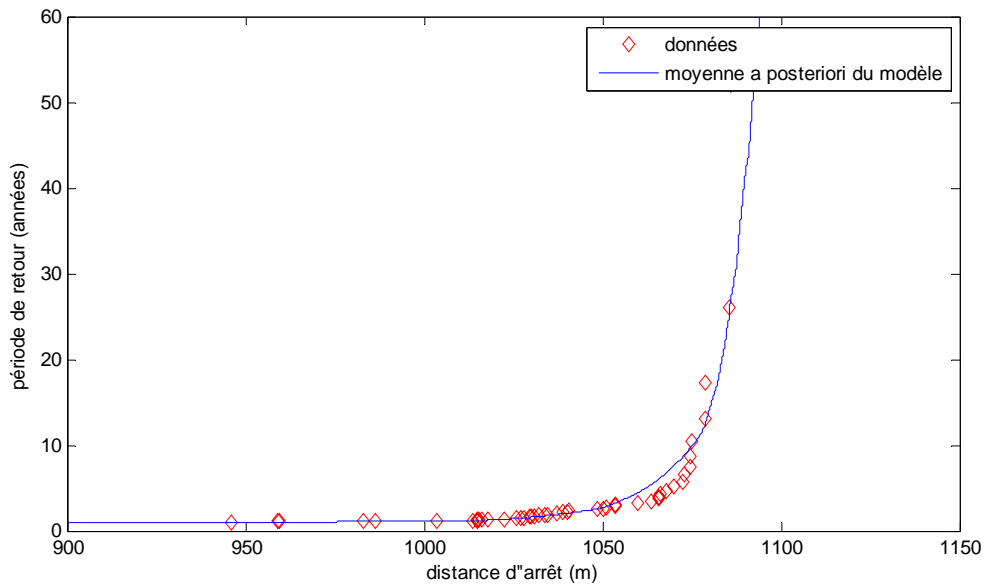


Topographie du couloir EPA 5 de la commune de Bessans, Savoie, France. 50 événements fictifs.

- *Adéquation entre modèle et données*

Enfin, l'adéquation du modèle aux données simulées en termes de périodes de retour associées aux distances d'arrêt mérite d'être soulignée (Figure 4.20). Les périodes de retour empiriques correspondent très bien à la moyenne du modèle. Cela traduit une très bonne adéquation entre les « données » de distance d'arrêt et la moyenne du modèle après inférence dont l'obtention est détaillé au chapitre 6.1. L'accord est en particulier remarquable pour les distances d'arrêt les plus élevées correspondant aux périodes de retour les plus élevées. Le problème de compensation entre coefficients de frottement ne semble donc pas affecter la capacité prédictive du modèle. On y reviendra là aussi au chapitre 4.2.5.

Figure 4.20 : Adéquation entre modèle et données simulées



Topographie du couloir EPA 5 de la commune de Bessans, Savoie, France. 50 événements fictifs. Les périodes de retour sont calculées avec $\lambda = 1$, c'est-à-dire avec une fréquence moyenne de une avalanche par an.

4.2.4. Application à des données réelles

4.2.4.1. Cas d'étude choisi

Le cas d'étude choisi est le même que pour le modèle de Coulomb bruité, à savoir le couloir EPA 13 de la commune de Bessans. Le problème principal pour l'utilisation du modèle est que les hauteurs de bloc, si tant est qu'elles existent, ne sont évidemment pas renseignées dans l'EPA. Etant donné qu'il s'agit d'une variable quantitative représentant l'intensité de l'avalanche en termes de matériau mobilisé, il paraît naturel de penser qu'elle est fortement liée au volume de neige en écoulement. Meunier et al. (2004) ont donc proposé d'obtenir les hauteurs de bloc à partir des volumes de dépôts renseignés dans l'EPA. Ils utilisent pour ce faire une transformation déterministe basée sur un raisonnement expert et un calage sur un couloir haut-savoyard. L'idée est que la dépendance de la hauteur de bloc avec le volume de dépôt est « en racine carré », c'est-à-dire que la hauteur du bloc détermine fortement le volume de dépôt. De plus, une hauteur minimale de 2.5 m est proposée pour un volume tendant vers zéro et une hauteur de 7.5 m est supposée raisonnable pour une très grosse avalanche de 500 000 m³ (eq. 4.16). Bien que nous travaillions sur un couloir différent de celui pour lequel cette relation a été obtenue, elle semble suffisamment générale pour pouvoir être transposée à d'autres couloirs de taille moyenne à importante tel le cas d'étude choisi. C'est en tout cas l'hypothèse effectuée.

$$h = 2.5 + 5 \times 10^{-3} \times \sqrt{V_{stop}} \quad (\text{eq. 4.16})$$

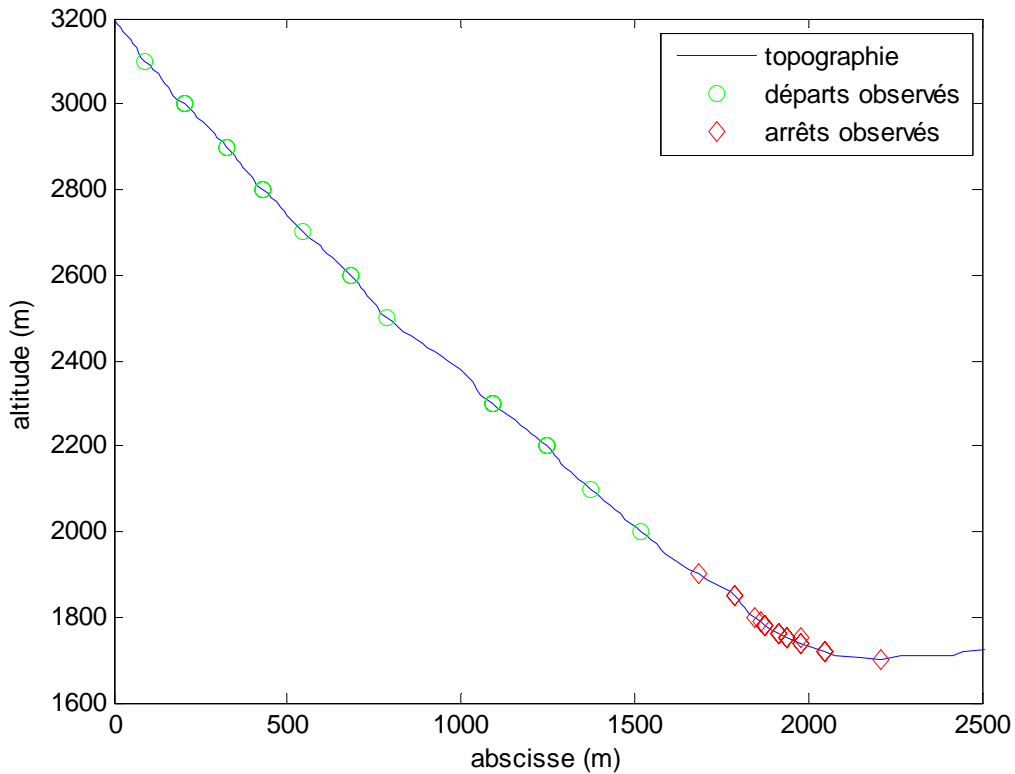
Les volumes de dépôts n'étant pas systématiquement disponibles, seuls $N=26$ événements peuvent être retenus pour l'inférence du modèle, soit 14 de moins que pour le modèle de Coulomb bruité. On est même contraint pour conserver une taille d'échantillon suffisante d'utiliser des avalanches beaucoup plus anciennes qu'avec le modèle de Coulomb bruité, ce qui n'est pas sans poser des problèmes de qualité des données. La répartition le long du couloir des abscisses de départ et d'arrêt reste toutefois assez analogue à celle décrite au chapitre 4.1.5 (Figure 4.21).

Tableau 4.14 : Données pour l'application du modèle de Voellmy bloc glissant

	Moy.	SD	q _{2.5}	q _{97.5}
x_{start} (m)	677.6	443.5	106.5	1496.9
v_{stop} (m ³)	69 226	169 590	5400	674 760
h (m)	3.37	1.01	2.59	6.6
x_{stop} (m)	1933.1	104.8	1701.3	2183.5

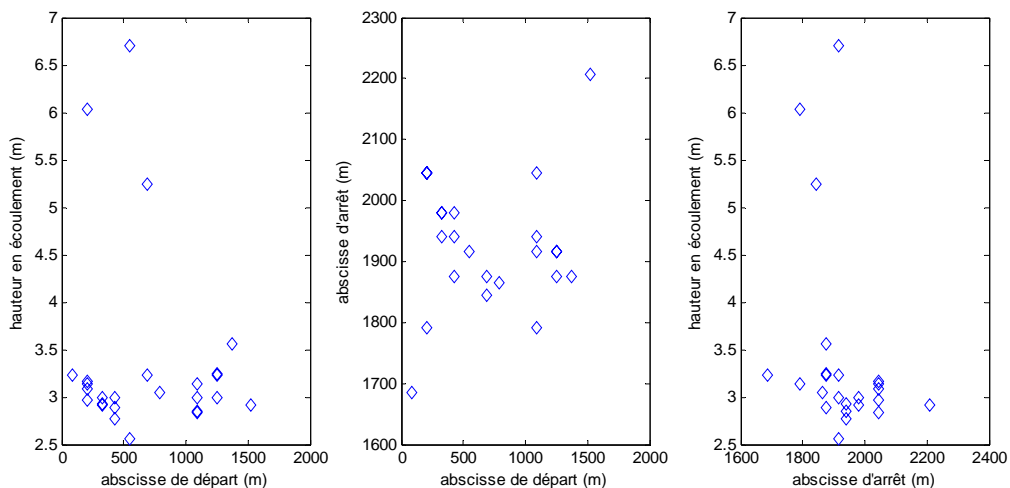
Les volumes de dépôt enregistrés s'échelonnent de 5400 à 750 000 m³. Les valeurs de hauteur de bloc obtenues s'échelonnent donc de 2.59 m à 6.6 m, avec une moyenne raisonnable de 3.37 m (Tableau 4.14). La dispersion des hauteurs en écoulement est principalement due à trois très gros événements présentant des hauteurs supérieures à 5 m alors que pour tous les 23 autres, les hauteurs sont comprises entre 2.5 m et 3.5 m (Figure 4.22).

Figure 4.21 : Cas d'étude pour les modèles de Voellmy bloc glissant et Saint Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Figure 4.22 : Scatter plot des données d'entrée du modèle de Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Bien que calculées uniquement sur 26 événements, les corrélations empiriques inter variables sont intéressantes (Tableau 4.15) et toutes significativement non nulles au seuil des 95%. Evidemment, la corrélation est extrêmement forte entre la hauteur et le volume de dépôt du fait du lien déterministe imposé entre les 2 variables ; le coefficient de corrélation n'est toutefois pas égal à 1 car le lien n'est

pas linéaire. Volume de dépôt et par conséquent hauteur en écoulement sont négativement corrélés à l'abscisse de départ, indiquant que, plus l'avalanche part de haut dans le couloir, plus la quantité de neige en écoulement est importante, résultat somme toute logique. De façon plus surprenante, la corrélation entre abscisse de départ et abscisse d'arrêt est positive tandis que celle entre hauteur de bloc et abscisse d'arrêt est faiblement négative. Les avalanches qui tendent à aller le plus loin sont donc, sur ce site, celles partant le plus bas et possédant les hauteurs en écoulement les plus faibles ! Un tel comportement est peut-être à relier d'un point de vue physique à un effet de l'altitude : à basse altitude les cumuls de précipitation sont moindres mais la neige est en moyenne plus humide de sorte que les avalanches peuvent malgré tout aller assez loin.

Tableau 4.15 : Corrélations empiriques entre variables du modèle de Voellmy bloc glissant

	x_{start}	v_{stop}	h	x_{stop}
x_{start}	1.00	-0.16	-0.13	0.07
v_{stop}	-0.16	1.00	0.98	-0.26
h	-0.13	0.98	1.00	-0.31
x_{stop}	0.07	-0.26	-0.31	1.00

4.2.4.2. Inférence

- *Lois a priori*

Des lois *a priori* assez proches de celles utilisées pour l'inférence des données simulées ont été employées. La principale différence concerne le coefficient de frottement ξ_{voel} pour lequel une distribution en moyenne plus élevée et légèrement plus dispersée a été préférée (Tableau 4.16).

Tableau 4.16 : Lois a priori utilisées avec le modèle Voellmy bloc glissant

$a_1 \sim U(0,10)$	$c \sim N(0.5, 0.2)$
$a_2 \sim U(0,10)$	$d \sim N(0, 0.25)$
$b_1 \sim N(3,1)$	$e \sim N(0, 0.25)$
$b_2 \sim N(0,1)$	$\sigma \sim Gamma(1,10)$
$\sigma_h^2 \sim Gamma(10,100)$	$\xi_{voel} \sim N(1200,100)$

- *Fonctionnement de l'algorithme*

Etant donné l'irrégularité topographique en zone d'arrêt, l'exigence en termes de calibration est un peu moins forte qu'avec le jeu de données simulées sur une topographie régulière. On fixe donc σ_{num} à 10 m, soit deux fois le pas de discrétisation du modèle de propagation.

Tableau 4.17 : Fonctions d'exploration pour le modèle Voellmy bloc glissant

$a_1(k) \sim N(a_1(k-1), 0.5)$	$c(k) \sim N(c(k-1), 0.05)$
$a_2(k) \sim N(a_2(k-1), 0.5)$	$d(k) \sim N(d(k-1), 0.1)$
$b_1(k) \sim N(a_1(k-1), 0.1)$	$e(k) \sim N(e(k-1), 0.015)$
$b_2(k) \sim N(b_2(k-1), 0.7)$	$\sigma(k) \sim \log N(\log(\sigma(k-1)), 0.5)$
$\sigma_h^2(k) \sim \log N(\log(\sigma_h^2(k-1)), 0.45)$	$\xi_{voel}(k) \sim \log N(\log(\xi_{voel}(k-1)), 0.004)$
	$\mu_{voel_i}(k) \sim N(\mu_{voel_i}(k-1), 0.0004)$

Tout comme pour les données simulées, les variances d'exploration ont été ajustées afin d'obtenir des taux d'acceptation satisfaisants (Tableau 4.17). Les différences entre les Tableaux 4.9 et 4.17 montrent que, dès lors que les données et la topographie changent, les variances d'exploration doivent être modifiées afin de conserver des taux d'acceptation optimaux (Tableau 4.18). Ce travail, bien que facilité par la construction séquentielle de l'algorithme, reste fastidieux et constitue de ce fait un frein important à l'utilisation systématisée de ce type d'algorithme pour la prédétermination.

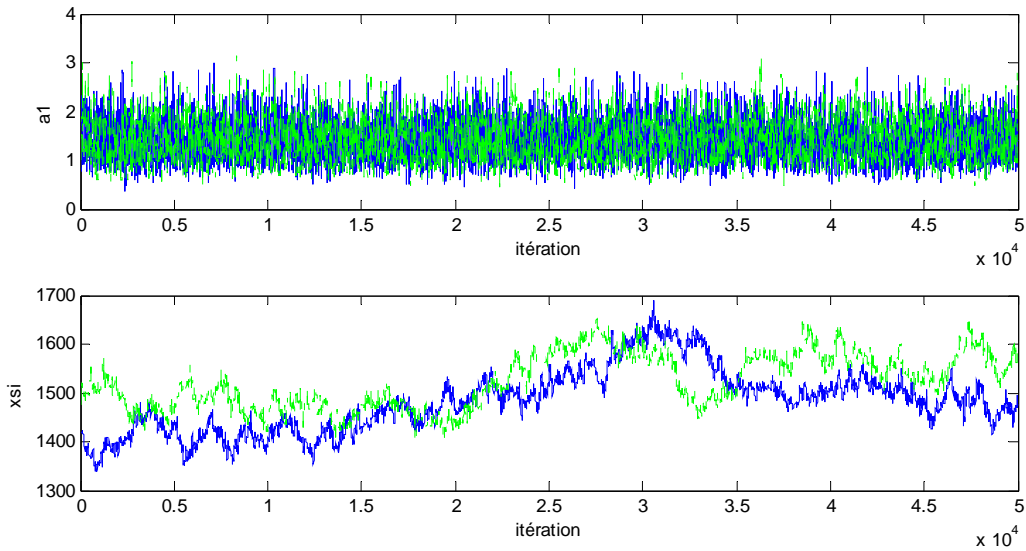
Comme pour les données simulées, 80 000 itérations de l'algorithme ont été effectuées avec deux chaînes démarrant en des points différents de l'espace des paramètres et les 30 000 premières ont été écartées pour la période de chauffe. Le temps de calcul nécessaire aux 80 000 itérations avec deux chaînes est de l'ordre du jour. C'est approximativement le même qu'avec les données simulées car la diminution du nombre d'événements à calibrer (26 contre 50) est compensée par l'augmentation importante de la taille du couloir et donc du temps nécessaire à la propagation de chaque avalanche.

Tableau 4.18 : Taux d'acceptation avec le modèle Voellmy bloc glissant

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel} et μ_{voel_i}
Chaîne 1	0.37	0.4	0.36	0.37	0.33	0.32	0.35	0.33	0.32	0.26
Chaîne 2	0.37	0.39	0.36	0.37	0.33	0.32	0.35	0.33	0.32	0.25

Les résultats graphiques (Figure 4.23) et les valeurs des variables de test de Gelman et Rubin (Tableau 4.19) montrent que, tout comme pour les données simulées, la convergence ne pose pas de problèmes pour tous les paramètres sauf pour ξ_{voel} . Pour ce dernier, la convergence est en effet encore plus difficile à atteindre sur ce couloir que pour les données simulées : les deux chaînes, bien qu'explorant la même gamme de valeurs, ont vraiment du mal à se mélanger et la statistique de test vaut encore 1.11 après 80 000 itérations. Pour la faire tendre vers 1 de très nombreuses itérations supplémentaires doivent être effectuées, de sorte qu'avec la valeur obtenue de 1.11, soit un écart en variance d'environ 10%, on supposera néanmoins la convergence atteinte par la suite.

Figure 4.23 : Valeurs successives de deux chaînes avec les vraies données



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 4.19 : Test de convergence de Gelman et Rubin avec le modèle Voellmy bloc glissant

a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}
1	1	1	1	1	1	1	1	1	1.11

- *Lois a posteriori des paramètres*

Les lois *a posteriori* obtenues sont présentées en Figure 4.24. Comme pour le modèle de Coulomb, elles présentent une forme unimodale bien marquée et une dispersion bien inférieure aux lois *a priori*. Cela indique que l'information contenue dans les données a bien été capturée par le modèle. Pour le paramètre de dispersion des hauteurs de bloc σ_h et pour le coefficient de frottement ξ_{voel} , la loi *a posteriori* marginale présente un mode bien distinct et significativement supérieur à celui de la loi *a priori*. Cela indique que la variabilité des hauteurs en écoulement autour de leur moyenne est plus forte que prévu. Cela indique également, si l'on en croit l'interprétation de la directive suisse (Salm et

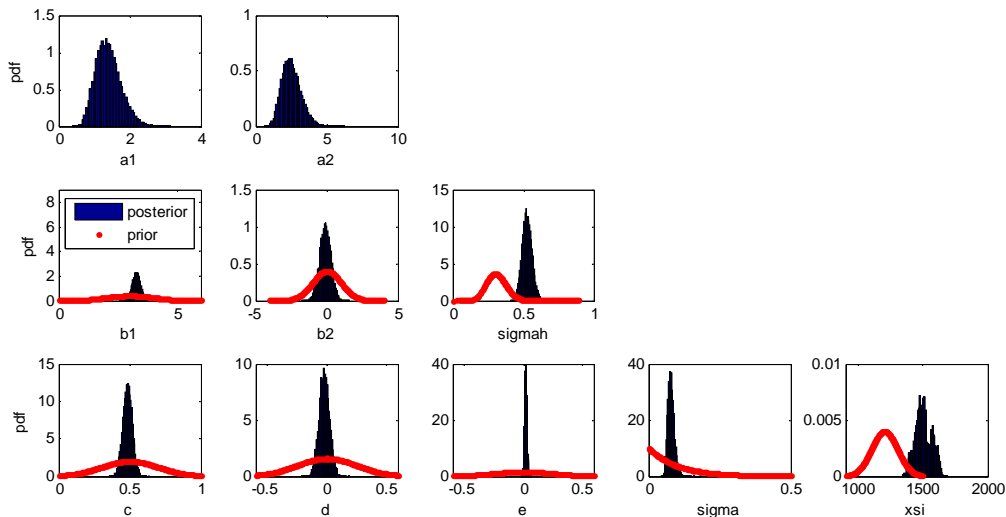
al., 1990), que le couloir présente une rugosité de surface plus faible que prévue. Pour les autres paramètres, on n'observe pas de différences très fortes entre les moyennes *a priori* et *a posteriori*, indiquant une bonne concordance entre données et expertise.

Le paramètre b_2 étant négatif mais pas significativement différent de zéro (Tableau 4.20), la corrélation entre hauteur de bloc et abscisse de départ est, conformément aux données, légèrement négative mais peu significative. Il en va de même pour le paramètre d représentant la corrélation entre le coefficient de frottement μ_{voel} et l'abscisse de départ x_{start} . Cela constitue une différence majeure avec le modèle de Coulomb, bruité ou non, puisque dans les deux cas d'étude traités avec une propagation coulombienne, une corrélation négative très significative avait été obtenue entre μ_{coul} et x_{start} . Enfin, la corrélation positive entre la hauteur en écoulement h et le coefficient de frottement μ_{voel} est significative puisque le paramètre e est significativement positif au seuil de 95%.

Tableau 4.20 : Statistiques descriptives des lois *a posteriori* pour, modèle de Voellmy bloc glissant

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}
Moy.	1.4	2.5	3.28	-0.11	0.53	0.49	-0.021	0.023	0.079	1503
SD	0.36	0.69	0.18	0.4	0.035	0.0032	0.044	0.01	0.012	63.5
q_{2.5}	0.8	1.35	2.94	-0.9	0.46	0.422	-0.107	0.003	0.06	1385
q_{97.5}	2.19	4	3.62	0.66	0.6	0.55	0.067	0.043	0.105	1623

Figure 4.24 : Lois *a priori* et *a posteriori* pour le modèle de Voellmy bloc glissant



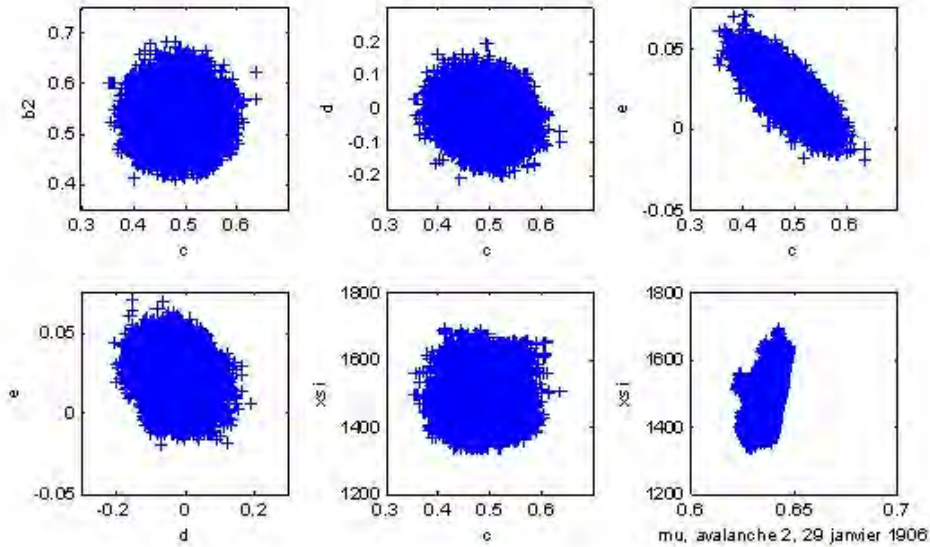
Couloir EPA 13 de la commune de Bessans, Savoie, France. Pour (a_1, a_2) les priors uniformes peu informatifs utilisés n'apparaissent pas sur la figure.

Les représentations de la loi *a posteriori* proposées jusqu'ici ont un peu fait oublier que le résultat de l'inférence bayésienne est la loi jointe *a posteriori* des inconnues. L'un des grands intérêts de l'inférence bayésienne est pourtant de pouvoir quantifier aisément la corrélation inter paramètres. La Figure 4.25 montre ainsi que certains couples d'inconnues ne sont pas corrélés du tout, tandis que d'autres le sont très fortement. Le coefficient de corrélation linéaire permet de quantifier la force du lien (Tableau 4.21) ; pour ce qui est des variables latentes, seuls les résultats concernant un événement, l'avalanche numéro 2 du jeu de données, sont présentés.

Etant donné que de très larges échantillons des lois *a posteriori* ont été simulés, la plupart des corrélations inter paramètres sont significatives au seuil des 95% (en rouge dans le Tableau 4.21), même si certaines d'entre elles sont très faibles. Les couples de paramètres décrivant la distribution d'une même variable sont en général fortement corrélés, la valeur de l'un pouvant assez facilement compenser celle de l'autre. Ainsi en va-t-il des couples (a_1, a_2) , (b_1, b_2) , (c, d) , (c, e) et (d, e) pour lesquels la valeur absolue du coefficient de corrélation linéaire est supérieure à 0.2. La

compensation explique également que les valeurs des coefficients de frottement latents μ_{voel_i} soient toutes corrélées avec le paramètre ξ_{voel} . Enfin, les paramètres indexant des distributions différentes tels a_1 et c ne sont pas du tout corrélés. Il en va de même des paramètres de variance, indépendants même des paramètres décrivant les moyennes des mêmes distributions.

Figure 4.25 : Corrélations a posteriori pour le modèle de Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 4.21 : Corrélations a posteriori des paramètres du modèle de Voellmy bloc glissant

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}	μ_{voel_2}
a_1	1.00	0.77	-0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.02
a_2	0.77	1.00	-0.01	0.01	-0.01	0.01	0.01	0.00	0.00	0.00	0.02
b_1	-0.01	-0.01	1.00	-0.81	0.04	-0.01	0.00	0.01	0.01	0.01	0.02
b_2	0.01	0.01	-0.81	1.00	0.01	0.02	0.00	-0.02	-0.01	-0.01	-0.02
σ_h	0.00	-0.01	0.04	0.01	1.00	0.00	-0.01	0.00	0.00	0.00	0.00
c	0.00	0.01	-0.01	0.02	0.00	1.00	-0.22	-0.80	0.07	0.01	-0.03
d	0.01	0.01	0.00	0.00	-0.01	-0.22	1.00	-0.22	0.07	-0.01	0.00
e	0.00	0.00	0.01	-0.02	0.00	-0.80	-0.22	1.00	-0.11	0.06	0.06
σ	0.00	0.00	0.01	-0.01	0.00	0.07	0.07	-0.11	1.00	-0.11	-0.05
ξ_{voel}	0.00	0.00	0.01	-0.01	0.00	0.01	-0.01	0.06	-0.11	1.00	0.44
μ_{voel_2}	0.02	0.02	0.02	-0.02	0.00	-0.03	0.00	0.06	-0.05	0.44	1.00

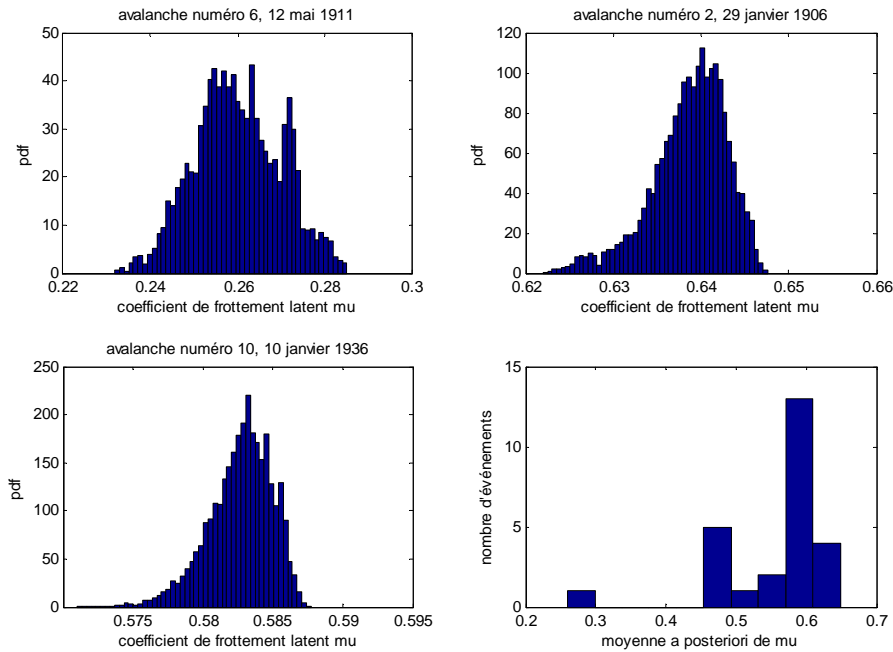
- Loi a posteriori des variables latentes

Les distributions obtenues pour les coefficients de frottement latents correspondant aux avalanches observées sont assez différentes les unes des autres. Pour l'avalanche du 12 mai 1911 correspondant à la plus forte distance d'arrêt observée, le coefficient de frottement estimé est de 0.26 ± 0.02 , avec une distribution de la variable latente unimodale, centrée sur son mode et symétrique (Figure 4.26, haut gauche). Pour les autres événements tels l'avalanche du 29 janvier 1906 (Figure 4.26, haut droite) ou celle du 10 janvier 1936 (Figure 4.26, bas gauche), des valeurs nettement plus élevées de coefficient de frottement sont obtenues. Les distributions sont toujours unimodales, mais parfois assez fortement asymétriques en fonction de la topographie au niveau de l'abscisse d'arrêt. Au final, les estimateurs ponctuels s'échelonnent entre 0.26 et 0.65, gamme de variation assez élevée mais néanmoins plausible par rapport aux valeurs usuellement recensées dans la littérature (Figure 4.26, bas droite).

Si l'on compare la distribution a posteriori du coefficient de frottement correspondant à un événement avec sa distribution théorique dans le modèle de mélange, c'est-à-dire une normale de

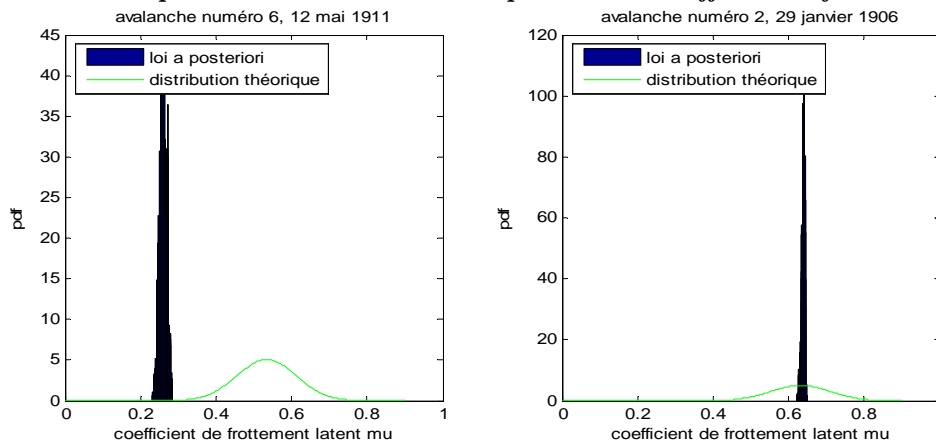
moyenne $E \left[\mu_{voel_i} \left| \hat{c}, \hat{d}, \hat{e}, x_{start_i}, \hat{h}_i \right. \right] = \hat{c} + \hat{d} x_{start_i} + \hat{e} \hat{h}_i$ et de variance $\hat{\sigma}^2$, on constate que la loi *a posteriori* estimée est bien moins dispersée pour tous les événements que la distribution de la variable latente prévue par le modèle (Figure 4.27). La raison en est que les paramètres de corrélation estimés sont faibles, de sorte que, quelle que soit l'abscisse de départ normalisée x_{start_i} et la hauteur de bloc h_i , la distribution théorique du coefficient de frottement latent varie peu. Lorsque l'événement considéré n'est pas un « événement moyen », le mode de la distribution *a posteriori* du coefficient de frottement ne correspond donc pas du tout à celui de la distribution théorique (Figure 4.27, gauche). Au contraire, pour un « événement moyen », les deux modes coïncident, mais avec toujours une dispersion bien moindre pour la loi *a posteriori* estimée (Figure 4.27, droite). Au final, il apparaît que si le modèle proposé semble bien capable de calibrer les coefficients de frottement associés aux événements observés, la calibration se fait essentiellement sur les distances d'arrêt. Les covariables modélisées sont donc peu prédictives, en tout cas beaucoup moins qu'avec le modèle de Coulomb.

Figure 4.26 : Loi *a posteriori* du coefficient de frottement latent du modèle Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Figure 4.27 : Comparaison entre modèle et loi *a posteriori* du coefficient de frottement latent

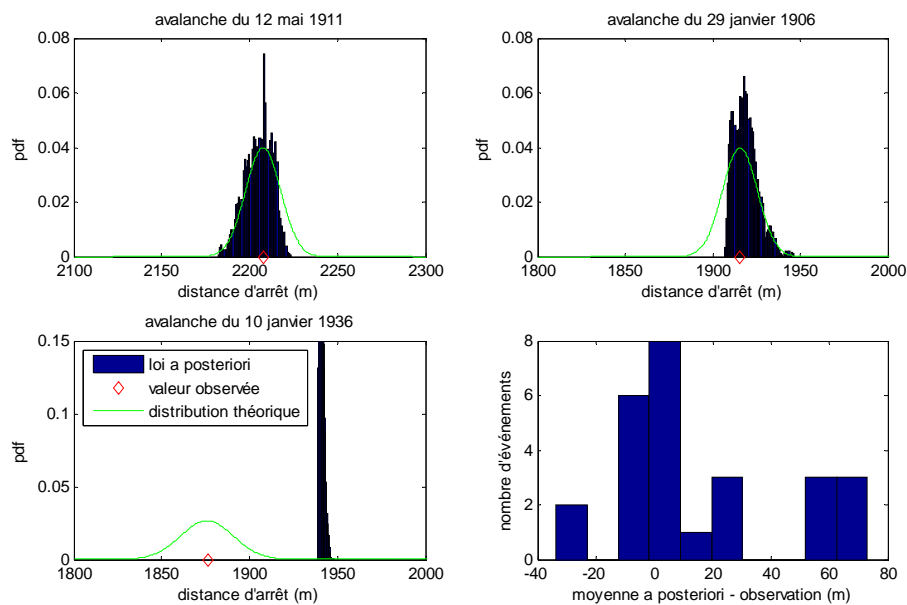


Couloir EPA 13 de la commune de Bessans, Savoie, France.

Pour la majorité des événements, la distance d'arrêt latente se distribue de façon conforme à la distribution théorique, c'est-à-dire selon une gaussienne centrée sur l'observation et de variance σ_{num}^2 . C'est par exemple le cas pour l'avalanche du 12 mai 1911 correspondant à la plus faible valeur calibrée du coefficient de frottement μ_{voel} (Figure 4.28, haut gauche). Pour un certain nombre d'événements, la distribution de la distance d'arrêt latente est toutefois un peu déformée par rapport à la distribution théorique, comme par exemple pour l'avalanche du 29 janvier 1906 (Figure 4.28, haut droite). Pour 6 événements, la distribution de la variable latente est très éloignée de la distribution théorique, comme par exemple pour l'avalanche du 10 janvier 1936 (Figure 4.28, bas gauche). Les abscisses d'arrêt de ces événements pour lesquels la calibration est difficile correspondent à la zone de ressaut en fin de couloir. Ainsi, même si le modèle de propagation utilisé est plus souple que le modèle de Coulomb, il n'est quand même pas capable de reproduire parfaitement toutes les distances d'arrêt observées, en particulier lorsqu'elles correspondent à des pentes moyennes succédant à des pentes faibles.

Au final, on voit néanmoins que la calibration des distances d'arrêt s'effectue de façon globalement satisfaisante puisque les erreurs de propagation moyennes sont relativement réduites. Si l'on excepte les 6 événements pour lesquels la calibration est difficile, on constate même que la distribution des erreurs de propagation moyennes est assez proche d'une gaussienne centrée sur zéro et d'écart type $\sigma_{num} = 10$ m (Figure 4.28, bas droite).

Figure 4.28 : Loi a posteriori de la distance d'arrêt latente avec le modèle Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

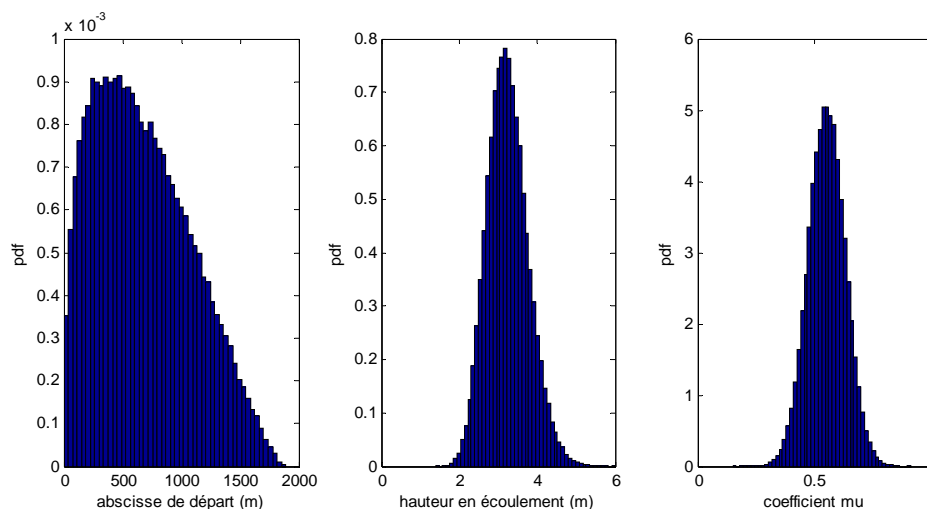
- *Moyenne a posteriori des variables d'entrée*

Tout comme pour le modèle de Coulomb, la moyenne a posteriori des variables d'entrée s'obtient à partir des estimateurs ponctuels des paramètres. Pour la hauteur en écoulement h et le coefficient de frottement μ_{voel} , la simulation des lois de mélange se fait par intégration sur la variabilité des variables explicatives, c'est-à-dire x_{start} uniquement pour h et x_{start} et h pour μ_{voel} (Figure 4.29). La simulation de la moyenne a posteriori des distances d'arrêt nécessitant en outre l'estimateur ponctuel de ξ_{voel} et la mise en œuvre de calculs de propagation, elle sera présentée au chapitre 6.

La moyenne a posteriori des abscisses de départ est assez conforme à celle obtenue lors de l'application du modèle de Coulomb bruité, ce qui est logique étant donné qu'une bonne partie des événements utilisés pour l'inférence du modèle de Voellmy l'avaient déjà été pour l'inférence sur le même site du modèle de Coulomb bruité. Pour la hauteur en écoulement h , la distribution est asymétrique conformément au modèle Gamma utilisé. Sa moyenne de 3.24 m est proche de celle des données (Tableau 4.22). En revanche, son écart-type est presque deux fois plus faible que l'écart-type

empirique, soit 0.53 m contre 0.98 m. Les 3 hauteurs les plus élevées « observées » ne sont donc pas dans l'intervalle de crédibilité à 95% du modèle. Ainsi, malgré l'utilisation d'une loi de mélange, le modèle proposé semble sous-estimer la variabilité des données et, en particulier, ne pas reproduire les événements les plus exceptionnels en termes de hauteur de bloc. Le constat fait pour x_{start} au chapitre 4.1.6 est donc généralisable : un travail important pourrait encore être effectué afin de proposer une représentation plus réaliste des variables d'entrée du ou des modèles de propagation.

Figure 4.29 : Distributions moyennes a posteriori des variables du modèle Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Pour le coefficient de frottement, la distribution obtenue, de moyenne 0.552, est très légèrement asymétrique mais beaucoup moins qu'avec le modèle de Coulomb bruité, ce qui traduit une corrélation moins significative avec x_{start} . Sa dispersion est par contre plus importante qu'avec le modèle de Coulomb bruité de façon à s'adapter à la variabilité des données. Néanmoins, la borne inférieure de l'intervalle de crédibilité à 95% est nettement supérieure à la plus faible valeur calibrée, c'est-à-dire 0.26 pour l'avalanche du 12 mai 1911. Deux explications sont possibles : soit cette dernière est réellement exceptionnelle, soit le modèle sous-estime aussi la variabilité des coefficients de frottement, ce qui d'un point de vue opérationnel est bien plus gênant que de sous-estimer la variabilité des hauteurs de bloc. On y reviendra au chapitre 6.

Tableau 4.22 : Distributions moyennes a posteriori des variables, modèle de Voellmy bloc glissant

	Moy.	SD	q _{2.5}	q _{97.5}
x_{start} (m)	679	411	62	1545
μ_{voel}	0.552	0.079	0.397	0.707
h (m)	3.24	0.53	2.29	4.36
x_{stop} (m)	1806.2	375.5	834.7	2109.9

Tableau 4.23 : Corrélations a posteriori entre les variables du modèle de Voellmy bloc glissant

	x_{start}	h	μ	x_{stop}
x_{start}	1.00	-0.04	-0.06	-0.06
h	-0.04	1.00	0.15	-0.04
μ_{voel}	-0.06	0.15	1.00	-0.70
x_{stop}	-0.06	-0.04	-0.70	1.00

Conformément aux données et aux paramètres de corrélation estimés, les corrélations inter variables sont assez faibles, légèrement négatives entre x_{start} et h de même qu'entre x_{start} et μ_{voel} et au contraire un peu plus nettement positives entre h et μ_{voel} (Tableau 4.23). Cela indique qu'utiliser des modèles de mélange plutôt que des modèles indépendants pour les variables d'entrée du modèle de Voellmy bloc glissant n'est pas décisif mais n'est tout de même pas totalement sans intérêt. Bien

entendu, la corrélation entre la distance d'arrêt simulée et μ_{voel} est quant à elle fortement négative du fait de la forte dépendance de la distance d'arrêt vis-à-vis du coefficient de frottement. Comme pour les paramètres, des échantillons très larges des variables du modèle sont simulés, de sorte que toutes les corrélations inter variables sont significativement non nulles au seuil des 95%.

4.2.5. Discussion

4.2.5.1. Interprétation des hyperparamètres c , d , e

L'obtention de paramètres de corrélation très faibles, voire non significatifs, entre le coefficient de frottement μ_{voel} et les variables d'entrée x_{start} et h est quelque peu surprenante. Une explication possible et plausible du point de vue de la physique est que cette corrélation n'existe pas sur le couloir étudié. Une autre explication, plus ennuyeuse, pourrait être qu'il est impossible d'inférer ces corrélations sans biais du fait des erreurs de propagation. Il est en effet bien connu en statistique que l'estimation de paramètres indexant la distribution de variables latentes n'est pas sans biais lorsque les observations sont bruitées. Kavetski et al. (2002) l'ont par exemple bien montré dans le cas de modèles hydrologiques à réservoirs, mais c'est également le cas avec un modèle de régression bruité très simple dès lors que le modèle d'erreur utilisé n'est pas le « vrai » modèle d'erreur.

Pour trancher entre les deux hypothèses, on a repris le jeu de données simulées mais on a introduit des « erreurs d'observation » non biaisées à l'aide d'un simple modèle de bruit blanc centré sur les valeurs simulées. On a ensuite refait tourner l'algorithme d'estimation comme décrit en 4.3.3. De façon très spectaculaire, les estimateurs des paramètres de corrélation d et e sont alors très proches de zéro et très éloignés des vraies valeurs utilisées pour la simulation avec des erreurs relatives de 4.7 et 3 écarts-types respectivement. L'estimateur du paramètre c est quant à lui également biaisé avec une erreur relative de 4.3 écarts-types car il a diminué pour compenser les variations des paramètres d et e . Bien entendu, les valeurs des coefficients de frottement latents μ_{voel_i} restent quant à eux quasi inchangés puisque la calibration se fait sans problème particulier sur les distances d'arrêt.

Il apparaît donc qu'un simple bruitage des observations est suffisant pour biaiser l'estimation des paramètres de corrélation et, en particulier, la rendre en apparence nulle alors qu'elle ne l'est pas. Comme de telles erreurs de propagation et/ou d'observation existent évidemment pour les données EPA et qu'il est illusoire de penser que l'on puisse introduire dans le modèle le vrai modèle d'erreur représentant fidèlement les différences entre le phénomène naturel et les valeurs enregistrées par l'observateur, il apparaît clairement que les paramètres de corrélation que l'on infère sur de vrais jeux de données sont en partie biaisés. Il convient par conséquent de les interpréter avec prudence et en particulier de ne pas nécessairement rejeter catégoriquement l'existence d'une corrélation entre μ_{voel} et les autres variables d'entrée même si elle n'est pas mise en évidence par l'inférence.

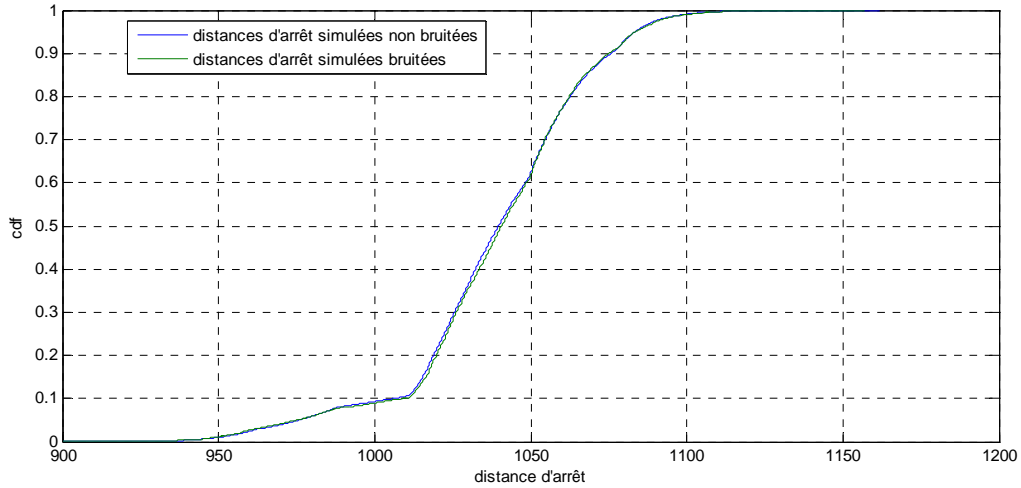
Tableau 4.24 : Statistiques descriptives des lois *a posteriori* avec les données simulées bruitées

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}
Moy.	1.04	2.13	2.98	-2.96	0.308	0.466	-0.13	-0.016	0.054	710
SD	0.19	0.43	0.07	0.16	0.028	0.031	0.036	0.012	0.006	45.8
q_{2.5}	0.71	1.39	2.83	-3.27	0.258	0.404	-0.2	-0.038	0.044	617
q_{97.5}	1.45	3.06	3.12	-2.63	0.367	0.527	-0.059	0.007	0.067	793
vraie valeur	1	2	3	-3	0.3	0.6	-0.3	-0.05	0.05	600
erreur	0.22	0.31	0.34	0.25	0.29	4.31	4.7	2.97	0.65	2.4

Afin de tester l'impact sur la prédiction de cette limite de la méthode d'inférence proposée, les calculs prédictifs détaillés au chapitre 6 ont été effectués sur le cas d'étude fictif avec les données simulées bruitées ou non. En termes de distribution moyenne *a posteriori*, le bruitage sans biais des distances d'arrêt observées n'introduit logiquement aucune différence (Figure 4.30). De façon plus surprenante étant donné que les calculs des distributions prédictives des aléas de référence font intervenir l'ensemble de la loi *a posteriori* des paramètres, les lois prédictives des distances d'arrêt de grandes périodes de retour ne semblent pas affectées non plus, ni en moyenne, ni même en variance (Tableau 4.25). Tout au plus, le fait de bruite les sorties semblerait conduire à des distances d'arrêt de projet très légèrement sous-estimées sans qu'une explication satisfaisante puisse être trouvée à ce résultat.

Au final, l'incapacité à inférer les paramètres de corrélation entre le coefficient μ_{voel} et les variables d'entrée x_{start} et h semble donc être davantage une limite à l'interprétation physique des résultats qu'à leur utilisation opérationnelle pour la caractérisation de l'aléa.

Figure 4.30 : Fonction de répartition des distances d'arrêt avec les données simulées



Topographie du couloir EPA 5 de la commune de Bessans, Savoie, France. 50 événements fictifs.

Tableau 4.25 : Distances d'arrêt de référence avec les données simulées bruitées ou non

Bruité	T ₁₀	T ₃₀	T ₁₀₀	T ₃₀₀	Non bruité	T ₁₀	T ₃₀	T ₁₀₀	T ₃₀₀
Moy.	1077.3	1088.8	1099.8	1109.7	Moy.	1077.9	1089.8	1101.8	1113.1
SD	5.2	5.7	7.8	10.5	SD	5.2	5.9	8.5	11.9
q_{2.5}	1066.4	1079.4	1086.7	1092.9	q_{2.5}	1067.2	1080.2	1088.2	1094.4
q_{97.5}	1087.1	1101.3	1118.1	1133.7	q_{97.5}	1087.7	1102.9	1121.7	1140.4

4.2.5.2. Identifiabilité et inférence sur le coefficient de frottement turbulent

Il est possible de calibrer une infinité de couples (μ_{voel}, ξ_{voel}) pour une distance d'arrêt donnée. Bien sûr, ce problème disparaît en théorie dès lors qu'un *prior* propre sur ξ_{voel} est utilisé. De plus, plus la quantité de données augmente et plus la loi *a posteriori* de ξ_{voel} est fortement contrainte de façon à recalculer de façon satisfaisante un maximum d'avalanches. Paradoxalement, c'est d'ailleurs d'autant plus vrai que la topographie du couloir est fortement tourmentée car la gamme de valeurs de ξ_{voel} permettant de s'approcher au mieux d'un maximum de distances d'arrêt est alors plus réduite que pour des couloirs très réguliers. Il n'en reste pas moins qu'en pratique l'utilisation d'un *prior* assez informatif sur ξ_{voel} est nécessaire pour faire converger l'algorithme à une vitesse raisonnable. La sensibilité à son choix se doit donc d'être investiguée.

Tableau 4.26 : Sensibilité au prior de la loi *a posteriori* de ξ_{voel}

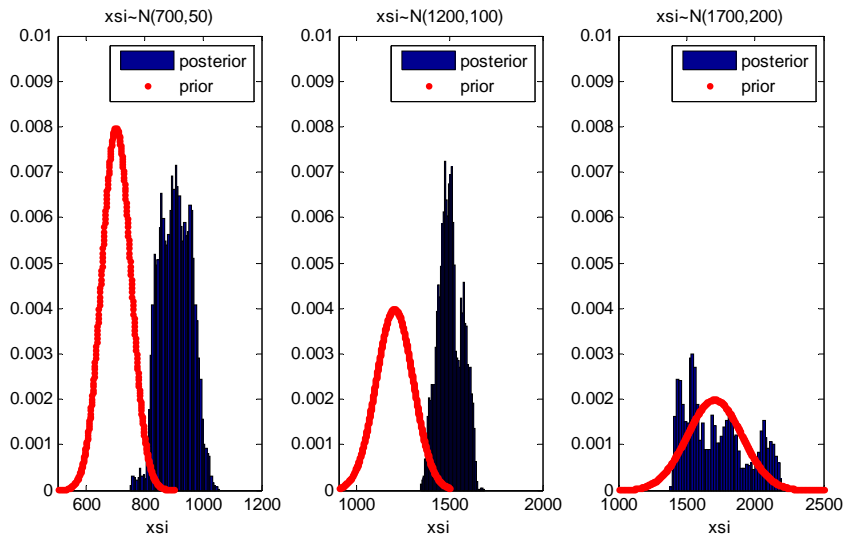
Prior	Moy.	SD	q _{2.5}	q _{97.5}
$\xi \sim N(700, 50)$	904.5	52.6	816	1001
$\xi \sim N(1200, 100)$	1503	63.5	1385	1623
$\xi \sim N(1700, 200)$	1722	226.4	1421	2151

L'inférence a donc également été effectuée sur le cas d'étude avec deux autres *priors*, l'un de moyenne et de variance plus faible, l'autre au contraire de moyenne et de variance plus forte. Par rapport aux calculs effectués avec le premier *prior*, seule la variance d'exploration sur ξ_{voel} a été modifiée, le reste des lois *a priori* et des réglages de l'algorithme restant inchangés.

Il apparaît que le *prior* a une influence forte sur le *posterior* obtenu : plus la variance du *prior* est forte, plus celle du *posterior* l'est (Tableau 4.26). De même, la moyenne du *prior* a tendance à tirer

celle du *posterior* vers le bas, sauf dans le cas du *prior* centré sur 1700 dont le mode semble correspondre à celui privilégié par les données. Notons toutefois que le *posterior* obtenu avec le troisième *prior* présente une forme peu satisfaisante. Même si la variable de test de Gelman et Rubin associée à ce paramètre est légèrement inférieure à 1.1, la convergence après 80 000 itérations n'est sans doute pas totalement atteinte. Cela indique que la loi *a priori* est vraisemblablement trop dispersée pour obtenir dans ce cas une convergence raisonnablement rapide (Figure 4.31).

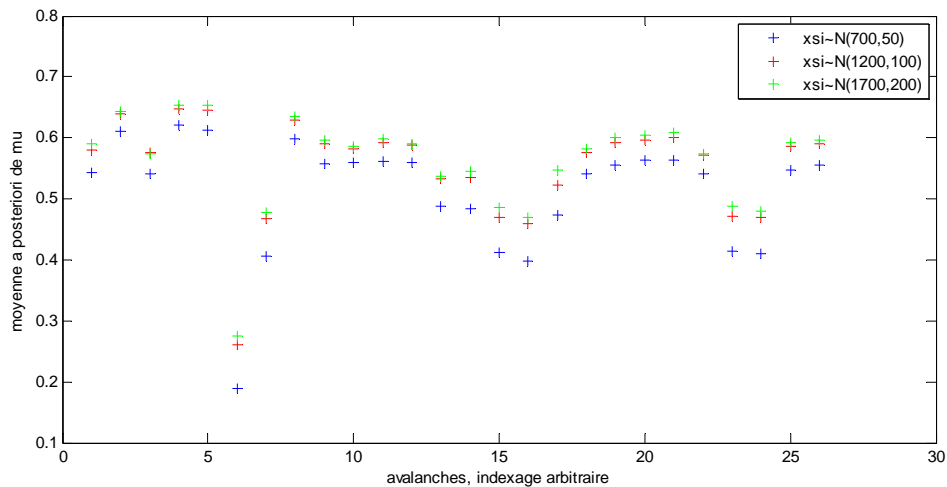
Figure 4.31 : Prior et posterior sur le paramètre ξ_{voel}



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant.

La compensation entre coefficients de frottement se visualise très bien si l'on compare les 26 valeurs calibrées des coefficients latents μ_{voel} avec les trois *priors* : lorsque ξ_{voel} augmente, toutes les valeurs de μ_{voel} augmentent également de façon à conserver la même distance d'arrêt (Figure 4.32). En revanche, l'augmentation des valeurs de μ_{voel} est plus forte lorsque ξ_{voel} passe de 700 à 1200 que lorsque ce dernier passe de 1200 à 1700 du fait de la non linéarité de la loi de frottement. On reviendra dans le chapitre 6 sur l'influence sur les valeurs de projet du choix du *prior* sur ξ_{voel} .

Figure 4.32 : Effet du prior sur ξ_{voel} sur les coefficients de frottement latents estimés



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant.

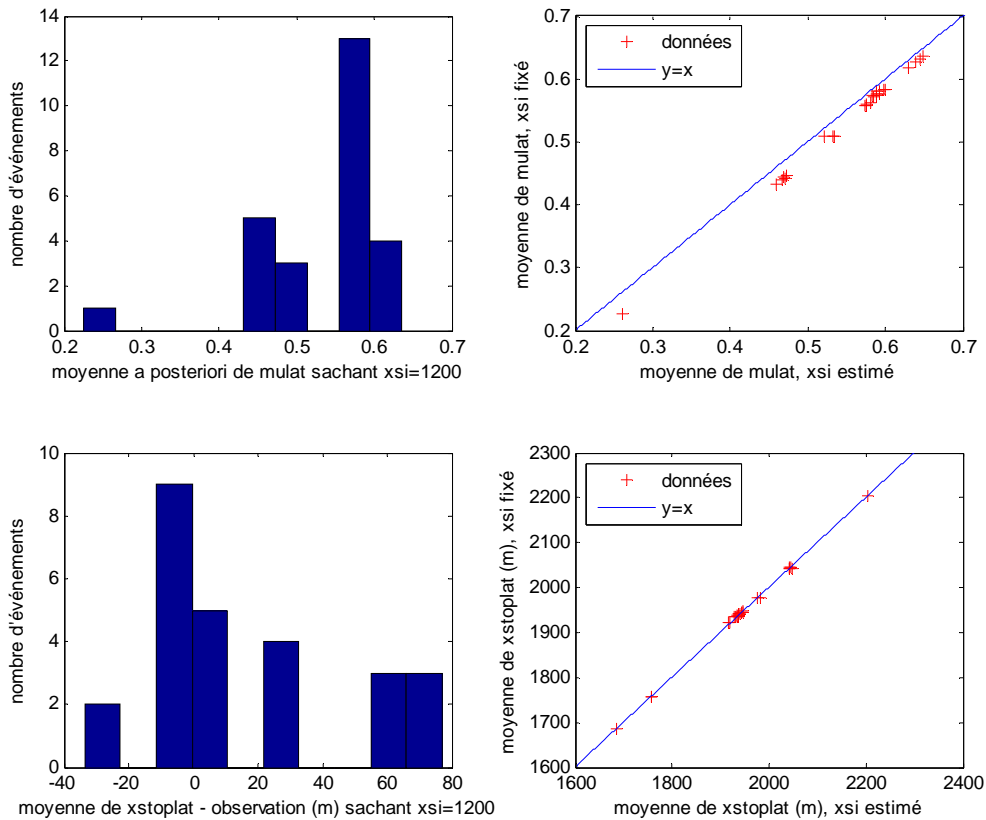
4.2.5.3. Et si on fixe ξ_{voel} ?

Puisque ξ_{voel} pose des problèmes de convergence et que, de toute façon, un *prior* informatif doit être utilisé, on a tenté de faire fonctionner l'algorithme en le fixant à sa moyenne *a priori*, c'est-à-dire 1200, le reste des lois *a priori* et des réglages de l'algorithme restant inchangé.

Le résultat s'avère très intéressant : 6000 itérations dont 1000 de chauffe se sont révélées largement suffisantes pour obtenir la convergence pour tous les paramètres et variables latentes, c'est-à-dire des valeurs de la variable de test de Gelman et Rubin inférieures à 1.1. Ainsi l'inférence du modèle n'a nécessité que deux heures. Les lois *a posteriori* des paramètres ainsi obtenues (Tableau 4.27) sont presque identiques à celles obtenues lorsque ξ_{voel} est estimé (Tableau 4.20). Quant aux distances d'arrêt calculées, on obtient exactement les mêmes estimateurs lorsque ξ_{voel} est fixé que lorsqu'il est estimé (Figure 4.33, bas droite), avec évidemment de ce fait la même répartition des erreurs de propagation (Figure 4.33, bas gauche et Figure 4.28, bas droite). La seule différence concerne les coefficients de frottement latents μ_{voel_i} qui sont significativement plus haut lorsque ξ_{voel} est fixé à 1200 que lorsqu'il est estimé à 1503 ± 120 (Figure 4.33, haut droite). Ainsi, l'histogramme des estimateurs des μ_{voel_i} est légèrement translaté vers les basses valeurs (Figure 4.33, haut gauche et Figure 4.26, bas droite). La raison en est bien sûr la compensation entre ξ_{voel} et μ_{voel} permettant de conserver les mêmes distances d'arrêt.

Fixer ξ_{voel} semble ainsi être une solution acceptable permettant d'assurer rapidement la convergence et de réduire considérablement le temps calcul. De ce fait, il pourrait s'agir d'une option envisageable lorsque l'on désire aller vite et que l'on dispose d'une forte confiance dans la valeur de ξ_{voel} qu'il convient d'employer. On y reviendra là aussi dans le chapitre 6 consacré à la prédiction.

Figure 4.33 : Variables latentes estimées avec ξ_{voel} fixé



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant.

Tableau 4.27 : Statistiques descriptives des lois a posteriori avec ξ_{voel} fixé

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ
Moy.	1.41	2.52	3.28	-0.1	0.53	0.48	-0.02	0.02	0.08
SD	0.37	0.68	0.17	0.38	0.03	0.03	0.05	0.01	0.01
q_{2.5}	0.81	1.41	2.93	-0.86	0.47	0.42	-0.13	0	0.06
q_{97.5}	2.22	4.05	3.59	0.67	0.6	0.55	0.07	0.04	0.11

4.3. Un modèle de type fluide

4.3.1. Modèle de propagation

Il s'agit maintenant de passer d'une description de type bloc glissant à une description de type fluide de l'avalanche. Le formalisme de Barré de Saint-Venant présenté au chapitre 3 est utilisé, de sorte que la hauteur de l'écoulement h et la vitesse v dépendent à présent de l'abscisse x et du temps t . On a donc $y = (x_{stop}, h_{xt}, v_{xt})$, mais on s'intéressera uniquement pour la caractérisation de l'aléa à h_x^{max} et v_x^{max} , c'est-à-dire à la hauteur et à la vitesse maximale de l'écoulement en chaque abscisse du profil. Une loi de frottement de type Voellmy est conservée avec un coefficient latent μ_{voel} et le coefficient ξ_{voel} considéré comme un paramètre.

En entrée, le modèle nécessite, outre les coefficients de frottement, la position et la dimension de la zone de départ. On note donc pour ce modèle x_{start} l'abscisse amont de la zone de départ sur le profil bidimensionnel. On note également $L_{plaque}, l_{plaque}, h_{plaque}$ les trois dimensions (longueur, largeur et hauteur) de la « plaque » de neige mise en mouvement. Quant à la pente moyenne de la zone de départ, elle est notée ϕ_{start} . Enfin, afin de limiter le temps calcul et la complexité du modèle stochastique, on ne prend toujours pas en compte les éventuels phénomènes d'érosion/dépôt souvent implémentés dans un modèle de propagation de type Saint-Venant (Naaïm et al., 2004).

La résolution des équations du mouvement se fait selon la méthode des éléments finis le long du profil topographique bidimensionnel selon un schéma proposé par Naaïm (1998) que l'on ne détaillera pas. Notons tout de même qu'à la différence des modèles de bloc glissant, l'emploi d'un modèle de type fluide pose le problème de la définition numérique de l'arrêt de l'avalanche. En effet, une fois la propagation terminée, des réarrangements du dépôt sont susceptibles de se produire pendant des durées très longues. Par conséquent, l'emploi du code de calcul nécessite la spécification d'un temps de propagation maximal ou d'un débit seuil en-deçà duquel la propagation est arrêtée. A noter également que du fait de la méthode de résolution utilisée, les distances d'arrêt sont, contrairement au modèle de bloc glissant, calculées au pas de 5 m.

4.3.2. Obtention des variables de départ à partir de l'EPA

Tout comme pour le modèle de bloc glissant, la principale difficulté de l'emploi de ce modèle avec les données EPA est que la variable de départ h_{plaque} traduisant la hauteur de la cassure n'y est pas renseignée, de même bien sûr que la longueur de plaque L_{plaque} .

On choisit d'obtenir les deux valeurs associées à chaque événement comme on a obtenu les hauteurs de bloc avec le modèle de bloc glissant, c'est-à-dire à l'aide d'une transformation déterministe de la seule variable quantitative disponible, i.e. le volume de dépôt. Pour ce faire, on suppose que les trois dimensions de la zone de départ ont une influence égale sur le volume de dépôt, soit une dépendance en puissance 1/3 des trois variables avec le volume de dépôt. Pour la hauteur et la longueur de plaque, on utilise comme pour la hauteur de bloc des considérations expertes concernant les hauteurs et longueurs de départ prévisibles des avalanches très petites et très grandes se produisant sur le couloir étudié. Cela qui permet de fixer les paramètres des deux lois puissance (eqs. 4.17 et 4.18). On en déduit une relation de dépendance linéaire entre h_{plaque} et L_{plaque} ainsi que les largeurs correspondantes sous hypothèse de conservation du volume et en assimilant la zone de départ à un tronçon de pente ϕ_{start} constante (eq. 4.19).

$$h_{plaque} = 0.5 + 0.04 \times v_{stop}^{1/3} \quad (\text{eq. 4.17})$$

$$L_{plaque} = 75 + 3.5 \times v_{stop}^{1/3} = 31.25 + 87.5 \times h_{plaque} \quad (\text{eq. 4.18})$$

$$l_{plaque} = \frac{v_{stop} \times \cos(\phi_{start})}{L_{plaque} \times h_{plaque}} \quad (\text{eq. 4.19})$$

La propagation bidimensionnelle ne tient pas compte des largeurs. Tout se passe pour chaque avalanche comme si la largeur de départ était identique. On peut donc calculer un volume de départ équivalent $v_{starteq}$ pour chaque événement à partir de la moyenne sur tous les événements de la largeur de départ (eq. 4.20). Sa variabilité représente mieux la variabilité introduite en entrée du modèle de propagation que celle des volumes de dépôt observés.

$$v_{starteq} = \frac{L_{plaque} \times h_{plaque}}{\cos(\phi_{start})} \times \frac{1}{N} \sum_{i=1}^N l_{plaque_i} \quad (\text{eq. 4.20})$$

On choisit d'utiliser le volume équivalent pour définir la condition d'arrêt de l'avalanche : on considère que l'avalanche est arrêtée lorsque le débit maximal propagé est inférieur à 0.01% du volume équivalent de l'avalanche (eq. 4.21). L'intérêt d'utiliser un débit seuil non fixe est de ne pas risquer d'arrêter une très petite avalanche trop tôt et, au contraire, de ne pas prolonger de façon irréaliste la propagation des très grandes avalanches.

$$q_{seuil} = v_{starteq} \times 10^{-4} \quad (\text{eq. 4.21})$$

4.3.3. Modèle stochastique

La longueur de plaque intervient en tant que variable d'entrée dans le modèle de propagation mais, du fait de la méthode de calcul choisie, c'est une fonction déterministe de h_{plaque} . On peut donc se contenter de modéliser la hauteur de fracture et se ramener ainsi à un modèle stochastique quasi identique à celui correspondant au modèle de bloc glissant. La représentation sous forme de DAG reste inchangée par rapport au modèle précédent (Figure 4.16) bien que l'interprétation de la variable h soit différente : hauteur de bloc avec le modèle de bloc glissant et hauteur de fracture avec le modèle de Saint-Venant. Le « même » algorithme d'estimation peut donc être employé.

4.3.4. Cas d'étude

Le cas d'étude choisi est le même que pour le modèle de Voellmy bloc glissant, c'est-à-dire les 26 avalanches sur le couloir EPA 13 de Bessans. Les données sont donc quasi identiques à celles décrites en 4.2.4.1, sauf pour la hauteur qui n'a pas la même signification (Tableau 4.28). Les hauteurs de fracture obtenues varient entre 75 cm et 4 m ce qui semble acceptable (Figure 4.34, gauche). Les longueurs de plaque varient quant à elles grossièrement de 100 m à 400 m (Figure 4.34, centre). Les largeurs de plaques qui en résultent présentent une importante variabilité résiduelle. Cette variabilité des entrées n'est pas prise en compte par le modèle de propagation puisqu'il ne fait pas la différence entre des largeurs de départ faibles et fortes. En conséquence, la variabilité du volume équivalent est bien inférieure à celle des volumes de dépôt observés (Figure 4.34, droite).

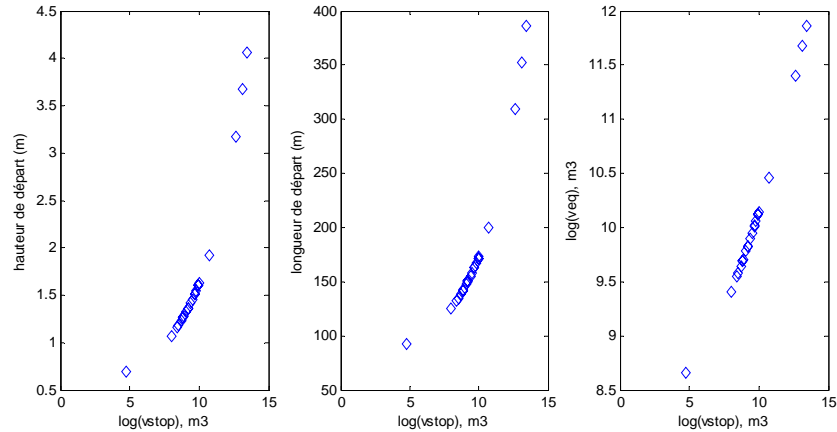
Tableau 4.28 : Données pour l'application du modèle de Saint-Venant

	Moy.	SD	q _{2.5}	q _{97.5}
x_{start} (m)	677.6	443.5	106.5	1496.9
v_{stop} (m ³)	69 226	169 590	5400	674 760
h_{plaque} (m)	1.64	0.78	0.75	4
L_{plaque} (m)	174.7	68.3	96.9	381.5
l_{plaque} (m)	72.3	88.4	3.8	350.9
$v_{starteq}$ (m ³)	30 587	32 898	6740	138 540
x_{stop} (m)	1933.1	104.8	1701.3	2183.5

Les corrélations inter variables (Tableau 4.29) sont cohérentes avec celles observées pour le modèle de bloc glissant ainsi qu'avec les choix de modélisation effectués. Bien que certaines d'entre elles soient

faibles, elles sont toutes significativement non nulles au seuil des 95%. Les corrélations entre les dimensions de départ $L_{plaque}, l_{plaque}, h_{plaque}$ et l'abscisse de départ x_{start} sont légèrement négatives. Elles sont toutes trois similaires à la valeur obtenue entre l'abscisse de départ et la hauteur de bloc avec le modèle de bloc glissant. Il en va de même pour les corrélations plus nettement négatives entre les dimensions de départ et l'abscisse d'arrêt x_{stop} . Les nombreux coefficients de corrélation égaux à 1 ou très proches de 1 traduisent les relations déterministes inter variables.

Figure 4.34 : Scatter plot des données d'entrée du modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 4.29 : Corrélations empiriques entre variables du modèle de Saint-Venant

	x_{start}	v_{stop}	h_{plaque}	L_{plaque}	l_{plaque}	$v_{start_{eq}}$	x_{stop}
x_{start}	1.00	-0.16	-0.12	-0.12	-0.13	-0.14	0.07
v_{stop}	-0.16	1.00	0.95	0.95	0.98	0.99	-0.26
h_{plaque}	-0.12	0.95	1.00	1.00	0.99	0.99	-0.33
L_{plaque}	-0.12	0.95	1.00	1.00	0.99	0.99	-0.33
l_{plaque}	-0.13	0.98	0.99	0.99	1.00	1.00	-0.32
$v_{start_{eq}}$	-0.14	0.99	0.99	0.99	1.00	1.00	-0.30
x_{stop}	0.07	-0.26	-0.33	-0.33	-0.32	-0.30	1.00

4.3.5. Inférence

4.3.5.1. Fonctionnement de l'algorithme

Bien que le passage d'un modèle de bloc glissant à un modèle fluide n'ait guère modifié le modèle, la comparaison des Tableaux 4.30 et 4.17 montre que les variances d'exploration ont à nouveau dû être ajustées afin d'obtenir des taux d'acceptation satisfaisants. Les valeurs optimales des variances d'exploration dépendent donc du modèle de propagation autant que du modèle stochastique, en tout cas pour certains paramètres. Cela illustre les difficultés pratiques liées à l'implémentation d'un algorithme de Metropolis Hastings.

La modification du réglage par rapport au modèle de bloc glissant a en particulier été nécessaire pour les coefficients de frottement ξ_{voel} et μ_{voel_i} . Cependant, ce travail n'a pas été inutile puisque les variances d'exploration ont pu être considérablement augmentées. L'intérêt pratique d'une variance d'exploration plus forte est de permettre une exploration plus efficace de l'espace des paramètres et donc d'assurer la convergence vers la loi cible beaucoup plus rapidement. Ainsi 3000 itérations dont 1000 de chauffe ont été suffisantes pour obtenir avec deux chaînes des valeurs de test de Gelman et Rubin correctes, c'est-à-dire inférieures à 1.1 pour toutes les inconnues. Les taux d'acceptation obtenus ne sont pas parfaits puisqu'un peu élevés pour certains paramètres scalaires (Tableau 4.31). Ils

ont néanmoins été estimés satisfaisants. Les ajuster plus précisément est difficile pour ce modèle du fait du temps important nécessaire pour effectuer quelques dizaines d'itérations.

Tableau 4.30 : Fonctions d'exploration avec le modèle de Saint-Venant

$a_1(k) \sim N(a_1(k-1), 0.5)$	$c(k) \sim N(c(k-1), 0.04)$
$a_2(k) \sim N(a_2(k-1), 0.6)$	$d(k) \sim N(d(k-1), 0.1)$
$b_1(k) \sim N(a_1(k-1), 0.15)$	$e(k) \sim N(e(k-1), 0.03)$
$b_2(k) \sim N(b_2(k-1), 0.4)$	$\sigma(k) \sim \log N(\log(\sigma(k-1)), 0.4)$
$\sigma_h^2(k) \sim \log N(\log(\sigma_h^2(k-1)), 0.4)$	$\xi_{voel}(k) \sim \log N(\log(\xi_{voel}(k-1)), 0.01)$
	$\mu_{voel_i}(k) \sim N(\mu_{voel_i}(k-1), 0.001)$

Tableau 4.31 : Taux d'acceptation avec le modèle de Saint-Venant

a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel} et μ_{voel_i}
0.37	0.34	0.4	0.48	0.41	0.49	0.41	0.41	0.38	0.26

Le temps de calcul reste avec ce modèle un réel problème. Effectuer 3000 itérations, c'est-à-dire 78 000 itérations du modèle de propagation avec 26 données, a ainsi nécessité près de 10 jours de calcul sur un PC récent. La forte diminution par rapport au modèle de bloc glissant du nombre d'itérations nécessaires à la convergence rend toutefois le calcul possible, bien que lourd. Cette diminution est en partie explicable par la tolérance d'une erreur numérique un peu plus forte. Afin de limiter un peu le temps de calcul, σ_{num} a en effet été cette fois fixée à 15 m, soit trois fois le pas de discrétisation du code numérique de propagation, contre 10 m avec le modèle de bloc glissant. Toutefois, la baisse du nombre d'itérations nécessaires reflète également une moindre sensibilité du modèle aux très faibles variations de coefficients de frottement. Avec le modèle de bloc glissant, une très faible variation des coefficients de frottement peut en effet se traduire, toutes choses étant égales par ailleurs, par une variation assez importante de la distance d'arrêt. L'acceptation des candidats est alors difficile et une très faible variance d'exploration doit être utilisée pour les coefficients de frottement. Pour le modèle de Saint-Venant, la sensibilité aux faibles variations de coefficients de frottement en termes de distance d'arrêt est un peu moins forte car une compensation s'effectue au cours de l'écoulement au travers de la répartition de la masse dans l'espace et dans le temps. On peut donc utiliser une variance d'exploration plus forte permettant de réduire le nombre d'itérations.

4.3.5.2. Lois a posteriori des paramètres

Les mêmes *priors* que pour le modèle de bloc glissant ont été utilisés. Les lois *a posteriori* obtenues (Figure 4.35) sont moins « lisses » qu'avec le modèle de bloc glissant car beaucoup moins d'itérations ont été effectuées. Néanmoins les formes des lois *a posteriori* marginales sont satisfaisantes : unimodales avec des modes assez bien marqués et des dispersions globalement bien inférieure à celles des lois *a priori*.

Tableau 4.32 : Statistiques descriptives des lois a posteriori pour le modèle de Saint-Venant

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}
Moy.	1.38	2.49	1.52	0.03	0.45	0.449	-0.013	0.025	0.11	1300
SD	0.35	0.7	0.17	0.39	0.037	0.028	0.05	0.017	0.02	53.6
q_{2.5}	0.8	1.29	1.18	-0.74	0.386	0.39	-0.12	-0.001	0.08	1200
q_{97.5}	2.13	4.05	1.84	0.8	0.535	0.5	0.08	0.06	0.15	1387

Les lois obtenues sont naturellement identiques à celles obtenues avec le modèle de bloc glissant pour les paramètres a_1 et a_2 puisque ni les données ni le modèle n'ont changé pour les abscisses de départ. Les lois *a posteriori* sont par contre assez différentes pour les paramètres du modèle de mélange de la variable h dont le terme constant b_1 et la variance σ_h sont plus faibles qu'avec le modèle de bloc glissant. Cela traduit simplement une moyenne et une variance plus faibles des hauteurs de fracture

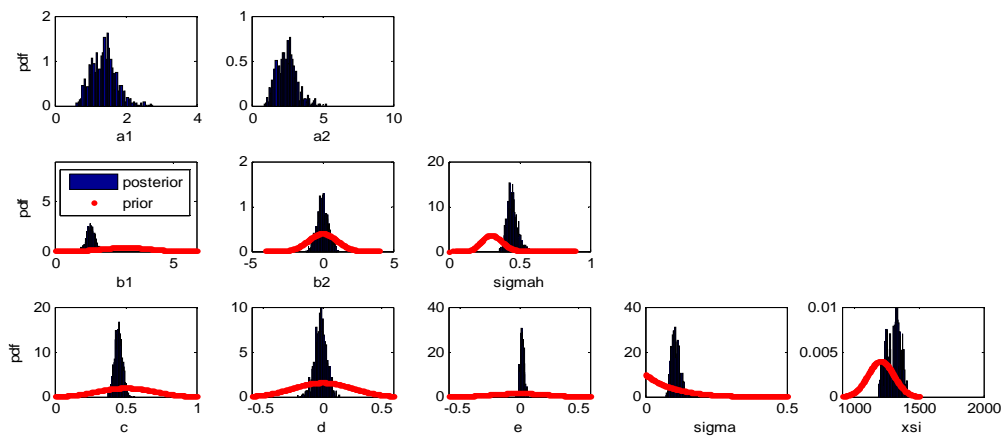
que des hauteurs de bloc (Tableau 4.32). Quant au paramètre de corrélation b_2 , il est encore plus faible qu'avec le modèle de bloc glissant, indiquant que les hauteurs de fracture sont quasiment indépendantes des abscisses de départ.

Pour le modèle de mélange du coefficient de frottement, les conclusions sont assez similaires sous réserve que les estimateurs obtenus sont vraisemblablement biaisés comme discuté en 4.2.5.1. Les

paramètres de corrélation \hat{d} et \hat{e} sont plus faibles qu'avec le modèle de bloc glissant et, à la différence du modèle de bloc glissant, aucun des deux n'est significatif au seuil des 95%. Au contraire, la variance de μ_{voel_i} autour de sa moyenne (mesurée par $\hat{\sigma}$) est plus importante qu'avec le modèle de bloc glissant, de sorte que les variables d'entrée semblent encore moins bien expliquer la variabilité du coefficient de frottement qu'avec le modèle de bloc glissant. Au final, aucun des paramètres de corrélation des deux modèles de mélange n'est donc significatif au seuil des 95%. Les variables d'entrée semblent donc être presque indépendantes sur le cas d'étude choisi.

Enfin, la loi *a posteriori* de ξ_{voel} est plus proche du *prior* qu'avec le modèle de bloc glissant, avec une moyenne plus faible de 1300 contre 1500 et une variance *a posteriori* similaire (Tableau 4.32). Ce résultat va à l'encontre de l'opinion de certains modélisateurs estimant qu'il convient d'utiliser des valeurs plus élevées de ξ_{voel} avec un modèle fluide qu'avec un modèle de bloc glissant. Cette assertion est toutefois généralement basée sur des considérations liées à la résolution des équations du mouvement plutôt que sur l'adéquation aux données. Le résultat obtenu montre donc qu'il convient d'être prudent à ce sujet.

Figure 4.35 : Lois a priori et a posteriori pour le modèle de Saint-Venant



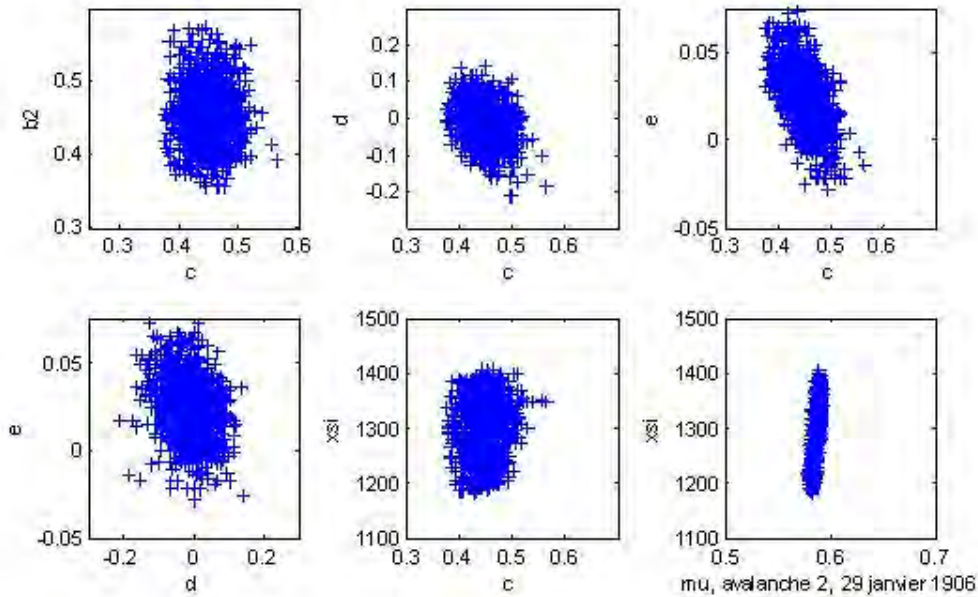
Couloir EPA 13 de la commune de Bessans, Savoie, France. Pour a_1 et a_2 , des priors non informatifs uniformes ont été utilisés.

Tableau 4.33 : Corrélation a posteriori des paramètres du modèle de Saint-Venant

	a_1	a_2	b_1	b_2	σ_h	c	d	e	σ	ξ_{voel}	μ_{voel_2}
a_1	1.00	0.79	0.08	-0.10	-0.01	-0.01	0.00	0.00	-0.05	-0.09	-0.09
a_2	0.79	1.00	0.07	-0.08	-0.02	-0.01	-0.03	0.02	-0.03	-0.11	-0.07
b_1	0.08	0.07	1.00	-0.85	0.07	0.03	-0.05	-0.04	0.01	-0.11	-0.05
b_2	-0.10	-0.08	-0.85	1.00	0.04	-0.05	0.04	0.05	-0.03	0.12	0.04
σ_h	-0.01	-0.02	0.07	0.04	1.00	0.06	-0.03	-0.04	0.04	-0.05	0.01
c	-0.01	-0.01	0.03	-0.05	0.06	1.00	-0.31	-0.57	0.28	0.16	0.08
d	0.00	-0.03	-0.05	0.04	-0.03	-0.31	1.00	-0.28	-0.11	-0.09	-0.08
e	0.00	0.02	-0.04	0.05	-0.04	-0.57	-0.28	1.00	-0.13	0.00	0.06
σ	-0.05	-0.03	0.01	-0.03	0.04	0.28	-0.11	-0.13	1.00	-0.03	-0.02
ξ_{voel}	-0.09	-0.11	-0.11	0.12	-0.05	0.16	-0.09	0.00	-0.03	1.00	0.66
μ_{voel_2}	-0.09	-0.07	-0.05	0.04	0.01	0.08	-0.08	0.06	-0.02	0.66	1.00

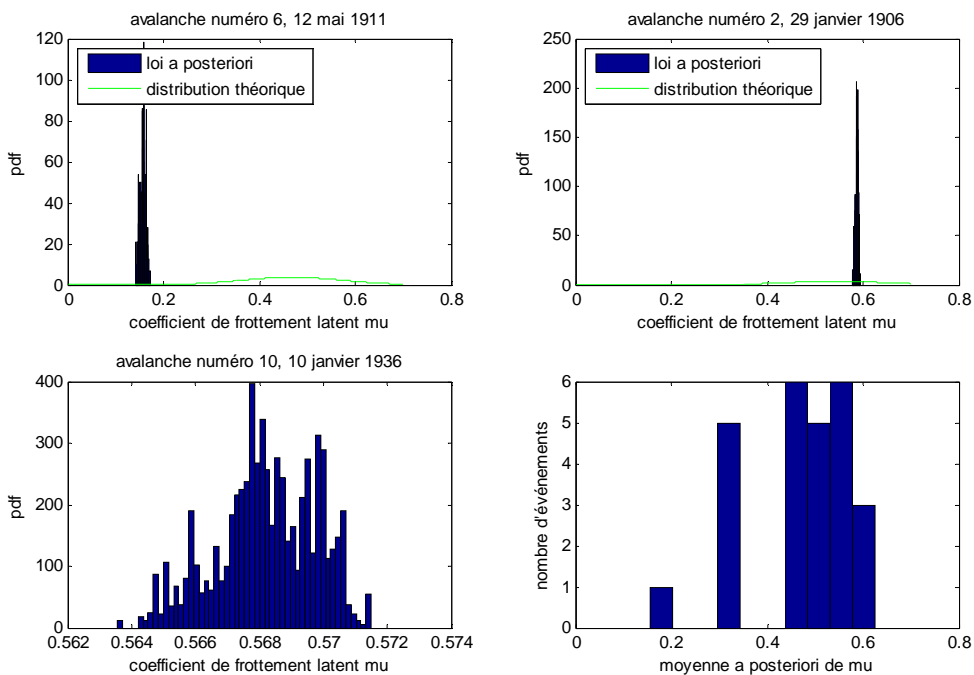
La Figure 4.36 présente les mêmes lois jointes *a posteriori* que la Figure 4.25 tandis que le Tableau 4.33 propose, tout comme le Tableau 4.21, les coefficients de corrélation linéaire inter paramètres. La quasi-totalité des corrélations sont significativement non nulles au seuil des 95% du fait de la taille des échantillons simulés. De plus, du fait des possibles compensations, les corrélations sont significatives entre paramètres des modèles de mélange et entre le paramètre ξ_{voel} et les valeurs de μ_{voel} correspondant aux différents événements. Enfin, pour les autres paramètres, l'indépendance *a posteriori* est manifeste. Des corrélations *a posteriori* inter paramètres assez analogues à celles présentées pour le modèle de bloc glissant sont donc obtenues avec le modèle de Saint-Venant.

Figure 4.36 : Corrélations *a posteriori* pour le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Figure 4.37 : Loi *a posteriori* du coefficient de frottement latent du modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

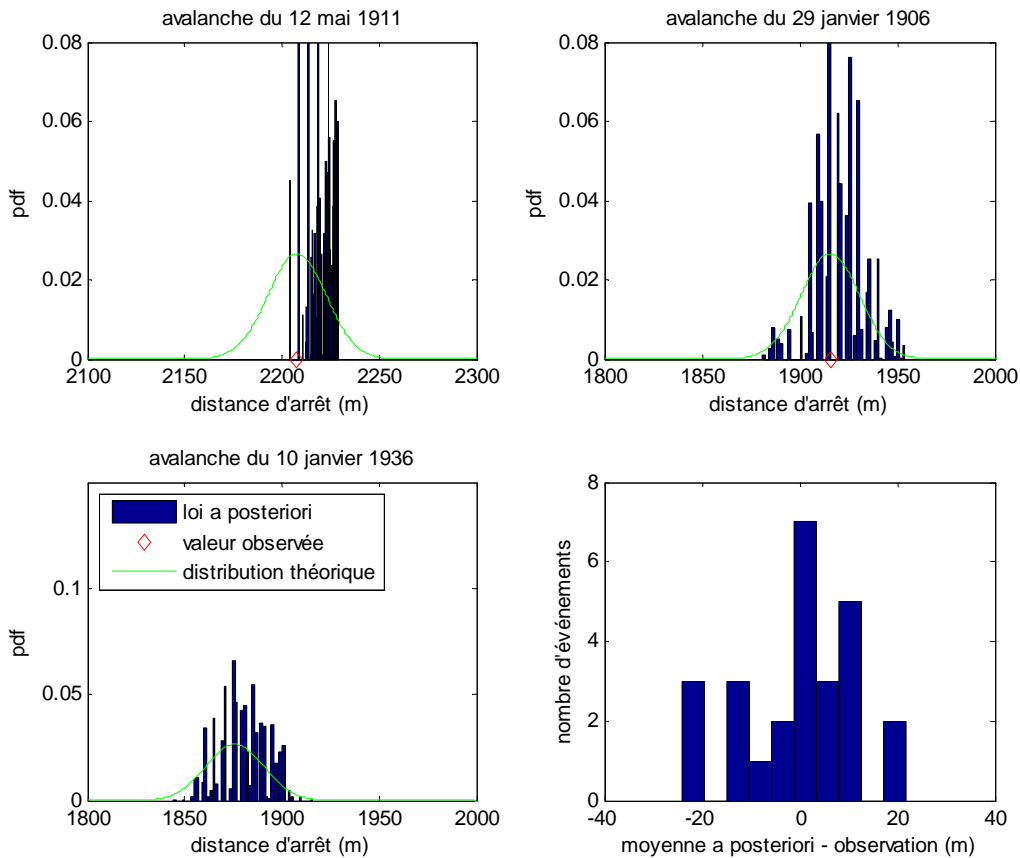
4.3.5.3. *Lois a posteriori des variables latentes*

Pour chaque événement, la distribution du coefficient de frottement latent présente une forme en cloche et une très faible dispersion (Figure 4.37, bas gauche) assez analogue à celle obtenue avec le modèle de bloc glissant. La forme de l'histogramme des estimateurs ponctuels obtenus (Figure 4.37, bas droite) est également analogue à celle obtenue avec le modèle de bloc glissant. Néanmoins, pour un événement donné, le coefficient de frottement estimé est significativement plus faible avec le modèle de Saint-Venant qu'avec le modèle de Voellmy. Ainsi, pour l'avalanche de distance d'arrêt exceptionnelle du 12 mai 1911, une valeur $\hat{\mu}_{voel} = 0.16$ est obtenue contre 0.26 avec le modèle de bloc glissant. Une partie de cet écart compense la diminution de $\hat{\xi}_{voel}$. Néanmoins, la valeur de 0.16 étant plus proche de celle proposée par la littérature pour les avalanches exceptionnelles que celle de 0.26, le remplacement d'un modèle de bloc glissant par un modèle fluide semble également aller dans le sens d'un meilleur réalisme physique de la description. On y reviendra au chapitre 6.

Enfin, l'adéquation entre la distribution *a posteriori* de chaque coefficient μ_{voel_i} et la distribution théorique du modèle de mélange, c'est-à-dire la normale de moyenne $E\left[\mu_{voel_i} \mid \hat{c}, \hat{d}, \hat{e}, x_{start_i}, h_i\right] = \hat{c} + \hat{d} x_{start_i} + \hat{e} h_i$ et de variance $\hat{\sigma}^2$, est encore moins forte qu'avec le modèle de bloc glissant. Du fait des très faibles valeurs moyennes des paramètres d et e , la distribution moyenne de μ_{voel_i} est en effet quasiment identique quelles que soient les valeurs de h_i et x_{start_i} .

Toute la variabilité de la distribution du coefficient de frottement est donc traduite par l'écart type σ . De plus, les lois *a posteriori* de chaque événement peuvent de ce fait correspondre au mode de la distribution théorique moyenne (Figure 4.37, haut droite) ou au contraire en être très éloignées (Figure 4.37, haut gauche).

Figure 4.38 : Distribution *a posteriori* des distances d'arrêt latentes du modèle de Saint-Venant

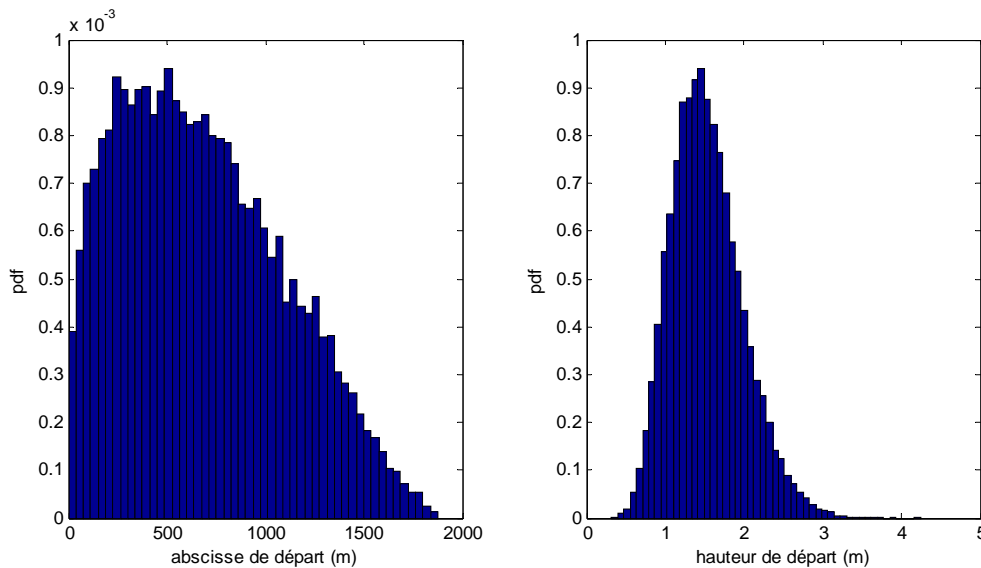


Couloir EPA 13 de la commune de Bessans, Savoie, France.

En termes de calibration des distances d'arrêt, des résultats très satisfaisants sont observés. Bien sûr, la distribution des distances d'arrêt de chaque événement est discrète du fait de la résolution numérique au pas de 5 m. Néanmoins, pour la plupart des événements, la répartition des distances d'arrêt calculées autour de la valeur observée se fait de manière assez conforme à la distribution théorique, c'est-à-dire une normale centrée sur l'observation et d'écart-type $\sigma_{num} = 15$ m. Ainsi en va-t-il de l'avalanche du 29 janvier 1906 (Figure 4.38, haut droite). C'est en particulier vrai pour certaines des avalanches que le modèle de bloc glissant n'arrivait pas à reproduire, comme celle du 10 janvier 1936 (Figure 4.38, bas gauche). Cependant, pour quelques événements comme celui du 12 mai 1911, la distribution des distances d'arrêt calculées est un peu éloignée de la distribution théorique, ce qui indique que la calibration n'est quand même pas parfaite (Figure 4.38, haut gauche).

Au final, on obtient des erreurs de propagation moyennes généralement faibles mais s'échelonnant globalement de -20 m à +20 m (Figure 4.38, bas droite). Il est ainsi remarquable de constater que tous les 6 événements qui étaient vraiment mal recalculés avec le modèle de bloc glissant le sont beaucoup mieux avec le modèle de Saint-Venant. Ce dernier est donc bien mieux capable de reproduire les distances d'arrêt correspondant à la zone de ressaut en milieu de zone d'arrêt. Cela s'explique par le fait qu'il offre à l'écoulement la possibilité de s'adapter au changement de pente en modifiant hauteur et répartition de la masse le long du profil topographique. La représentation plus réaliste de l'écoulement comme un phénomène continu dans l'espace et dans le temps plutôt que comme le déplacement d'un solide ponctuel offre donc un intérêt indiscutable en termes de caractérisation de l'aléa.

Figure 4.39 : Distributions moyennes des variables d'entrée du modèle de Saint-Venant (1)



Couloir EPA 13 de la commune de Bessans, Savoie, France.

4.3.5.4. Moyenne *a posteriori* des variables du modèle

Le calcul des moyennes *a posteriori* des variables du modèle de Saint-Venant s'effectue, comme pour les autres modèles, à l'aide des estimateurs ponctuels. Les distributions des hauteurs (Figure 4.39) et longueurs (Figure 4.40) de départ présentent des formes très voisines puisque l'une est une fonction linéaire de l'autre. Leurs formes et dispersions (Tableau 4.34) correspondent assez bien à celles des données présentées dans le Tableau 4.28. Les distributions estimées semblent en particulier moins sous estimer la variabilité des données qu'avec le modèle de bloc glissant.

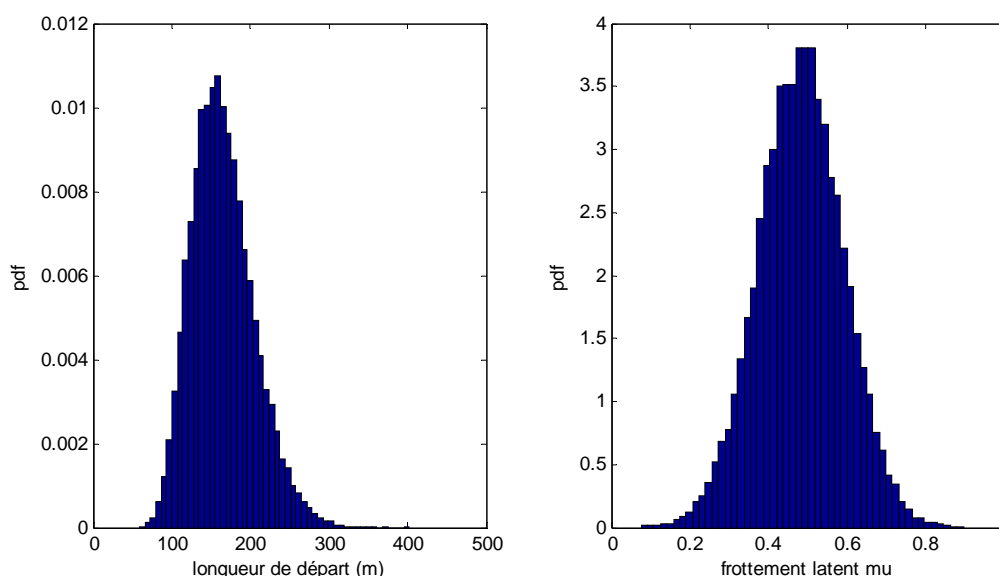
Pour le coefficient de frottement latent, le principal point notable est le caractère quasi symétrique de la distribution moyenne *a posteriori*. On est très proche d'une gaussienne de moyenne \hat{c} et de variance $\hat{\sigma}^2$, ce qui traduit la très faible corrélation avec les autres variables du modèle. La corrélation *a posteriori* entre μ_{voel} et x_{start} ainsi qu'entre μ_{voel} et les dimensions de départ h_{plaque}, L_{plaque} est donc très faible (Tableau 4.35). Les autres coefficients de corrélation *a posteriori*

traduisent assez fidèlement les données, en particulier la faible dépendance des dimensions de départ h_{plaque}, L_{plaque} avec x_{start} . Malgré tout, du fait encore une fois de la taille des échantillons simulés, toutes les corrélations a posteriori inter variables sont significativement non nulles au seuil des 95%

Tableau 4.34 : Distributions moyennes a posteriori des variables du modèle de Saint-Venant

	Moy.	SD	q _{2.5}	q _{97.5}
x_{start} (m)	681.7	412.9	58.8	1550.3
h_{plaque} (m)	1.53	0.45	0.77	2.54
μ_{voel}	0.482	0.109	0.268	0.695
x_{stop} (m)	1805.8	343.3	889	2137.3
L_{plaque} (m)	165.2	39.8	98.7	253
$v_{starteq}$ (m ³)	24568	13472	6890	58262

Figure 4.40 : Distributions moyennes des variables d'entrée du modèle de Saint-Venant (2)



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 4.35 : Corrélation a posteriori pour les variables du modèle de Saint-Venant

	x_{start}	h_{plaque}	μ_{voel}	x_{stop}	L_{plaque}	$v_{starteq}$
x_{start}	1.00	0.02	-0.03	0.23	0.02	0.02
h_{plaque}	0.02	1.00	0.12	0.13	1.00	0.98
μ_{voel}	-0.03	0.12	1.00	-0.79	0.12	0.11
x_{stop}	0.23	0.13	-0.79	1.00	0.13	0.12
L_{plaque}	0.02	1.00	0.12	0.13	1.00	0.98
$v_{starteq}$	0.02	0.98	0.11	0.12	0.98	1.00

4.4. Conclusion du chapitre 4

- *Trois modèles d'intensité*

Dans ce chapitre consacré au modèle d'intensité, différentes hypothèses simplificatrices concernant la nature des variables modélisées, le modèle de propagation, les opérateurs stochastiques décrivant la variabilité du phénomène ainsi que les *priors* ont permis d'utiliser le formalisme général proposé au

chapitre 3 pour mettre en œuvre l'inférence bayésienne de trois modèles de propagation d'avalanche différents sur des cas concrets.

Un modèle très simplifié inversible a tout d'abord été appliqué sur un site très régulier. Un avatar de ce modèle avec un terme d'erreur et une vitesse initiale non nulle a également été proposé pour traiter, au prix d'un réalisme physique très faible, le cas d'un site EPA à la géométrie raisonnablement complexe. Deux modèles non inversibles proposant une description de plus en plus réaliste de l'écoulement avalancheux ont ensuite été proposés et inférés.

Pour les trois modèles proposés, l'essentiel de l'effort de modélisation a porté sur le coefficient de frottement latent. Les variables d'entrée ont au contraire été modélisées de façon assez fruste, et ce d'autant plus que ce sont des transformations des vraies observations qui ont été utilisées. Les éventuelles corrélations existant entre les différentes variables ont en revanche toujours été explicitement prises en compte sous la forme de lois de mélange.

- *Des algorithmes d'estimation adaptés mais délicats*

Des algorithmes de calibration adaptés de type MCMC ont été développés pour les trois modèles. Pour le modèle de Coulomb et son avatar bruité, l'inférence ne pose pas de problèmes car le modèle est inversible. Pour les deux autres modèles, inférence et propagation ne peuvent être dissociées et des algorithmes de Metropolis-Hastings adaptés ont dû être construits. L'utilisation d'un cas d'étude fictif a permis d'en vérifier le bon fonctionnement. Relativement simples en principe, ils sont en pratique délicats à manier du fait de la nécessité d'ajuster les fonctions d'exploration au cas d'étude traité. De plus, les temps de calcul pour l'inférence sont relativement longs, en particulier pour le modèle fluide. Le nombre d'itérations nécessaires à la convergence peut toutefois éventuellement être nettement réduit en fixant le coefficient de frottement turbulent.

- *Résultats d'inférence*

Pour les trois modèles, les lois *a posteriori* des différents paramètres ont pu être obtenues, donnant accès à des estimateurs ponctuels, aux incertitudes associées et aux coefficients de corrélation inter paramètres. Les phénomènes de compensation entre certains couples de paramètres et entre coefficients de frottement ont pu être établis.

Les corrélations inter variables se sont révélées fortement significatives avec le modèle de Coulomb et son avatar bruité. Avec les deux autres modèles, au contraire, elles ont été de plus en plus faibles au fur et à mesure que le modèle de propagation se complexifiait. Cependant, il est apparu que les estimateurs des paramètres de corrélation pouvaient être biaisés par les erreurs de discrétisation numérique et d'observation, sans toutefois affecter significativement le calcul des aléas de référence.

L'adéquation entre données d'entrée et modèles s'est révélée satisfaisante pour les trois modèles, en particulier pour le coefficient de frottement. Pour les deux modèles non inversibles, la distribution des variables latentes inférée sur les cas d'études s'est cependant révélée être bien plus resserrée que la distribution moyenne *a posteriori*. Celle-ci correspond alors à un comportement moyen du fait de la faiblesse des corrélations inter variables.

Avec le modèle de Saint-Venant, au-delà des problèmes de temps calcul, des résultats particulièrement satisfaisants ont été obtenus : valeurs de coefficient de frottement correspondant bien à celles données par la littérature pour les avalanches extrêmes ainsi que diminution par rapport au modèle de bloc glissant du nombre d'itérations nécessaires à la convergence du fait d'une moindre sensibilité à de faibles changements des coefficients de frottement.

- *Et maintenant...*

Les résultats de ce chapitre devront être complétés au chapitre 6 par les calculs de propagation permettant d'obtenir la distribution des variables de sortie. Ils devront également être consolidés sous la forme d'analyses de sensibilité, notamment vis-à-vis des structures de modèles et de l'information *a priori* utilisées. Cependant, quels que soient les résultats qui seront obtenus, il restera une importante marge de progrès pour la construction de modèles d'intensité offrant une représentation plus réaliste des zones de départ d'avalanches à partir des données de l'EPA. Inclure davantage de connaissance experte sous la forme de lois *a priori* plus informatives que celles utilisées jusqu'ici restera également une perspective future de développement. Il convient avant cela d'effectuer un travail similaire sur le modèle fréquentiel de façon à disposer de tous les éléments permettant le calcul de périodes de retour.

5. Construction et estimation du modèle fréquentiel

Préambule

Ce chapitre est consacré au modèle fréquentiel. Les occurrences avalanches sont modélisées par un modèle poissonien très simple. Comme pour le modèle d'intensité, seule l'information locale est tout d'abord utilisée pour traiter les mêmes cas d'étude qu'au chapitre 4. Ensuite, conformément au cadre décrit au chapitre 3.5, l'estimation des fréquences avalanches est effectuée à l'aide d'un modèle spatial puis spatio-temporel régional de plus en plus sophistiqué. Enfin, le meilleur modèle est appliqué à plusieurs départements des Alpes françaises de façon à y visualiser la structure spatio-temporelle des occurrences avalanches.

5.1. Un modèle poissonien local

5.1.1. Modèle et inférence

On s'intéresse à présent aux nombres d'avalanches se produisant chaque année sur le site étudié au cours des T_{obs} hivers d'observation. L'occurrence d'une avalanche étant un phénomène discret et rare, le modèle le plus naturel est un modèle poissonien dont l'unique paramètre λ représente à la fois la moyenne et l'écart-type de la distribution des nombres d'avalanches hivernaux (eq. 5.1).

L'inférence bayésienne d'un modèle poissonien est très aisée à condition de choisir une loi *a priori* de type Gamma pour λ (eq. 5.2). Les propriétés de conjugaison permettent en effet d'obtenir, pour une fois analytiquement, la loi *a posteriori*, celle-ci étant toujours de type Gamma. De plus, si l'on nomme (a_λ, b_λ) les deux paramètres de la loi *a priori* de λ , les deux paramètres (a'_λ, b'_λ) de la loi *a posteriori* s'expriment simplement en fonction de (a_λ, b_λ) , du nombre total P d'avalanches et de la durée d'observation T_{obs} selon $a'_\lambda = a_\lambda + T_{obs}$ et $b'_\lambda = b_\lambda + P$.

Pour des valeurs assez larges de T_{obs} et P , l'information *a priori* n'a que peu d'influence sur le *posterior*, tout particulièrement si des valeurs faibles sont choisies pour les paramètres (a_λ, b_λ) . λ tend alors vers P/T_{obs} c'est-à-dire la fréquence avalancheuse empirique moyenne. Cela traduit la convergence de l'estimateur bayésien vers l'estimateur classique du maximum de vraisemblance.

$$l(a_t | \lambda) = \frac{\lambda^{a_t}}{a_t!} \times \exp(-\lambda) \quad (\text{eq. 5.1})$$

$$p(\lambda | a_\lambda, b_\lambda) = \frac{a_\lambda^{b_\lambda}}{\Gamma(b_\lambda)} \times \lambda^{b_\lambda - 1} \times \exp(-a_\lambda \times \lambda) \quad (\text{eq. 5.2})$$

5.1.2. Application aux cas d'étude

5.1.2.1. Données

Les cas d'étude traités sont les mêmes qu'au chapitre 4, c'est-à-dire les couloirs EPA 5 et 13 de Bessans. Pour ces deux couloirs, la fréquence avalancheuse est estimée sur la période 1960-2003 où la collecte des données est supposée régulière et exhaustive. Les Figures 5.1 et 5.2 montrent en effet qu'il ne semble n'y avoir aucun « trou » (période de plusieurs années sans observation), ni de tendance, ni de rupture marquée. Des tests statistiques standards de type corrélation de rang de Spearman n'infirmes pas ces hypothèses.

Sur les deux couloirs, de 0 à 3 avalanches ont été enregistrées chaque hiver (Tableau 5.1). Le couloir 5 est toutefois un peu plus actif que le couloir 13, avec une fréquence avalancheuse empirique moyenne de 33% supérieure. Pour les deux couloirs, l'écart-type empirique est assez comparable à la moyenne empirique, ce qui semble indiquer qu'un modèle poissonien est effectivement acceptable (Tableau 5.1).

Tableau 5.1 : Activité avalancheuse observée sur les deux sites étudiés

	P	Période	T _{obs}	Moy.	SD	min	max
Bessans 5	53	1960-2003	44	1.2	1	0	3
Bessans 13	40	1960-2003	44	0.91	0.86	0	3

Figure 5.1 : Nombres d'avalanches observées sur le couloir EPA 5 de Bessans

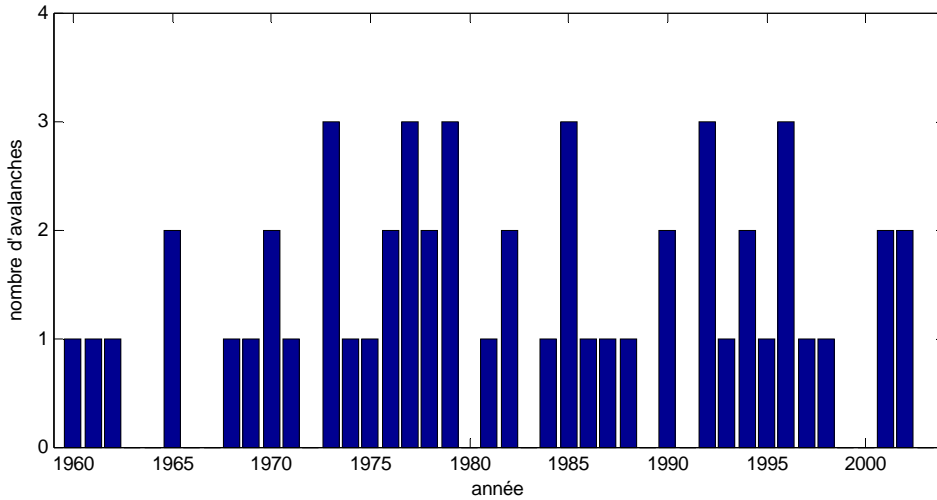
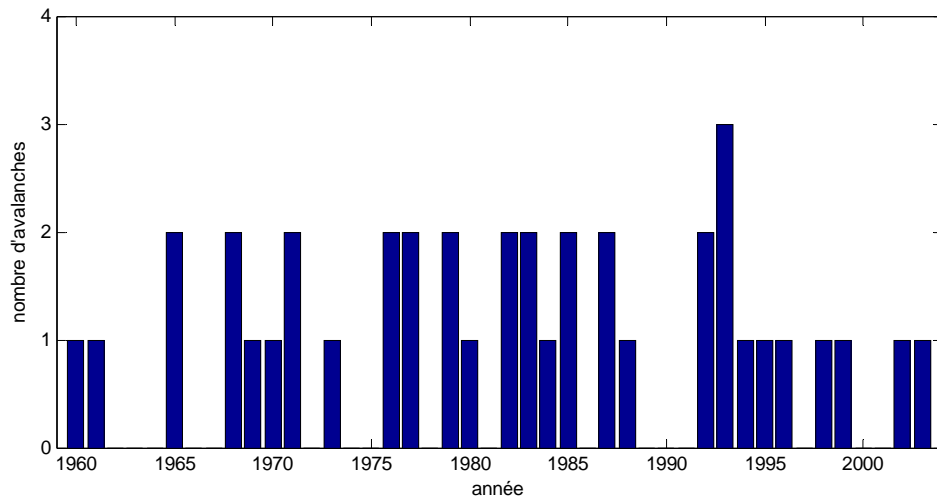


Figure 5.2 : Nombres d'avalanches observées sur le couloir EPA 13 de Bessans



5.1.2.2. Inférence

L'inférence des deux cas d'étude est effectuée avec un *prior* très peu informatif (eq 5.3). Les estimateurs obtenus sont proches des fréquences avalancheuses empiriques moyennes, avec une dispersion modérée reflétant la quantité déjà assez importante de données (Tableau 5.2).

$$\lambda \sim \text{Gamma}(0.01, 0.001) \text{ (eq 5.3)}$$

Tableau 5.2 : Lois a posteriori sur les deux sites étudiés

	Moy.	SD	q _{2.5}	q _{97.5}
Bessans 5	1.2	0.166	0.91	1.56
Bessans 13	0.93	0.15	0.67	1.24

Les Figures 5.3 et 5.4 présentent, pour chaque couloir, un échantillon simulé de la loi *a posteriori* de λ ainsi qu'une comparaison entre les fonctions de répartition des données et de la moyenne *a posteriori* du modèle. L'échantillon simulé de λ sera utilisé au chapitre 6 pour le calcul numérique des distributions des quantiles annuels de distance d'arrêt. Quant aux comparaisons entre données et modèles, elles confirment la relative bonne adaptation entre modèle et variabilité des données sur les

deux sites. Pour le couloir 5 de Bessans, l'accord est excellent pour les probabilités de non dépassement faibles et la seule différence notable entre données et modèle est que ce dernier prévoit une probabilité non nulle d'observer plus de 3 avalanches par hiver. Pour le couloir 13 de Bessans, l'accord est un peu moins bon mais reste tout de même très satisfaisant étant donné la simplicité du modèle utilisé.

Figure 5.3 : Inférence sur le couloir EPA 5 de Bessans

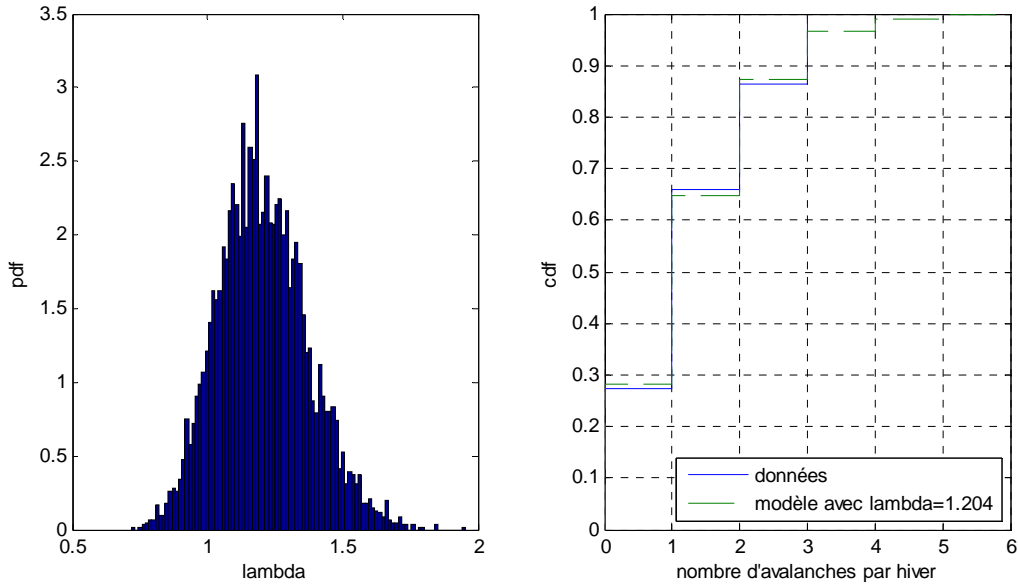
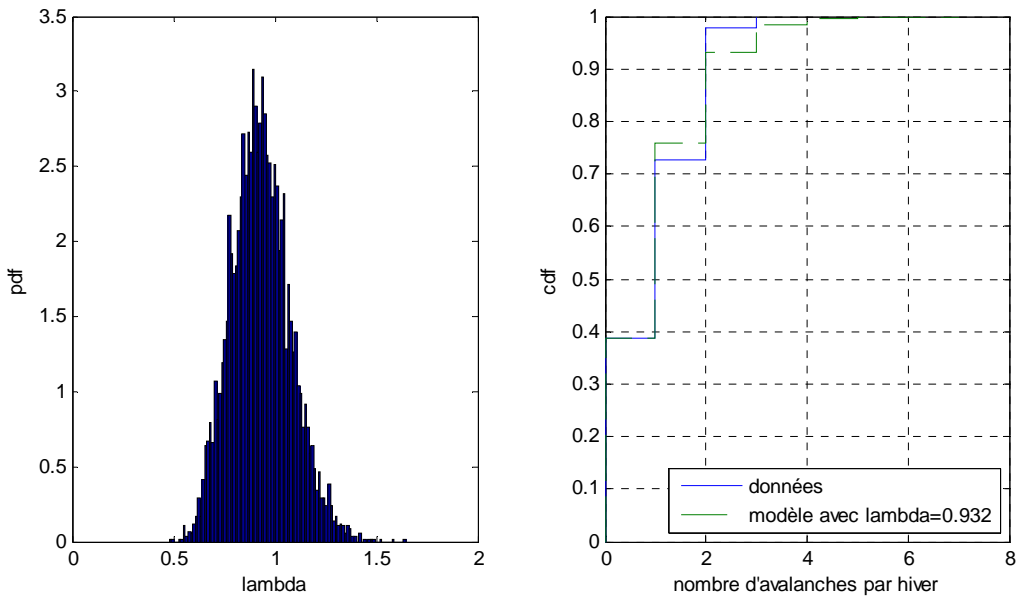


Figure 5.4 : Inférence sur le couloir EPA 13 de Bessans



5.2. Intégrer la dimension spatiale

5.2.1. Un modèle de base

5.2.1.1. Principe

On s'intéresse à présent, conformément au cadre défini au chapitre 3.5.2, au développement d'une méthode de calcul des fréquences avalanches faisant intervenir la modélisation spatiale

hiérarchique. Son principe est de considérer la coordonnée spatiale comme un substitut à des covariables non observées dont la répartition explique les variations locales de l'activité avalancheuse. La structure spatiale est ainsi utilisée pour améliorer l'estimation en transférant de l'information d'un site à un autre.

Le choix d'une échelle spatiale bien adaptée constitue un problème crucial. L'échelle de la commune a été retenue car on a estimé que la localisation de nombreux couloirs était par le passé insuffisamment précise pour travailler directement avec les coordonnées cartésiennes des sites EPA. La structure de données ainsi obtenue, dite en lattice, rend possible l'adaptation au domaine des avalanches de modèles spatiaux sur grille irrégulière utilisés en épidémiologie statistique, domaine dans lequel l'analyse de données de comptage pour des phénomènes rares et discrets est fondamentale.

Il est impossible de présenter tous les modèles spatiaux qui pourraient être utilisés pour décrire les variations spatiales des occurrences avalancheuses. Des revues détaillées existent en épidémiologie statistique (Best et al., 2004), de sorte que l'on se contentera de présenter le modèle proposé par Besag et al. (1991) que l'on a adapté au problème étudié pour obtenir un modèle spatial de base. Ce modèle sera ensuite, conformément au cadre décrit au chapitre 3.5.3, amélioré par touches successives afin de coller au plus près de la réalité du phénomène modélisé et, en particulier, identifier la composante climatique régionale. L'ensemble des résultats détaillés dans ce chapitre a été obtenu avec la contribution active de Robert Kies qui a effectué sur ce sujet son stage de master 2 Ingénierie Statistique de l'Université Joseph Fourier (Kies, 2007) sous ma direction.

5.2.1.2. *Modèle de vraisemblance, normalisation et risque multiplicatif*

Conformément aux notations introduites au chapitre 3.5.2, une région composée de M communes est considérée. On note les nombres d'avalanches observés à l'échelle régionale $(a_{11}, a_{12}, \dots, a_{jt}, \dots, a_{MT_{obs}})$, $j \in [1, M]$, $t \in [1, T_{obs}]$ au cours des T_{obs} années d'observation. Un modèle d'observation local de Poisson est conservé, puisque, s'il est valide au niveau d'un couloir, il doit l'être également au niveau communal étant donné ses propriétés d'invariance d'échelle. Les paramètres λ_{jt} , $j \in [1, M]$, $t \in [1, T_{obs}]$ résument alors l'activité avalancheuse locale et ses variations au cours du temps (eq. 5.4). Sous l'hypothèse de stationnarité temporelle, le modèle poissonien se ramène à l'observation de répétitions indépendantes (eq. 5.5). Les deux hypothèses seront discutées au chapitre 5.2.3.

$$l(a_{jt} | \lambda_{jt}) = \frac{\lambda_{jt}^{a_{jt}}}{a_{jt}!} \times \exp(-\lambda_{jt}) \quad (\text{eq. 5.4})$$

$$l(a_{j1}, a_{j2}, \dots, a_{jT_{obs}} | \lambda_j) = \prod_{t=1}^{T_{obs}} l(a_{jt} | \lambda_j) \quad (\text{eq. 5.5})$$

L'espérance mathématique de la variable aléatoire a_j est simplement λ_j . Et connaissant les λ_j , on peut calculer les fréquences annuelles moyennes f_j , $j \in [1, M]$, c'est-à-dire les nombres moyens d'avalanches par an et couloir dans chaque commune, en utilisant les nombres c_j de couloirs d'avalanches observés dans chaque commune (eq. 5.6).

$$f_j = \frac{\lambda_j}{c_j} \quad (\text{eq. 5.6})$$

Afin de pouvoir comparer les observations d'une commune à l'autre et détecter les structures spatiales, une normalisation par un nombre e_j , $j \in [1, M]$ d'avalanches attendu chaque hiver dans chaque commune sous hypothèse d'une homogénéité régionale de l'activité avalancheuse est introduite. Si les moyennes empiriques $a_{j.} = \frac{1}{T_{obs}} \times \sum_{t=1}^{T_{obs}} a_{jt}$ ne sont pas significativement différentes des nombres

annuels attendus e_j , $j \in [1, M]$, l'activité avalancheuse est effectivement homogène dans l'espace. Un modèle multiplicatif est de ce fait très commode pour comparer les nombres attendus et les observations afin de détecter les hétérogénéités spatiales (eq. 5.7). Selon les conventions utilisées en

épidémiologie spatiale, les RR_j sont « les risques relatifs » qui distinguent les communes où la fréquence avalancheuse est significativement plus haute ou plus faible que la moyenne régionale. Si RR_j est significativement supérieur à 1, un excès de risque existe dans la commune j . Réciproquement, si RR_j est significativement inférieur à 1, la commune considérée est sous-exposée.

$$\lambda_j = e_j \times RR_j \text{ (eq. 5.7)}$$

5.2.1.3. Niveaux hiérarchiques

Les variations du risque relatif dépendent fortement de la taille de la population considérée. Les valeurs les plus extrêmes sont généralement associées aux communes avec les plus petites populations (Elliott et al., 2000). Par conséquent, l'estimation directe des risques relatifs à l'aide des moyennes

empiriques des observations et des nombres prévus selon $\hat{RR}_j = \frac{a_j}{e_j}$ n'est pas adapté aux communes

de petite taille. Cette difficulté peut, en revanche, être surmontée en utilisant un modèle hiérarchique qui explique ce qui est observé dans une commune par ses voisins et améliore ainsi l'estimation locale en partageant l'information entre les différentes communes. Pour éviter des problèmes de positivité pendant l'inférence, le modèle hiérarchique est proposé pour les log risques relatifs (eq. 5.8).

$$\ln(\lambda_j) = \ln(e_j) + \ln(RR_j) \text{ (eq. 5.8)}$$

Le modèle hiérarchique proposé par Besag et al. (1991) combine un terme local non structuré et un terme spatialement structuré (eq. 5.9). Le premier terme v_j permet de représenter un défaut ou un excès local dans le risque relatif qui n'est pas corrélé avec ce qui est observé dans les autres communes. Il est modélisé par un bruit blanc gaussien avec un paramètre σ_v à estimer (eq. 5.10). Le second terme u_j vise à améliorer l'estimation locale grâce à l'information transmise par les autres communes. La corrélation spatiale est prise en considération en utilisant un modèle gaussien Conditionnel AutoRégressif (CAR) dit intrinsèque (Mollié, 1996). Ce modèle avec seulement un paramètre σ_u à estimer suppose que, sachant toutes les autres valeurs $u_{-j} = u_{k \neq j}$ du champ spatial

$u = (u_1, u_2, \dots, u_M)$, la valeur u_j est issue d'une distribution normale dont la moyenne est égale à la moyenne de u_{-j} pondérée par un indice de distance ω_{jk} entre la commune considérée et toutes les autres communes. L'écart-type est, quant à lui, proportionnel à la somme des indices de distance $\omega_{j+} = \sum_{k \neq j} \omega_{jk}$ entre la commune considérée de toutes les autres communes. La valeur u_j est ainsi

lissée par tous les autres u_{-j} avec un niveau de lissage dépendant de l'inverse de la distance totale séparant la commune considérée et toutes les autres (eq. 5.11). Le modèle vérifie $\sum u_j = 0$, de sorte que la moyenne régionale α_{moy} des log risques relatifs qui modélise un éventuel écart régional entre

nombres attendus et observés s'obtient simplement selon $\alpha_{moy} = \frac{1}{M} \sum_{j=1}^M v_j$.

Une bonne mesure de la significativité de la structure spatiale est donnée par la comparaison des contributions respectives du bruit blanc et du terme spatial à la variabilité totale au travers du calcul du rapport r_s (eq. 5.12). Avec un processus spatial aléatoire, r_s peut évidemment prendre n'importe quelle valeur entre 0 et 1, mais quand il est au-dessus de 0.5, des motifs spatiaux significatifs existent dans la réalisation observée du processus (Mollié, 1996).

$$\ln(RR_j) = v_j + u_j \text{ (eq. 5.9)}$$

$$V_j \sim N(0, \sigma_v^2) \text{ i.e. } p(v_j | \sigma_v) = \frac{1}{\sqrt{2\pi} \times \sigma_v} \times \exp\left(-\frac{v_j^2}{2 \times \sigma_v^2}\right) \text{ (eq. 5.10)}$$

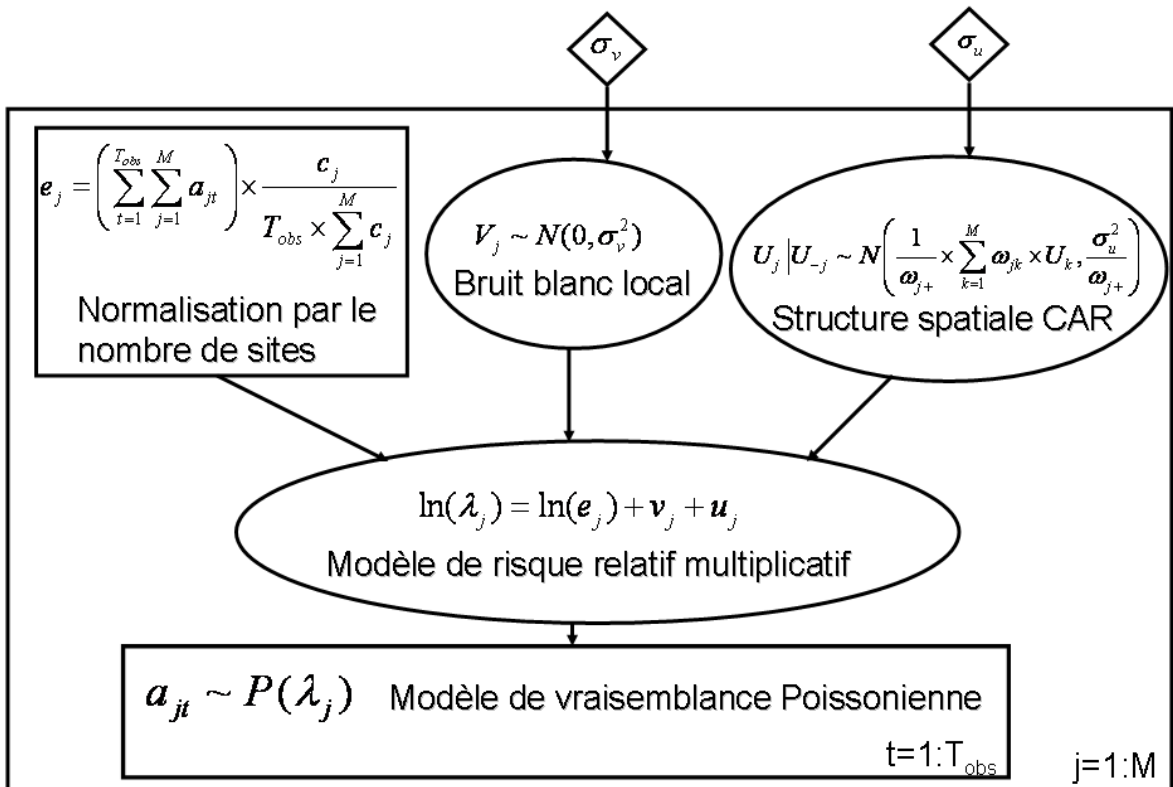
$$U_j / U_{-j} \sim N\left(\frac{1}{\omega_{j+}} \times \sum_{k=1}^M \omega_{jk} \times U_k, \frac{\sigma_u^2}{\omega_{j+}}\right)$$

$$\text{i.e. } p(u_j | u_{-j}, \sigma_u) = \frac{\sqrt{\omega_{j+}}}{\sqrt{2\pi} \times \sigma_u} \exp\left(-\frac{\omega_{j+} \times \left(u_j - \frac{1}{\omega_{j+}} \times \sum_{k=1}^M \omega_{jk} \times u_k\right)^2}{2 \times \sigma_u^2}\right) \quad (\text{eq. 5.11})$$

$$r_s = \frac{\text{VAR}[U]}{\text{VAR}[U] + \sigma_v^2} \quad \text{avec } \text{VAR}[U] = \frac{1}{M} \times \sum_{j=1}^M \left(\left(U_j - \frac{1}{M} \times \sum_{j=1}^M U_j \right)^2 \right) \quad (\text{eq. 5.12})$$

Pour résumer, le modèle spatial de base proposé implique trois champs latents (λ_j, u_j, v_j) , $j \in [1, M]$ et les hyperparamètres (σ_u, σ_v) . Par rapport aux notations introduites au chapitre 3.5.2, on a donc $\theta_{F_j} = (\lambda_j, u_j, v_j)$ et $\theta_{FR} = (\sigma_u, \sigma_v)$. En outre, les trois champs latents appartiennent à deux niveaux hiérarchiques différents, de sorte qu'un modèle à quatre niveaux est obtenu. Il est résumé en Figure 5 sous la forme d'un DAG.

Figure 5.5 : DAG du modèle spatial de base



5.2.1.4. Inférence bayésienne et triages prédictifs du modèle

L'équation 5.13 applique le théorème de Bayes au modèle proposé. Pour en simplifier l'expression, il est écrit de façon un peu abusive pour une seule commune. $p(\sigma_u, \sigma_v)$ est le *prior* joint pour les deux

paramètres de variance des champs latents, $\prod_{t=1}^{T_{obs}} l(a_{jt} | \lambda_j)$ est la vraisemblance des observations dans la commune considérée et $p(\lambda_j | e_j, u_j, v_j)$ ainsi que $p(u_j, v_j | \sigma_u, \sigma_v)$ sont les probabilités d'occurrence des variables latentes. A noter que $p(\lambda_j | e_j, u_j, v_j) = \delta(e_j \times \exp(u_j + v_j))$. Le terme de gauche $p(\lambda_j, u_j, v_j, \sigma_u, \sigma_v | a_{jt}, e_j)$ est bien-sûr la loi *a posteriori* jointe des inconnues sachant les observations et la variable explicative e_j .

$$p(\lambda_j, u_j, v_j, \sigma_u, \sigma_v | a_{jt}, e_j) \propto p(\sigma_u, \sigma_v) \times p(u_j, v_j | \sigma_u, \sigma_v) \times p(\lambda_j | e_j, u_j, v_j) \times \prod_{t=1}^{T_{obs}} l(a_{jt} | \lambda_j) \quad (\text{eq. 5.13})$$

La structure du modèle rend l'algorithme de Gibbs particulièrement bien adapté à l'inférence puisque les lois conditionnelles nécessaires sont directement accessibles. Un exemple évident est $V_j | \sigma_v \sim N(0, \sigma_v)$. Les autres lois conditionnelles complètes sont données par Mollié (1996). Les calculs d'inférence ne nécessitent pas, à la différence du modèle d'intensité, de propagation numérique. Ils ont donc pu, comme indiqué au chapitre 3, être menés à l'aide du logiciel spécialisé Winbugs. L'inférence de tous les autres modèles qui seront présentés aux paragraphes suivants de ce chapitre s'est effectuée de manière complètement similaire.

Conformément à l'équation 3.19, la simulation prédictive des nombres d'avalanches par commune et par hiver a_j^{pred} s'effectue en moyennant la vraisemblance sur la loi *a posteriori* (5.14). Néanmoins, comme on a affaire à un modèle hiérarchique, le calcul s'effectue, comme détaillé au chapitre 3 pour le modèle d'intensité, par intégration sur la variabilité des variables latentes. Numériquement c'est toutefois « sans douleur » puisqu'il suffit d'échantillonner le modèle à chaque itération en fonction des valeurs prises par les inconnues.

Enfin, la loi prédictive de la fréquence avalancheuse communale f_j^{pred} peut être obtenue en normalisant a_j^{pred} par le nombre de couloirs observés (eq. 5.15). Sa moyenne peut être considérée comme le meilleur pari sachant les observations tandis que l'intervalle de crédibilité à 95% donne des bornes maximale et minimale vraisemblables.

$$p(a_j^{pred} | data) = \int l(a_j | \lambda_j) \times p(\lambda_j, u_j, v_j, \sigma_u, \sigma_v | data) \times d\sigma_u \times d\sigma_v \quad (\text{eq. 5.14})$$

$$f_j^{pred} = \frac{a_j^{pred}}{c_j} \quad (\text{eq. 5.15})$$

5.2.2. Application à la Savoie

5.2.2.1. Données

Pour la modélisation à l'échelle régionale des occurrences avalancheuses, un prétraitement manuel tel qu'il a été effectué pour sélectionner les chroniques les plus fiables pour l'application des modèles de prédétermination locale n'est pas envisageable. Heureusement, pour des données de comptage telles les occurrences avalancheuses, seules deux sources d'erreur sont à considérer : d'une part la mauvaise localisation de certains sites dans le passé, d'autre part les observations manquantes.

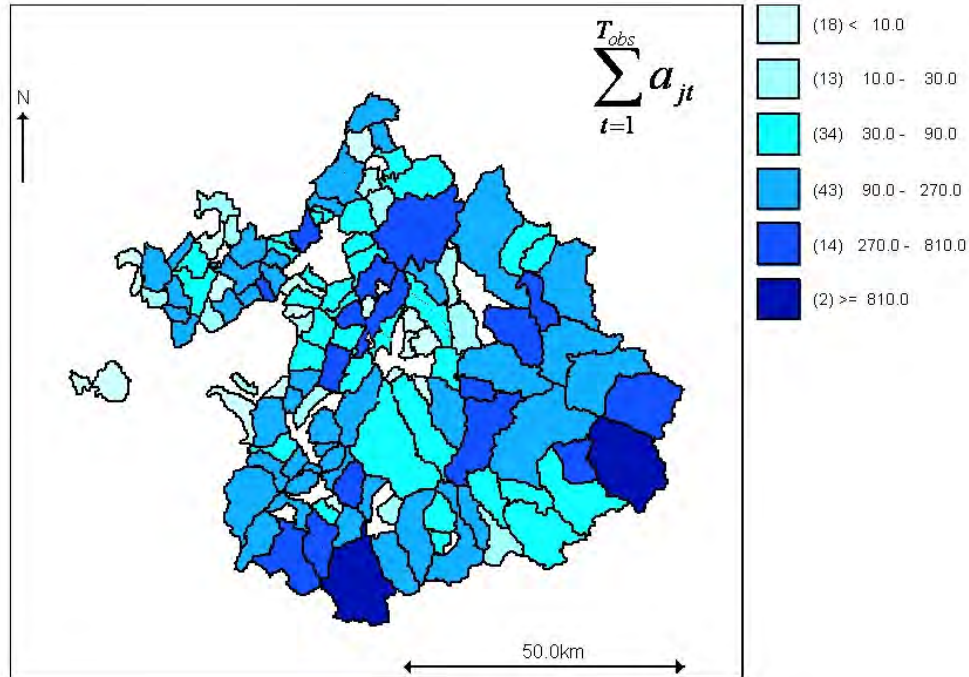
L'erreur de localisation disparaît à l'échelle communale car les sites ont été rattachés à la bonne commune depuis le démarrage de l'EPA. Pour limiter le nombre d'événements manquants, on a, dans un premier temps, appliqué le modèle uniquement dans le département de la Savoie où l'activité avalancheuse est la plus intense en France (environ 30% des événements enregistrés dans l'EPA) et où la qualité des données est supposée relativement satisfaisante, tout au moins au cours des dernières décennies.

Dans ce département, des avalanches ont été enregistrées dans l'EPA depuis 1902 dans 124 communes sur 1347 sites, avec de 1 à 49 sites observés par commune. Au début du 20^{ème} siècle, le nombre annuel

total d'avalanches à l'échelle du département $a_{iR} = \sum_{j=1}^M a_{jt}$ présente une tendance positive significative

parce que le protocole d’observation s’est mis progressivement en place. Sur les 60 dernières années, cette tendance a disparu, de sorte que l’hypothèse d’un processus stationnaire peut être effectuée. Cela correspond à 18 755 avalanches, avec de 0 à 1575 avalanches par commune durant $T_{obs} = 60$ ans (Figure 5.6).

Figure 5.6 : Nombres totaux d’avalanches en Savoie



Les nombres d’avalanches sont donnés à l’échelle communale en six classes. Entre parenthèses, le nombre de communes correspondant à chaque classe. La période d’étude correspond à 1943-2003.

5.2.2.2. Normalisation, voisinage et regroupement par tranches de 5 ans

En épidémiologie statistique, des taux de référence pour l’occurrence des maladies rares ont pu être calculés à large échelle en fonction de l’âge et du sexe (Elliott et al., 2000). Pour les avalanches, de tels taux n’existent évidemment pas, et il n’existe pas de méthode permettant de calculer un nombre d’avalanches attendu par couloir et par an en fonction de caractéristiques topographiques et/ou nivologiques. Par conséquent, les e_j ont été calculés selon l’équation 5.16 à partir du nombre total d’avalanches observées sur le département durant les 60 ans d’observation et du nombre de couloirs observés dans chaque commune (Figure 5.19, gauche). Ainsi, grâce au modèle de risque multiplicatif utilisé, les observations sont comparées à un processus totalement homogène dans l’espace. De plus, puisque la somme des nombres attendus est égale à la somme des nombres observés, la moyenne régionale des log risques relatifs α_{moy} est ainsi nécessairement faible. On le vérifiera *a posteriori*.

$$e_j = \left(\sum_{j=1}^M \sum_{t=1}^{T_{obs}} a_{jt} \right) \times \frac{c_j}{T_{obs} \times \sum_{j=1}^M c_j} \quad (\text{eq. 5.16})$$

Dans un premier temps, une matrice de distance très simple a été employée pour calculer les indices de distance ω_{jk} . La corrélation spatiale a été supposée limitée aux voisins directs avec un poids égal pour tous les voisins. Ceci se traduit par $\omega_{jk}=1$ si les communes j et k ont une frontière commune et par $\omega_{jk}=0$ sinon. Cependant, quelques ω_{jk} ont été également fixés à 1 pour des communes séparées par des « trous » dans la carte de la Figure 5.6. Ces trous correspondent à de petites communes généralement situées dans les fonds de vallée et qui ont été exclues de l’étude parce qu’elles ne contiennent aucun site EPA. On a supposé que deux communes séparées par un tel trou pouvaient partager de l’information même si elles n’avaient pas une frontière commune.

Enfin, une étude descriptive préliminaire a montré que les nombres d'avalanches annuels à l'échelle communale étaient fortement variables. Physiquement, cela s'explique par l'occurrence simultanée d'un grand nombre d'avalanches sur les différents sites d'une même commune au moment des fortes tempêtes. D'un point de vue statistique, cela empêche l'utilisation d'un modèle poissonien homogène pour les décomptes annuels car les hivers très avalancheux sont trop différents des hivers peu avalancheux. Afin de pouvoir malgré tout appliquer ce premier modèle au département de la Savoie, $N_{T_{obs}} = 12$ « paquets » de 5 ans ont été constitués afin de lisser la variabilité interannuelle. Les nombres d'avalanches observés sur 5 ans dans chaque commune sont notés

$$a_{jN_t} = \sum_{t=5 \times (N_t - 1) + 1}^{5 \times N_t} a_{jt}, N_t \in [1, 12].$$

Quant aux nombres attendus sur 5 ans dans chaque commune, ils sont simplement égaux à cinq fois les nombres attendus annuellement. L'ensemble de ces hypothèses sera discuté puis remis en cause au cours des paragraphes suivants.

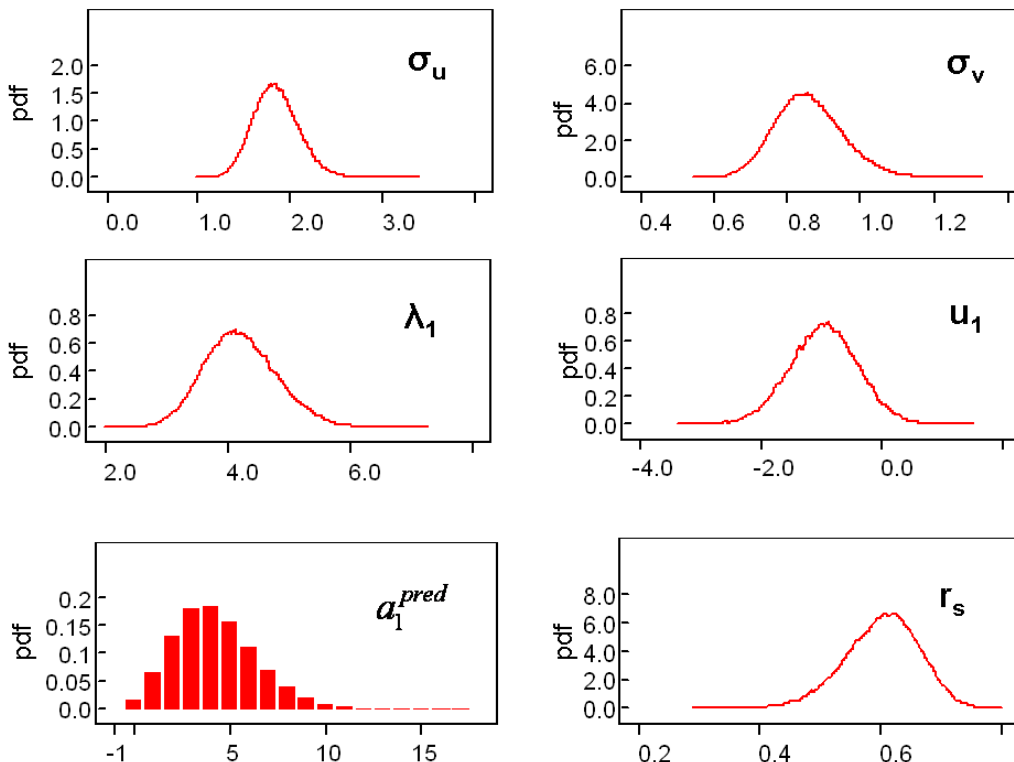
5.2.2.3. Inférence

De façon à laisser les données s'exprimer, des *priors* peu informatifs ont été choisis pour (σ_u, σ_v) :

les précisions $1/(\sigma_u)^2$ et $1/(\sigma_v)^2$ ont toutes deux été tirées selon des lois *Gamma*(0.01, 0.001)

indépendantes. 120 000 itérations de l'algorithme de Gibbs ont été effectuées. Les 20 000 premières ont constitué la période de chauffe et les 100 000 autres ont permis d'obtenir les lois *a posteriori* marginales des inconnues. Les lois *a posteriori* des autres grandeurs d'intérêt telles les risques relatifs et le rapport spatial ont été obtenues en calculant leur valeur à chaque itération. De même, le modèle a été échantillonné à chaque itération conditionnellement aux valeurs prises par les inconnues pour simuler les lois prédictives. La Figure 5.7 présente quelques unes des lois *a posteriori* marginales obtenues. Elles sont résumées dans le Tableau 5.3. L'index 1 correspond à la commune d'Aigueblanche, tandis que l'index 43 correspond à la commune de La Chapelle. Les moyennes *a posteriori* des trois champs latents sont présentées sous forme de cartes dans les Figures 5.8 à 5.10.

Figure 5.7 : Lois marginales *a posteriori* pour le modèle spatial de base



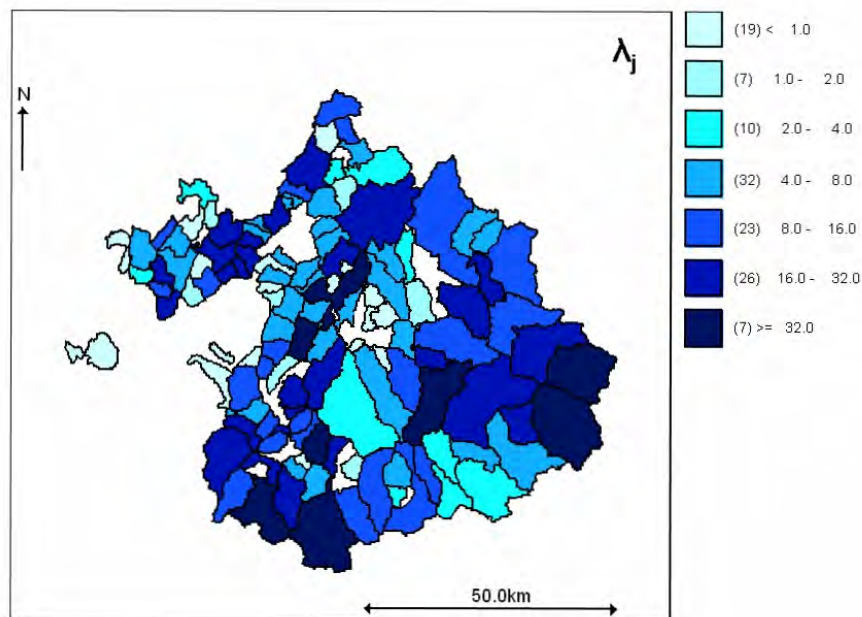
Modèle spatial de base appliqué à la Savoie. Variabilité spatiale σ_u , bruit blanc σ_v et rapport spatial r_s . Les variables latentes et lois prédictives correspondent à la commune d'Aigueblanche (indexée 1).

Les moyennes *a posteriori* des λ_j s'échelonnent de 0.3 à 131.4 avalanches par commune et période de 5 ans. Les intervalles de crédibilité correspondants montrent que l'activité avalancheuse à l'échelle communale n'est pas du tout un processus spatialement homogène, ce qui est bien sûr naturel puisque le nombre de sites avalancheux varie fortement d'une commune à l'autre. Ainsi le Tableau 5.3 montre que le nombre d'avalanches sur 5 ans sont significativement différents dans les communes 1 et 43 au seuil des 95%. Des motifs spatiaux s'observent également (Figure 5.8), avec en particulier de fortes valeurs de λ_j concentrées à l'est, ce qui suggère que l'activité avalancheuse est, au moins partiellement, spatialement structurée. Des motifs spatiaux relativement identiques apparaissent également en Figure 5.6, ce qui reflète simplement le fait que le nombre moyen d'avalanches observé dans chaque commune sur cinq ans est l'estimateur du maximum de vraisemblance de λ_j si l'on ignore les niveaux hiérarchiques du modèle.

Tableau 5.3 : Modèle spatial de base, paramètres et variables latentes pour les communes 1 et 43

	Moy.	SD	q _{2.5}	q _{97.5}
σ_u	1.84	0.24	1.39	2.35
σ_v	0.86	0.09	0.69	1.05
r_s	0.60	0.06	0.47	0.71
λ_1	4.23	0.59	3.15	5.46
u_1	-0.96	0.56	-2.08	0.13
v_1	-0.47	0.56	-1.56	0.66
RR_1	0.24	0.03	0.18	0.31
f_1	0.06	0.01	0.04	0.07
a_1^{pred}	4.24	2.14	1.00	9.00
f_1^{pred}	0.06	0.03	0.01	0.12
λ_{43}	15.32	1.13	13.19	17.63
RR_{43}	1.65	0.12	1.42	1.90
f_{43}	0.38	0.03	0.33	0.44
a_{43}^{pred}	15.34	4.09	8.00	24.00
f_{43}^{pred}	0.38	0.10	0.20	0.60

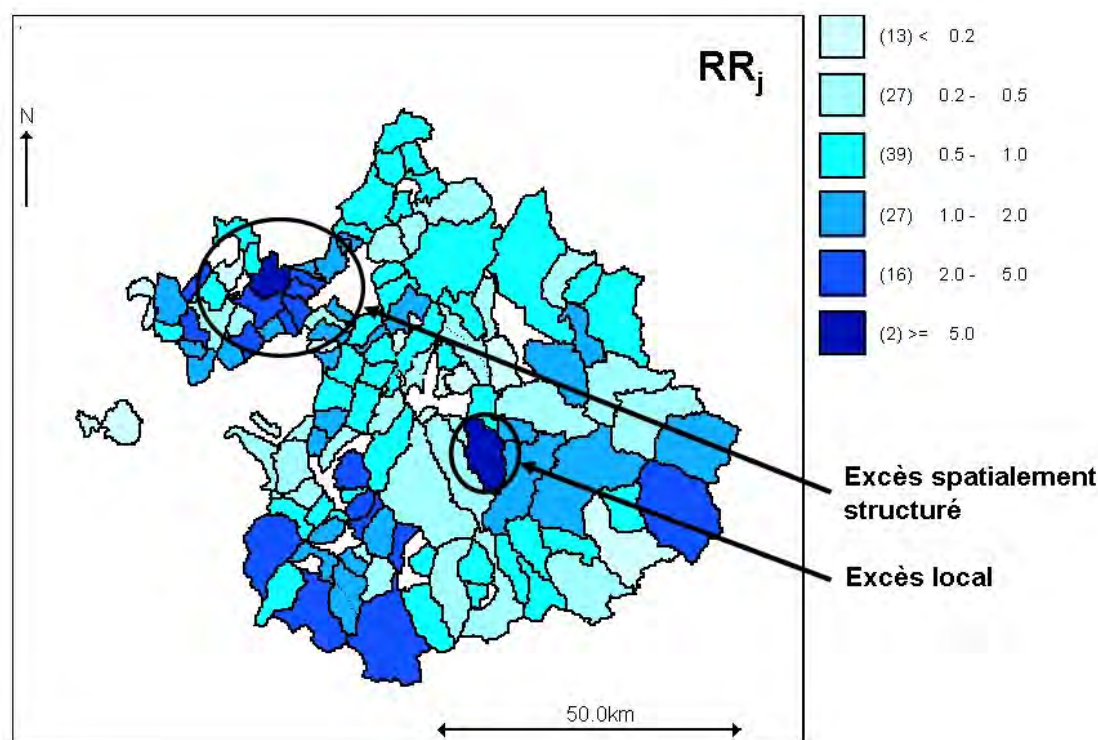
Figure 5.8 : Moyenne *a posteriori* des fréquences avalancheuses communales moyennes



Modèle spatial de base appliqué à la Savoie. Les nombres moyens sont donnés par commune pour une période de 5 ans.

L'hétérogénéité spatiale des fréquences avalancheuses apparaît plus nettement grâce à l'analyse des risques relatifs. Ils s'échelonnent de 0.04 à 6.2, ce qui indique malgré la normalisation une très forte variabilité intercommunale autour de la moyenne régionale. Ainsi le Tableau 5.3 montre un excès de risque significatif dans la commune de La Chapelle ($RR_{43} > 1$ au seuil des 95%) et, au contraire, un déficit de risque significatif dans la commune d'Aigueblanche ($RR_1 < 1$ au seuil des 95%). La distribution spatiale des risques relatifs (Figure 5.9) montre clairement des groupes de communes avec des valeurs fortes de RR_j , en particulier dans le nord-ouest, ainsi que des groupes de communes avec des faibles valeurs de RR_j , en particulier dans le nord-nord-est. Ces groupes de communes traduisent l'existence d'une structure spatiale de l'activité avalancheuse. D'autres communes présentent cependant des risques relatifs très forts ou très faibles avec très peu de corrélation spatiale avec leurs voisines. Elles indiquent que l'activité avalancheuse présente donc également une composante locale non nulle. La moyenne *a posteriori* du rapport spatial s'établit de ce fait à 0.6, avec un intervalle de crédibilité à 95% de [0.47, 0.71] (Figure 5.7 et Tableau 5.3). Cela traduit effectivement une composante spatiale du processus étudié légèrement prédominante et une composante locale non négligeable.

Figure 5.9 : Moyenne *a posteriori* des risques relatifs



Modèle spatial de base appliqué à la Savoie.

La carte du champ latent u_j donne la vision la plus claire de la structure spatiale de l'activité avalancheuse puisque celle-ci a été séparée du bruit (Figure 5.10). Elle montre trois groupes de communes avec une composante spatiale forte dans le risque relatif. Ces groupes correspondent grossièrement aux massifs montagneux des Bauges (au nord ouest), de la Vanoise (au sud est) et des Aiguilles d'Arves (au sud ouest). Dans le reste du département, la composante spatiale est relativement homogène et faible. La carte du bruit local (Figure 5.13, gauche) permet quant à elle de bien visualiser les communes avec de forts excès ou déficit locaux. On vérifie par ailleurs que

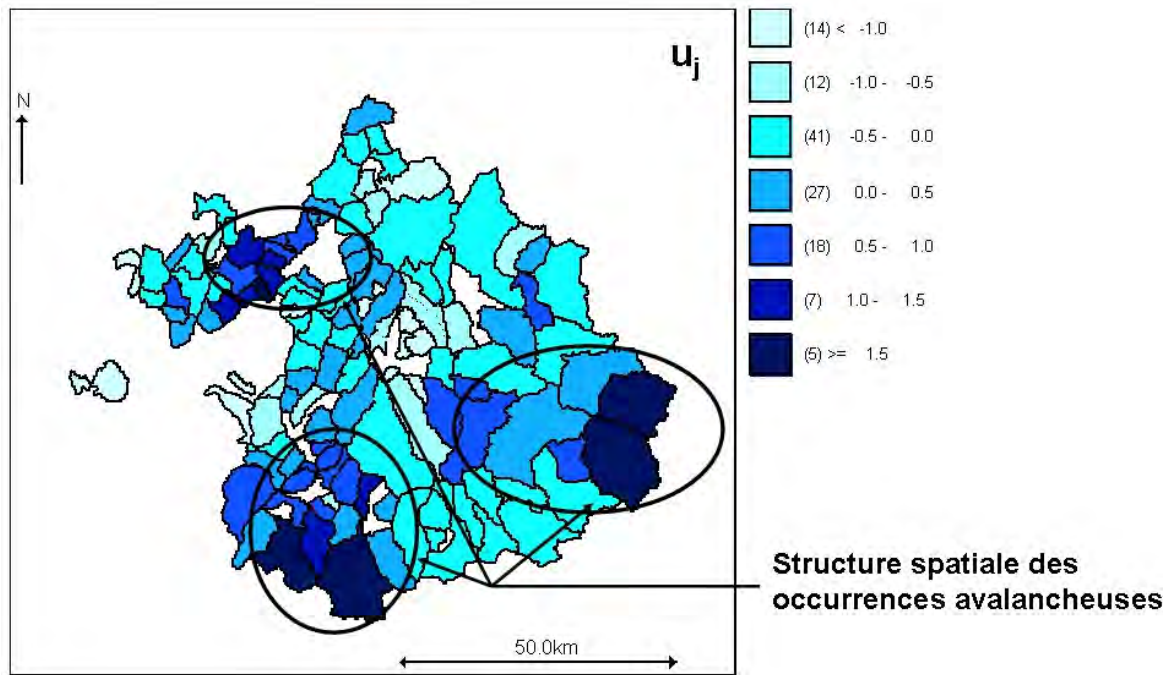
$$\alpha_{moy} = \frac{1}{M} \sum_{j=1}^M v_j$$

n'est pas significativement différent de zéro, ce qui est cohérent puisque, à cause de

la méthode de normalisation utilisée, la somme des nombres attendus est nécessairement égale à la somme des nombres enregistrés d'avalanches.

Enfin, les fréquences avalancheuses moyennes f_j sont les quantités d'intérêt pour la prédétermination locale en association avec les modèles d'aléa proposés au chapitre 4. Afin de les obtenir, les λ_j doivent être normalisés par les nombres de sites de chaque commune et divisés par 5 pour se ramener à des fréquences annuelles. Leurs moyennes *a posteriori* s'échelonnent selon la commune de 0.011 à 1.43 avalanches par an et par couloir. Une carte n'en est pas proposée parce que les motifs spatiaux sont presque identiques à ceux observés pour les risques relatifs. En effet, les moyennes *a posteriori* des f_j et des RR_j sont presque proportionnelles pour toutes les communes.

Figure 5.10 : Moyenne *a posteriori* de la composante spatiale



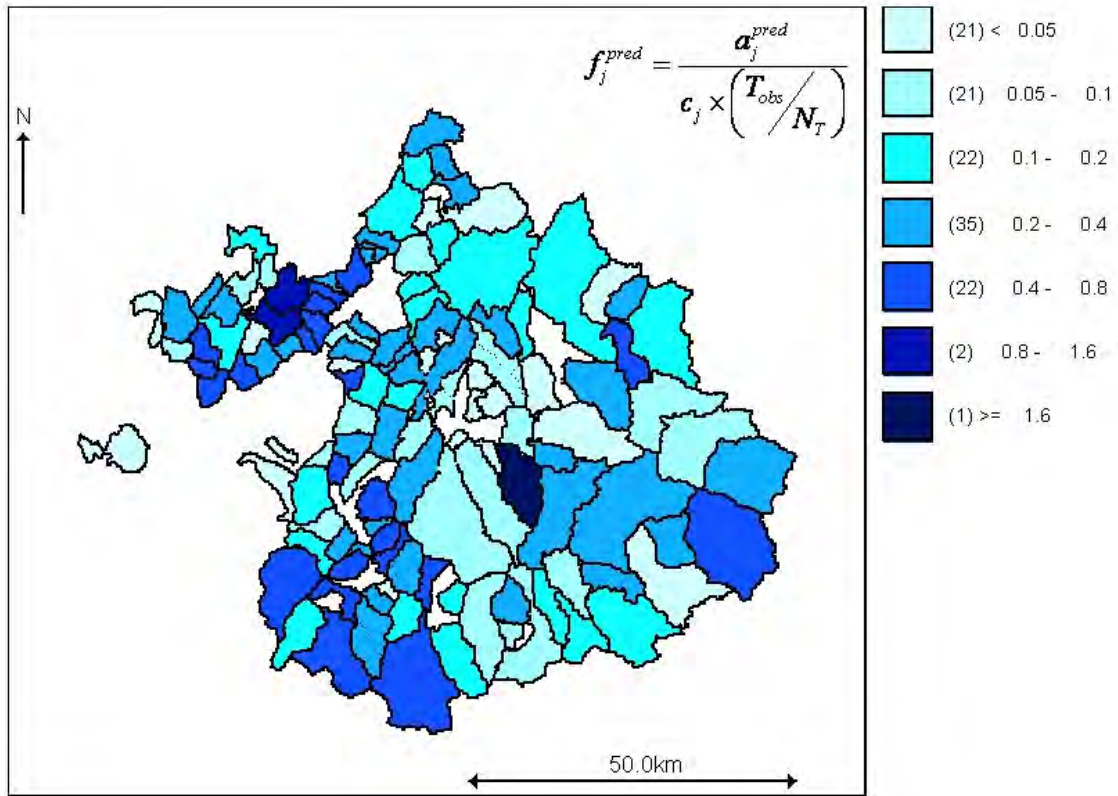
Modèle spatial de base appliqué à la Savoie.

5.2.2.4. Prédiction

Les lois prédictives locales sont des mélanges de lois de Poisson et sont de ce fait intrinsèquement discrètes. Les moyennes prédictives sur 5 ans dans les différentes communes sont logiquement très proches des moyennes des λ_j , de sorte que les motifs spatiaux sur les cartes des a_j^{pred} et des λ_j sont presque identiques. Néanmoins, les intervalles de crédibilité des a_j^{pred} sont bien plus larges que ceux des λ_j parce qu'à la fois la variabilité poissonnienne et l'erreur d'estimation sont prises en compte. Ainsi le Tableau 5.3 montre que, pour les communes 1 et 43, l'intervalle de crédibilité à 95% est environ quatre fois plus large pour la loi prédictive a_j^{pred} que pour λ_j . La Figure 5.7 permet quant à elle grâce à l'exemple de $p(a_1^{pred} | data)$ de visualiser la lourdeur de la queue de la distribution de la loi prédictive due au mélange poissonnien, cette dernière permettant à des valeurs très supérieures aux moyennes d'être générées.

De même, les fréquences annuelles moyennes prédites f_j^{pred} sont très proches des fréquences annuelles moyennes f_j . Comme attendu, les motifs spatiaux correspondant aux f_j^{pred} (Figure 5.11) sont donc très similaires à ceux observés pour les f_j et les RR_j . Toutefois, les intervalles de crédibilité sont là aussi bien plus larges pour les valeurs prédites que pour les moyennes *a posteriori*.

Figure 5.11 : Prédiction moyenne pour la fréquence avalancheuse annuelle



Modèle spatial de base appliqué à la Savoie. La moyenne de la loi prédictive communale est normalisée par le nombre de couloirs et ramenée à 1 an.

5.2.3. Discussion

5.2.3.1. Adéquation au modèle

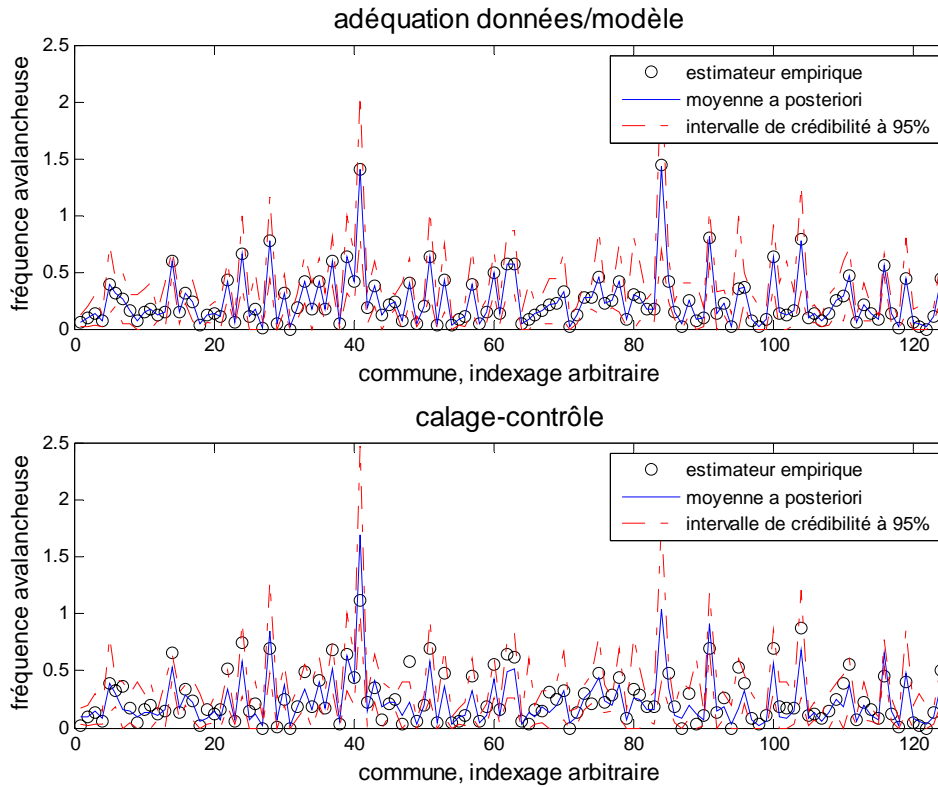
La Figure 5.11 présente une comparaison entre les estimateurs empiriques des fréquences

avalancheuses moyennes par couloir $\hat{f}_j = \frac{\sum_{t=1}^{T_{obs}} a_{jt}}{c_j \times T_{obs}}$ et les lois prédictives des f_j^{pred} . Quand

l'inférence est effectuée sur le jeu de données entier, l'adéquation entre les données et le modèle est très bonne puisque tous les estimateurs empiriques sont très proches des moyennes des lois prédictives (Figure 5.12 haut). Le jeu de données a également été séparé en deux échantillons de calage et de contrôle en affectant aléatoirement chacune des 12 périodes à l'un des deux échantillons.

La comparaison des données de contrôle et des lois prédictives dérivées de l'analyse de l'échantillon de calibration est encourageante (Figure 5.12 bas) : presque tous les estimateurs empiriques sont dans l'intervalle de crédibilité à 95% et les différences entre les moyennes prédictives et les observations sont faibles pour une grande majorité des communes. Ces résultats suggèrent que le modèle s'ajuste bien aux données. Néanmoins, ils montrent également que le modèle ne prédit des avalanches que là où des avalanches ont été observées car la physique du processus spatial n'y est pas prise en compte sous la forme de variables explicatives. Par conséquent, ce modèle de base est vraisemblablement très sensible aux erreurs d'observation (i.e. aux observations manquantes), ce qui limite son utilisation aux départements présentant la meilleure qualité de données.

Figure 5.12 : Adéquation entre données et modèle spatial de base



Modèle spatial de base appliqué à la Savoie. Pour l'adéquation données/modèle, toutes les données sont utilisées pour l'inférence. Pour le calage-contrôle l'estimation est effectuée sur un échantillon de calage et la loi prédictive du modèle est comparée à l'échantillon de contrôle.

5.2.3.2. Intérêt du terme spatial

Le modèle M_0 utilisé jusqu'ici décompose le risque relatif en une composante locale et une composante spatiale selon $\ln(\lambda_j) = \ln(e_j) + u_j + v_j$. Deux modèles plus simples ont été testés et comparés à ce modèle sur la base du DIC présenté au chapitre 2 pour voir si un modèle plus simple pouvait être utilisé. Le premier, noté M_1 , modélise les log risques relatifs par un bruit blanc selon $\ln(\lambda_j) = \ln(e_j) + v_j$. Le second suppose au contraire que les log risques relatifs peuvent être entièrement décrits par un modèle conditionnel autorégressif selon $\ln(\lambda_j) = \ln(e_j) + u_j$.

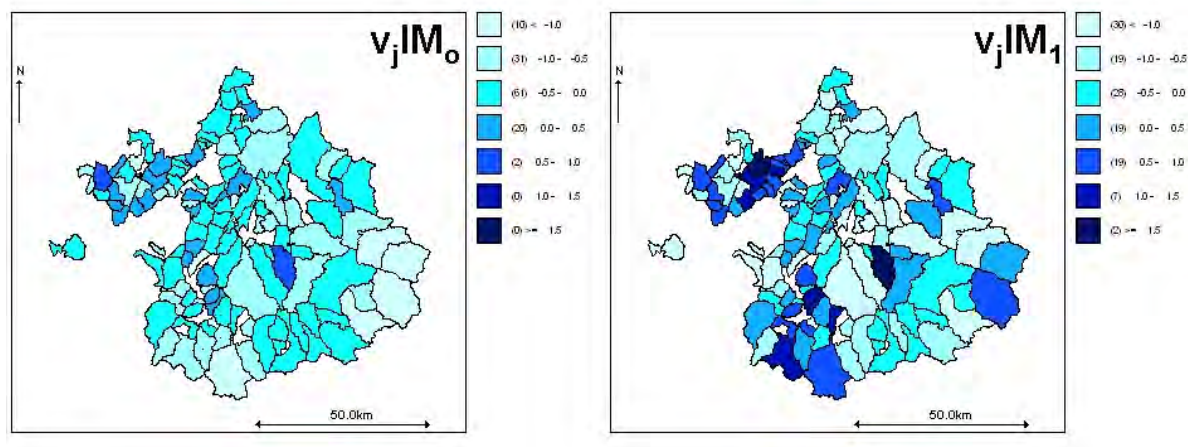
D'après les DIC obtenus (Tableau 5.4), il apparaît que le modèle M_0 est le meilleur des trois. Le modèle M_2 est nettement moins bon que les deux autres. Cela indique que la variabilité locale non structurée spatialement est trop forte pour qu'un modèle "purement spatial" s'ajuste bien aux données. Au contraire, le modèle très simple M_1 est presque aussi bon que le modèle M_0 du point de vue du DIC, ce qui pourrait suggérer que la composante spatiale du modèle apporte peu. Néanmoins, la valeur élevée du rapport spatial plaide en faveur d'une modélisation explicite de la structure spatiale. De plus, les motifs spatiaux observés sur la carte des v_j plaident également pour conserver M_0 (Figure 5.13).

En effet, avec le modèle M_1 , i.e. sans la composante spatiale, les v_j présentent une structure spatiale assez nette alors qu'ils sont supposés distribués selon un bruit blanc indépendant de l'espace. Avec le modèle M_0 , les motifs spatiaux résiduels sur la carte des v_j sont nettement moins clairs, ce qui indique que le modèle M_0 permet la séparation de la structure spatiale et du bruit grâce à la combinaison des deux termes u_j et v_j . Néanmoins, une légère structure spatiale résiduelle semble toujours exister sur la carte des v_j même avec le modèle M_0 . Cela indique que ce modèle de base présente encore des limites et peut par conséquent être amélioré. Ce sera justement l'objectif des paragraphes suivants.

Tableau 5.4 : DIC, modèle spatial de base M_0 contre les modèles simplifiés M_1 et M_2

	M_0	M_1	M_2
DIC	17039	17045	19830

Figure 5.13 : Moyenne a posteriori de la composante locale sans et avec la composante spatiale



Application à la Savoie.

5.2.3.3. Limites et améliorations

- *Exhaustivité de l'observation*

On a jusqu'ici supposé l'observation exhaustive en appliquant le modèle uniquement au département où la qualité de la donnée était la meilleure. Evidemment, cette hypothèse est fallacieuse et peut amener à sous-estimer fortement les fréquences avalanches moyennes si beaucoup d'événements ont été manqués. Une première amélioration indispensable du modèle est donc la prise en compte de l'effort d'observation dans le modèle. La solution retenue est le développement d'une procédure de prétraitement automatisée permettant d'identifier les observations insuffisantes sur la base de tests simples et de les considérer comme manquantes.

- *Covariables et calcul des nombres attendus*

Jusqu'à présent, aucune variable explicative n'a été utilisée dans le modèle. La structure spatiale obtenue fait pourtant ressortir les massifs montagneux. Cela montre que la topographie et la climatologie, à condition d'en disposer d'une mesure à l'échelle communale, pourraient facilement apporter une information utile à l'analyse des variations spatiales des fréquences avalanches. La prédiction des fréquences avalanches moyennes dans les communes possédant peu de sites observés serait ainsi améliorée.

Un problème très proche de celui des covariables est celui du calcul des nombres d'avalanches attendus. Jusqu'à présent, seuls les nombres de couloirs suivis dans chaque commune ont été utilisés pour la normalisation. L'hypothèse sous-jacente est que ces couloirs représentent parfaitement l'activité avalancheuse de la commune. Cela reste bien entendu discutable étant donné que le choix des sites suivis dans EPA reste, pour partie au moins, subjectif. Cette limite pourra être surmontée une fois que des covariables significatives auront été introduites dans le modèle en permettant de recalculer les nombres attendus sur la base de critères plus objectifs que le seul nombre de sites.

- *Echelle spatiale et matrice de distance*

L'échelle de la commune a été choisie en raison de l'imperfection de la base EPA. Les résultats obtenus jusqu'ici suggèrent qu'elle est suffisamment petite pour détecter des variations spatiales significatives de l'activité avalancheuse. Néanmoins, la commune reste une subdivision administrative mal adaptée à l'étude d'un phénomène naturel. En particulier, l'hypothèse d'un comportement homogène de tous les couloirs des plus grandes communes partagées entre différents massifs montagneux et/ou différentes expositions aux vents dominants est certainement douteuse. Une échelle plus adaptée comme celle du site EPA ne pourra cependant être utilisée que lorsqu'un échantillon

suffisant de données bien localisées et donc postérieures au projet de rénovation 2001-2006 aura été collecté.

La formulation de la matrice de distance utilisée jusqu'ici est également un peu fruste avec des poids 0-1. Il paraît donc intéressant de revenir sur cette hypothèse en testant d'autres types de dépendance que le voisinage ou en investiguant de possibles corrélations plus longues afin de préciser la portée de la structure spatiale des occurrences avalanches.

- *Homogénéité du modèle d'observation poissonien*

L'hypothèse d'observations poissoniennes homogènes est également très discutable. Les arguments en faveur d'un modèle poissonien sont son utilisation usuelle pour les événements discrets rares et, plus prosaïquement, l'interprétation directe de son paramètre unique. De plus, il est apparu qu'il fonctionnait bien à l'échelle du couloir, ce qui, du fait de ses propriétés d'invariance d'échelle, devrait théoriquement le rendre également bien adapté à une échelle plus large. Néanmoins, du fait de la forte variabilité interannuelle du nombre d'avalanches à l'échelle communale, cela n'est vrai que si l'on considère des « paquets » d'années permettant de lisser l'observation. Ainsi que l'a montré le calage-contrôle sur la région étudiée, un modèle de Poisson homogène est alors assez efficace. Toutefois, cela n'est pas très satisfaisant conceptuellement et cela ne permet pas d'extraire de l'information sur la variabilité interannuelle.

Une approche possible pour progresser est de remplacer le modèle d'observation poissonien par un modèle de mélange Gamma de lois de Poisson, c'est à dire $a_{jt} \sim P(\lambda_{jt})$ et $\lambda_{jt} \sim \text{Gamma}(a_{\lambda_j}, b_{\lambda_j})$.

Cela revient en fait à modéliser les occurrences avalanches par une loi binomiale négative. Une reparamétrisation utile consiste alors à poser $E[\lambda_j] = \frac{a_{\lambda_j}}{b_{\lambda_j}} = e_j \times RR_j$ et

$\text{VAR}[\lambda_j] = \frac{a_{\lambda_j}}{b_{\lambda_j}^2} = (\sigma_\lambda)^2$, ce qui permet d'exprimer les inconnues a_{λ_j} et b_{λ_j} à l'aide du modèle

de risque utilisé jusqu'ici et d'un paramètre de variance global σ_λ supplémentaire quantifiant la variabilité interannuelle. Des champs latents supplémentaires $\lambda_{jt}, t \in [1, T_{obs}]$ et le paramètre σ_λ sont ainsi introduits. Le nouveau modèle est identique à l'ancien seulement si tous les $\lambda_{jt}, t \in [1, T_{obs}]$ sont identiques, c'est-à-dire si σ_λ est très proche de zéro. Le modèle poissonien est ainsi remplacé par une classe de modèles plus large. Néanmoins, l'inférence MCMC est alors délicate du fait de problèmes numériques lorsque $a_{jt} = 0$. De plus, un tel modèle de mélange permet uniquement de surdispenser le modèle d'observation poissonien sans améliorer beaucoup l'information qui peut en être extraite. Le choix qui a été effectué a donc plutôt été d'introduire la non homogénéité au niveau des risques relatifs, c'est-à-dire une fois effectuée la normalisation par les nombres attendus.

- *Homogénéité des risques relatifs*

En effet, un modèle spatio-temporel peut également être proposé pour les risques relatifs de façon à visualiser les variations au cours de l'espace et du temps des occurrences avalanches par rapport à une moyenne régionale annuelle selon $\ln(\lambda_{jt}) = \ln(e_j) + u_j + v_j + g_t + h_{jt}$, le terme g_t quantifiant la variabilité interannuelle globale à l'échelle de la région étudiée et le terme h_{jt} l'effet spatio-temporel local. Les paragraphes suivants vont être consacrés au développement d'un tel modèle.

5.3. Améliorer la modélisation spatiale

5.3.1. Prendre en compte les observations manquantes

5.3.1.1. Détection des données manquantes

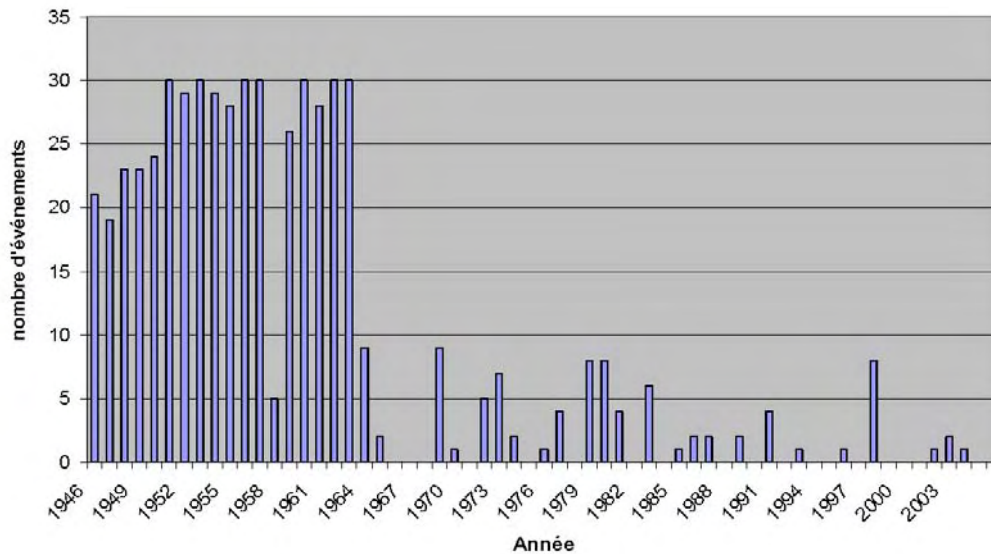
Premièrement, les communes subissant en moyenne au moins 3 événements par année ont été étudiées. Si l'activité avalancheuse est effectivement poissonienne, ces communes ont moins de 5% de

chance de n'enregistrer aucun événement une année donnée. Les décomptes nuls ont donc été déclarés comme des années où l'observation est manquante dans ces communes. Cette méthode a été utilisée afin de détecter pour les communes où l'activité avalancheuse est importante les années où il n'y a pas eu d'observation, par exemple le temps du remplacement d'un observateur partant à la retraite. Bien sûr, pour des communes où l'activité est plus faible, elle n'a pas pu être employée car une année sans événements n'a rien d'extraordinaire.

Deuxièmement, les moyennes mobiles sur 21 ans ont été calculées dans toutes les communes, la moyenne mobile pour l'année t étant bien évidemment la moyenne des nombres d'avalanches sur l'intervalle $[t-10, t+10]$. Ensuite, le maximum de ces moyennes mobiles sur la période d'étude a été détecté et un intervalle de confiance pour cette valeur maximale a été construit. Enfin, chaque année pour laquelle la moyenne mobile était plus petite que la borne inférieure de cet intervalle de confiance a été déclarée comme une année où l'observation n'a pas été effectuée. Si c'était le cas pour t égal à 10 et/ou $t = T_{obs} - 10$, c'est-à-dire pour la première et/ou la dernière moyenne mobile calculée, les 10 premières et/ou les 10 dernières années ont également été déclarées comme manquantes. La fenêtre de 21 ans a été choisie car il a été considéré grâce à la connaissance du réseau d'observation que l'on avait au moins eu dans toutes les communes 20 années consécutives de bonne observation, c'est-à-dire sans observations manquantes. Cette méthode a été utilisée afin de détecter, quelle que soit la taille de la commune, des longues périodes sans observation, à cause d'un observateur local peu motivé par exemple. La Figure 5.14 illustre de façon caricaturale un tel cas de figure avec la commune de Sixt Fer à Cheval pour laquelle les nombres d'événements annuels sont compris entre 20 et 30 jusqu'en 1965 avant de devenir très faibles. Une telle rupture n'a évidemment pas d'origine physique, de sorte qu'il semble effectivement légitime de déclarer les observations manquantes durant toute la seconde partie de la période d'étude.

Au final, deux procédures complémentaires ont donc été employées pour détecter de façon automatisée les observations manquantes dans les chroniques des nombres d'événements. Néanmoins elles restent rudimentaires et des approches plus élaborées prenant notamment en compte la structure spatiale des observations auraient également pu être mises en œuvre pour mettre en évidence les communes avec un déficit d'événements entourées de commune très actives.

Figure 5.14 : Nombres d'événements par hiver sur la commune de Sixt Fer à Cheval



5.3.1.2. Application

Puisque l'on prend en compte les observations manquantes, on peut travailler sur un échantillon plus large et non plus seulement en Savoie. On passe donc à l'ensemble Savoie-Haute-Savoie, soit $M=204$ communes avec au moins un site EPA. La période d'étude choisie correspond aux campagnes 1946 à 2005, soit toujours $T_{obs} = 60$ ans. Le nombre total d'événements s'élève cette fois-ci à 30 082 avalanches. L'application de la méthode de prétraitement des données conduit toutefois à considérer

4 975 valeurs de a_{jt} comme non observées, soit 41%, des $T_{obs} \times M = 12\,240$ valeurs totales, et à ne retenir ainsi que 21 682 avalanches, soit 72% des événements observés, pour l'inférence du modèle. Cela peut sembler un « sacrifice » important, mais le nombre d'événements disponible pour l'inférence du modèle reste substantiel. De plus, il a été considéré qu'il valait mieux d'un point de vue opérationnel avoir des intervalles de confiance plus larges du fait d'un échantillon moins important que sous-estimer les fréquences avalancheuses moyennes utiles à la prédétermination. Signalons également que les nombres attendus e_j ont été recalculés en pondérant dans chaque commune les valeurs précédemment utilisées au prorata des nombres d'années d'observation manquantes.

Il n'a pas été mentionné jusqu'ici que l'évaluation du nombre de sites par commune pouvait également poser problème, celui-ci ayant été susceptible de varier au cours du temps. En effet, le nombre de sites EPA en cours d'observation dans chaque commune en 2007 est connu précisément grâce à la rénovation du protocole d'observation effectuée entre 2002 et 2006. Mais un nombre de sites plus élevé a pu être observé dans certaines communes au cours des 60 années précédentes sans que l'on puisse le savoir avec certitude et encore moins quantifier de façon exacte les variations au cours du temps du nombre de sites observés dans chaque commune. Il a de ce fait été décidé que l'on ne pouvait guère faire mieux que considérer le nombre de sites de chaque commune comme constant en le prenant égal pour chaque commune au maximum entre le nombre de sites ayant enregistré au moins un événement une année donnée de la période 1946-2005 et le nombre de sites observés en 2006. On obtient ainsi 2166 sites, avec de 1 à 91 sites par commune.

Comme précédemment, l'inférence a été effectuée par paquets de 5 ans puisque, pour l'instant, le problème de la variabilité interannuelle trop forte pour un modèle poissonien homogène n'a pas été traité. Durant l'inférence, les observations manquantes sont traitées comme des paramètres supplémentaires du modèle, c'est-à-dire que leur loi *a posteriori* sachant les données disponibles est calculée. L'inférence a également été effectuée sur ces deux départements sans prétraitement des données de façon à montrer l'effet de la méthode de prétraitement proposée.

Tableau 5.5 : Rapport spatial pour les différents modèles (Savoie-Haute-Savoie)

	Moy.	SD	q_{2.5}	q_{97.5}
Modèle spatial sans prétraitement des données	0.57	0.04	0.49	0.64
Modèle spatial avec prétraitement des données	0.73	0.04	0.63	0.80
Intégration de covariables dans le modèle	0.32	0.10	0.13	0.53
Non-stationnarité dans le temps	0.65	0.13	0.37	0.87
Modèle avec dépendance spatiale d'ordre 2	0.67	0.16	0.29	0.92
Modèle à 3 régimes	0.74	0.10	0.53	0.91
Modèle d'interaction + 3 régimes	0.83	0.05	0.71	0.91

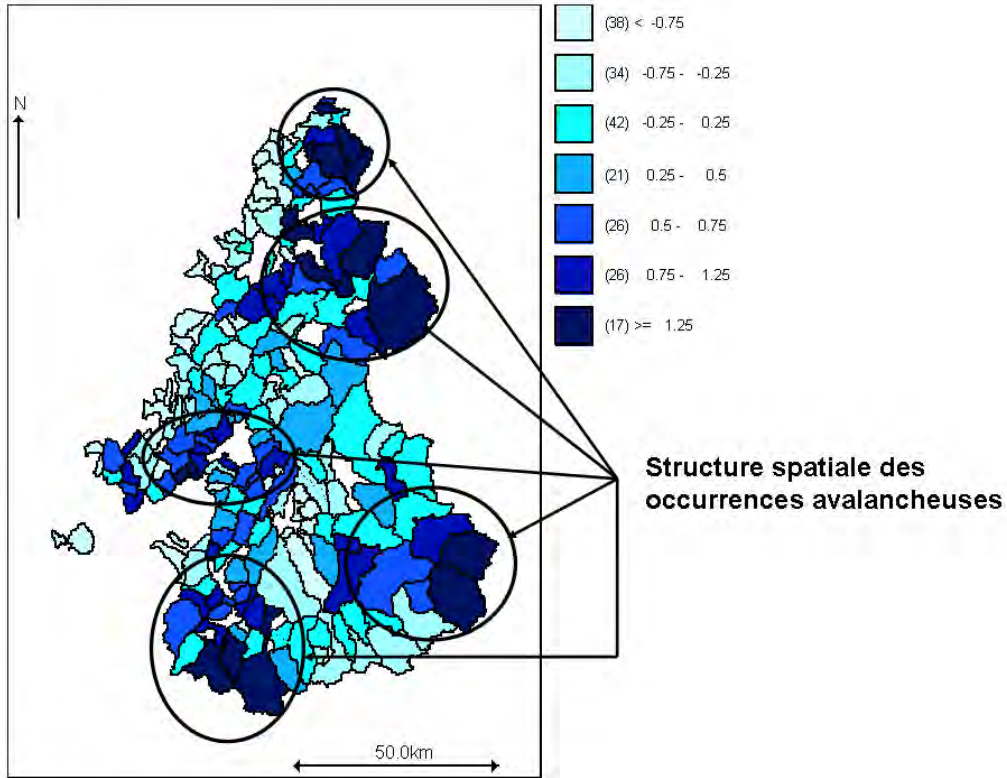
Sans prétraitement, le rapport spatial s'établit pour la région d'étude formée par les deux départements à 0.57 (Tableau 5.5), ce qui est très proche de la valeur de 0.6 qui avait été obtenue précédemment sur la Savoie seule. Après prétraitement, ce rapport s'établit à 0.73, ce qui indique une structure spatiale plus forte par rapport au bruit local. Cela s'explique par le fait qu'on a retiré du bruit en retirant les événements manqués par les observateurs. La composante locale reste bien évidemment non nulle.

Toutefois, on vérifie *a posteriori* que $\alpha_{moy} = \frac{1}{M} \sum_{j=1}^M v_j$ reste nul, c'est-à-dire que la moyenne régionale

des risques relatifs reste 1, même si les nombres d'avalanches observés et attendus ne coïncident plus tout à fait.

La carte des u_j confirme que la procédure de prétraitement a permis un filtrage des observations. La structure spatiale des occurrences avalancheuses (Figure 5.15) apparaît en effet plus clairement qu'avant, avec un contraste intercommunal plus fort. Dans le détail, on retrouve les trois groupes de communes savoyardes des Bauges, de la Vanoise et des Aiguilles d'Arves précédemment mentionnées comme ayant une composante spatiale importante. Pour la Haute-Savoie, le Chablais (à l'extrême nord) ainsi que la Vallée du Giffre et le massif du Mont-Blanc ressortent. Dans les autres communes des deux départements, on observe des composantes spatiales relativement basses et homogènes.

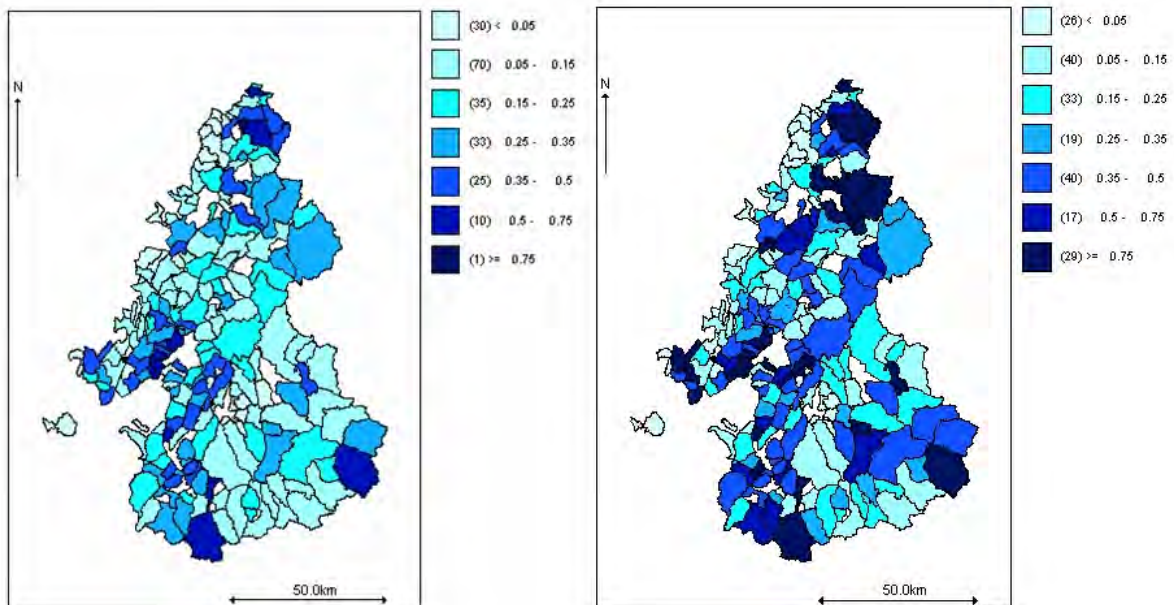
Figure 5.15 : Composante spatiale après prétraitement des données



Modèle spatial de base appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données.

Si l'on compare les moyennes prédictives des fréquences avalanches annuelles f_j^{pred} avec et sans prétraitement des données, l'apport du prétraitement apparaît encore plus nettement (Figure 5.16). Avant le traitement des données, une seule commune affichait une fréquence annuelle supérieure à 0.75 avalanches par couloir et par an. Après prétraitement, 29 communes présentent une fréquence supérieure à 0.75 avalanches par couloir et par an. Elles se situent à peu près toutes dans les cinq massifs énoncés précédemment. En retirant les observations manquantes et en rectifiant le nombre de sites par commune, on évite donc bien de sous-estimer les valeurs utiles à la prédétermination.

Figure 5.16 : Moyenne prédictive sans et avec prétraitement des données



Modèle spatial de base appliqué à l'ensemble Savoie-Haute-Savoie.

5.3.2. Intégrer des covariables nivo-météorologiques

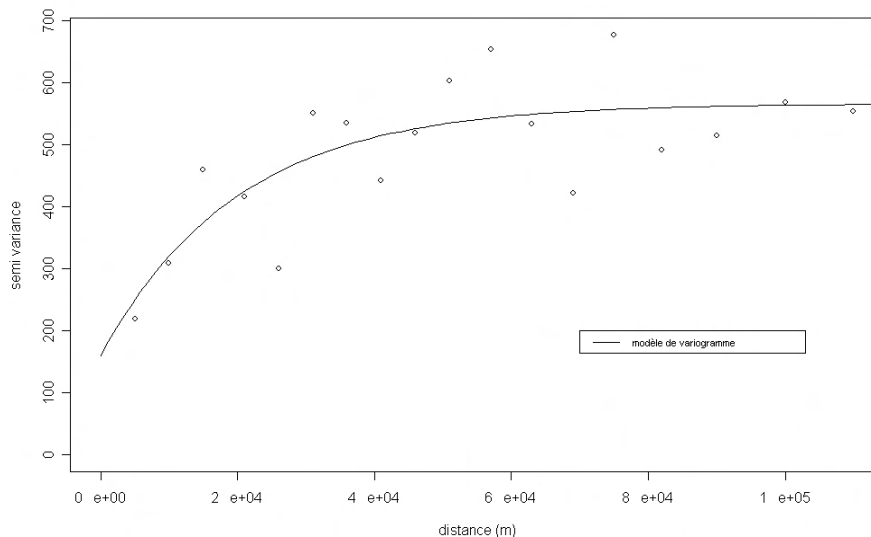
5.3.2.1. Obtention d'une covariable nivo-météorologique à l'échelle communale

Expliquer les variations spatiales de l'activité avalancheuse par les variations de l'intensité des précipitations neigeuses est tentant mais difficile, l'intégration d'une covariable nivo-météorologique à l'échelle communale posant un problème d'hétérotopie. En effet, les précipitations sont mesurées ponctuellement sur un réseau de pluviomètres ne correspondant évidemment pas à la grille irrégulière définissant les limites communales. Une méthode assez fruste a donc été utilisée afin de calculer une covariable représentant à l'échelle communale l'intensité du forçage nivo-météorologique. Etant donné que l'activité avalancheuse est supposée d'après la connaissance des nivologues être bien corrélée aux fortes précipitations, c'est la précipitation décennale à l'échelle communale qui a été retenue.

Pour ce faire, on est parti des hauteurs de précipitations hivernales cumulées sur 3 jours correspondant à une période de retour de 10 ans sur chacun des 78 pluviomètres du réseau Météo France dont on disposait dans les Alpes. Les pluviomètres sont situés à des altitudes comprises entre 291 et 2010 m. Les précipitations hivernales couvrent la période de la mi-novembre à la mi-mai. Bien qu'exprimées en mm d'eau, elles sont supposées se produire sous forme de neige à haute altitude. Les séries de hauteurs de précipitation maximales hivernales sur ces 78 postes ont été achetées à Météo France par le Cemagref. Les précipitations décennales ont été obtenues par ajustement d'une loi GEV sur chacune des séries. Elles ont toutes été ramenées à une altitude de 2500 m grâce à l'application d'un gradient orographique hivernal, avec une valeur de gradient par massif choisie d'après Durand et al. (2006). Ensuite, un terme constant et un gradient linéaire nord/sud ont été soustraits aux « observations » et les résidus ont permis l'ajustement d'un variogramme empirique. Enfin les valeurs des précipitations décennales à 2500 m ont pu être prédites au niveau des centroïdes des communes étudiées par krigeage. Une telle méthode séparant l'inférence du modèle de valeurs extrêmes de celle du modèle spatial est assez « laide » d'un point de vue statistique, mais il a été estimé qu'étant donné les nombreuses autres approximations de la méthode, sa précision était suffisante sachant que, de toute façon, seule une covariable assez grossière était recherchée.

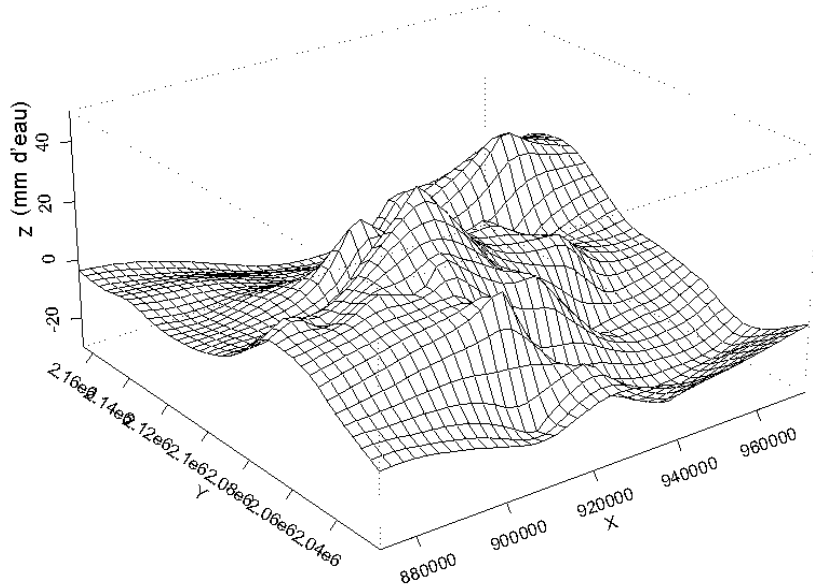
La Figure 5.17 présente l'ajustement du variogramme empirique aux résidus de précipitations décennales hivernales à 2500 m d'altitude. Le modèle de variogramme retenu a été de type exponentiel avec un effet de pépite significatif. Même si l'ajustement est loin d'être parfait, l'existence d'une portée de l'ordre de 60 km apparaît assez clairement. Celle-ci correspond plutôt bien à l'échelle de distance des massifs montagneux des Alpes. La Figure 5.18 présente la prédiction par krigeage sur une grille régulière correspondant à la région étudiée. La valeur de la variable prédite présente une amplitude de variation de l'ordre de 40 mm d'eau. Les « creux » et les « bosses » représentent la structure spatiale du forçage météorologique.

Figure 5.17 : Ajustement d'un variogramme exponentiel sur les précipitations décennales



Les données pluviométriques sont issues du réseau Météo France dans les Alpes françaises. Le variogramme empirique est obtenu à partir des précipitations décennales ramenées à 2500 m et auxquelles on a retranché un terme constant et un gradient linéaire nord/sud

Figure 5.18 : Prédiction de la hauteur de précipitation décennale par krigeage

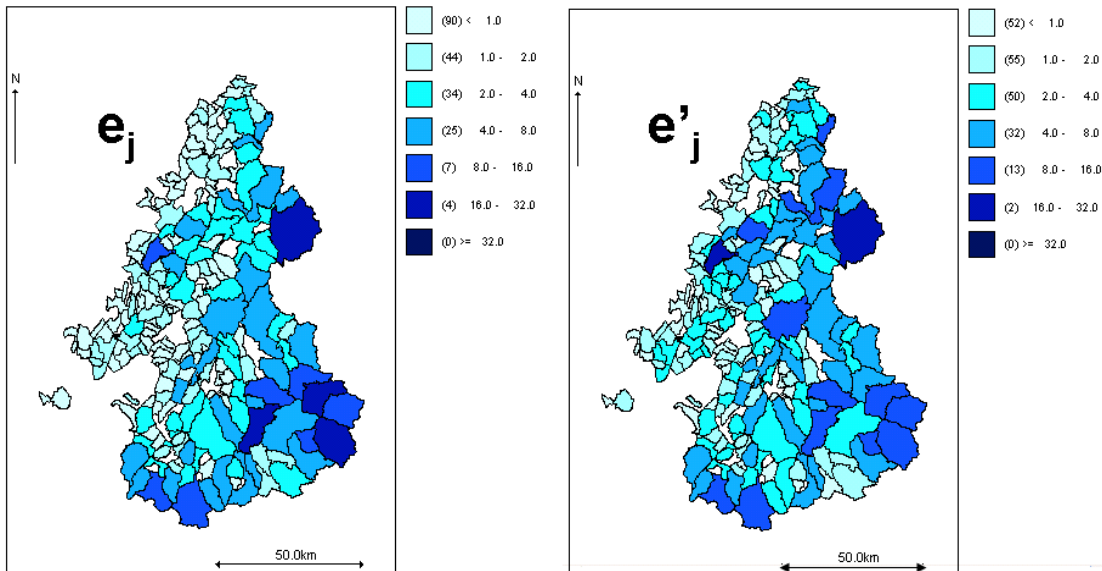


(x, y) sont les latitudes et longitudes.
 La surface de prédiction z correspond au résidu uniquement. Pour la prédiction au niveau des centroïdes des communes, on lui rajoute le terme constant et le gradient linéaire nord/sud retranchés aux observations avant ajustement du variogramme.

5.3.2.2. Nouveaux nombres d'avalanches attendus

Outre la précipitation décennale notée $neige_{T_{10}}$, une covariable topographique a été introduite dans le modèle afin de représenter la structure des massifs montagneux. Après plusieurs essais avec l'altitude minimale, l'altitude maximale et la superficie, c'est l'altitude moyenne z_{moy} de chaque commune qui s'est avérée la plus significative.

Figure 5.19 : Nombres attendus avant et après introduction des covariables



Les covariables sont l'altitude moyenne et la précipitation décennale. Application à l'ensemble Savoie-Haute-Savoie.

L'introduction des covariables dans le modèle de risque relatif s'effectue sous la forme d'une régression avec deux paramètres l_1 et l_2 quantifiant la force de la corrélation entre les covariables et l'activité avalancheuse (eq. 5.17). Cela revient en fait à utiliser des nombres attendus tenant compte non seulement du nombre de sites de chaque commune, mais aussi de sa topographie et de l'intensité du forçage climatique hivernal. Plutôt que de réestimer à chaque fois les nombres attendus, on utilisera donc par la suite les nombres e'_j donnés par l'eq. 5.18.

$$\ln(\lambda_j) = \ln(e_j) + l_1 \times neige_{T_{10j}} + l_2 \times z_{moyj} + u_j + v_j \quad (\text{eq. 5.17})$$

$$\ln(\hat{e}_j) = \ln(e_j) + \hat{l}_1 \times neige_{T_{10j}} + \hat{l}_2 \times z_{moyj} \quad (\text{eq. 5.18})$$

Les estimateurs obtenus pour l_1 et l_2 sont significativement positifs au seuil de 95%, on attend donc plus d'avalanches lorsque l'altitude et les précipitations augmentent, ce qui est logique. La Figure 5.19 compare les nombres attendus avant et après introduction des covariables. Dans les deux cas, on reconnaît la structure en massifs montagneux, avec plus d'avalanches attendues à l'est, c'est-à-dire là où sont situés les sommets les plus élevés. Cette structure est donc déjà bien traduite par les nombres de sites suivis qui permettent le calcul des e_j . Néanmoins, introduire les covariables n'est quand même pas sans effet puisque l'on attend désormais davantage d'avalanches dans certaines communes, notamment dans le Chablais et la vallée du Giffre.

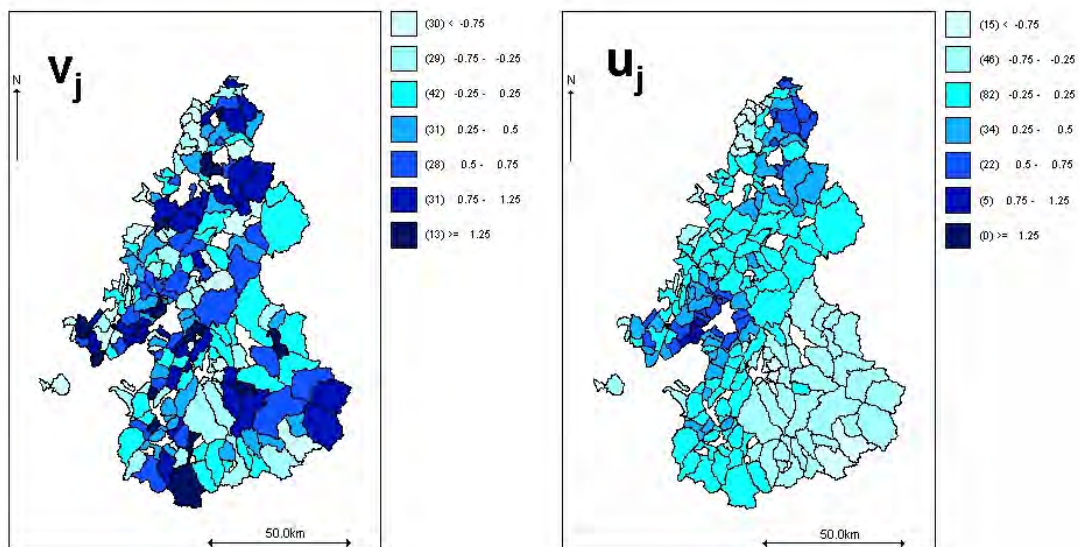
5.3.2.3. Inférence du modèle spatial avec covariables

Après la prise en compte des covariables dans le calcul des nombres attendus, on ne compare plus les observations à une moyenne régionale basée uniquement sur les observations, mais à une moyenne pondérée localement par la topographie et la climatologie. Sur la carte des u_j , il ne reste plus, après intégration des covariables, que deux groupes de communes ayant une composante spatiale importante (Figure 5.20, droite). Il s'agit des communes des Bauges et du Chablais, zones de forte activité mais d'altitude relativement modeste pour lesquelles une activité forte n'est pas prévue par le modèle de régression. Pour les massifs d'altitude élevée, toute la structure spatiale visible sur la Figure 5.15 a été capturée dans les nombres attendus.

Ainsi, même si les \hat{e}_j sont assez peu différents des e_j , l'effet de l'introduction des covariables est assez net. La précipitation décennale et l'altitude moyenne expliquent donc bien une partie de la structure spatiale de l'activité avalancheuse, de sorte que le ratio spatial chute à 0.32 (Tableau 5.5). Toutefois, c'est positif parce que l'information qui était précédemment simplement inférée par le modèle est maintenant explicitée, ce qui augmente le pouvoir prédictif du modèle en le rendant moins dépendant des observations. Le bruit local reste néanmoins élevé (Figure 5.20, gauche). Il nous indique qu'une large information est toujours non exploitée, de sorte qu'il est possible d'améliorer

encore le modèle en proposant une structure plus souple. Enfin $\alpha_{moy} = \frac{1}{M} \sum_{j=1}^M v_j$ reste non significativement différent de zéro même si, à présent, la somme des nombres attendus est différente du nombre total d'avalanches.

Figure 5.20 : Composantes locale et spatiale après introduction des covariables



Modèle spatial de base appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

5.3.3. Introduire un modèle non stationnaire de base

5.3.3.1. Un modèle sans structure temporelle

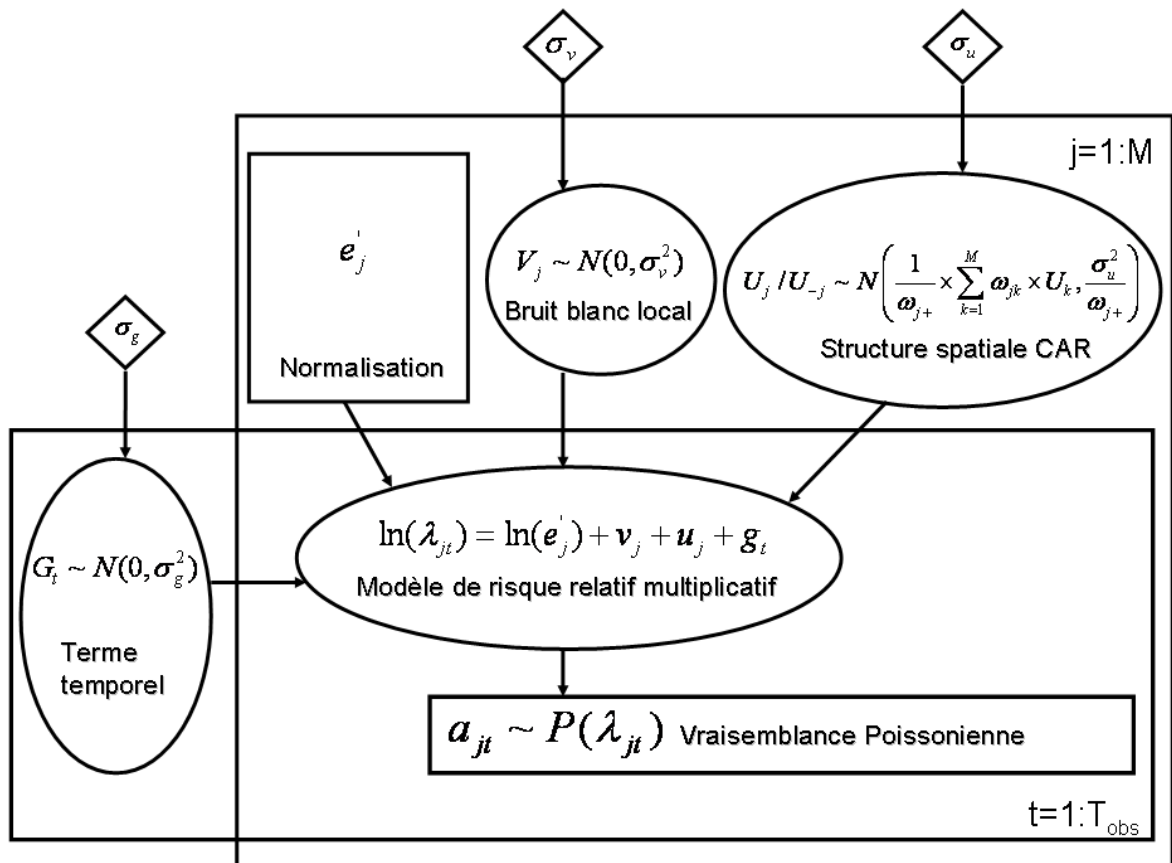
Jusqu'à présent, on n'a modélisé les occurrences avalancheuses qu'à l'échelle de l'espace. Ce faisant on a été obligé de faire l'inférence du modèle par paquets de cinq ans afin de lisser les fortes variations interannuelles de l'activité avalancheuse. On introduit donc à présent un terme temporel supplémentaire g_t dans les risques relatifs (eq. 5.19). Il modélise de possibles variations d'une année sur l'autre des risques relatifs, de sorte que les occurrences avalancheuses sont désormais considérées comme un processus non stationnaire.

$$\ln(\lambda_{jt}) = \ln(e'_j) + u_j + v_j + g_t \quad (\text{eq. 5.19})$$

Dans un premier temps, on considère le terme temporel non structuré sur la période d'observation. On le suppose donc gaussien et de moyenne nulle quelle que soit l'année avec un paramètre de variance σ_g^2 quantifiant la variabilité interannuelle. Enfin, afin d'assurer l'identifiabilité des différents termes,

on fixe $\sum_{t=1}^{T_{obs}} g_t = 0$. Le DAG obtenu permet de visualiser l'ajout de la variable latente G_t et la modélisation du risque relatif comme un champ latent spatio-temporel. Par rapport aux notations introduites au chapitre 3.3, on a donc $\theta_{F_{jt}} = \lambda_{jt}$, $\theta_{F_t} = g_t$, $\theta_{F_j} = (u_j, v_j)$, $\theta_{FR} = \sigma_g$ et toujours $\theta_{FR} = (\sigma_u, \sigma_v)$ (Figure 5.21).

Figure 5.21 : DAG du modèle spatio-temporel de base



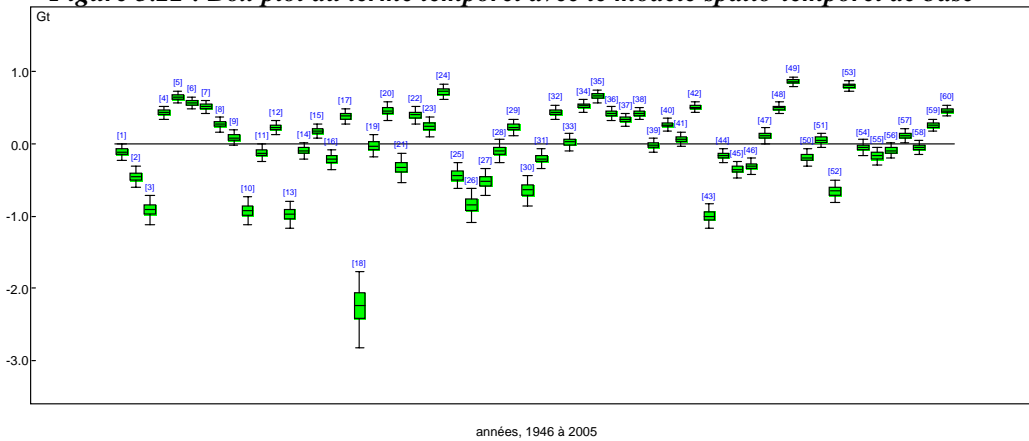
5.3.3.2. Application

On peut à présent considérer des décomptes annuels, soit 60 répétitions du processus. Les variations de g_t permettent de faire ressortir ce qui change d'une année à l'autre mais pas d'une commune à l'autre, c'est-à-dire l'effet annuel global du climat sur la région étudiée. On est donc tenté d'essayer

d'identifier des tendances, des ruptures ou des effets de mémoire pouvant marquer un éventuel changement climatique.

La Figure 5.22 présente les valeurs successives prise par g_t au cours des 60 années étudiées sous la forme d'un box-plot. On rappelle que les chiffres allant de 1 à 60, représentent les années 1946 à 2005. On vérifie que la moyenne sur 60 ans est nulle. Les intervalles de crédibilité à 95% de chaque g_t sont petits, sauf pour g_{18} qui représente un cas extrême, l'année 1963. Cela s'explique par le fait que la modélisation hiérarchique réduit l'intervalle de crédibilité sur chaque g_t en les tirant tous dans la même loi. En conséquence, de nombreuses valeurs de g_t sont significativement différentes les unes des autres, ce qui prouve que les occurrences avalanches sont bien un processus non stationnaire.

Figure 5.22 : Box plot du terme temporel avec le modèle spatio-temporel de base



Modèle spatio-temporel de base appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

Evidemment, les variations de g_t sont fortement liées aux variations du nombre d'événements observés chaque année. g_{18} qui se distingue par une valeur très négative (-2.2 en moyenne) représente ainsi une année (1963), où seulement 5 avalanches ont été observées à l'échelle régionale contre 360 en moyenne sur 60 ans. De même, l'année qui présente le maximum d'avalanches observées est l'année 1994, qui correspond bien à la valeur la plus élevée de g_t . Enfin, les années où g_t est proche de la valeur 0 représentent les années de fréquence avalancheuse moyenne. C'est le cas par exemple des années 1954, 1984 ou 2003.

Graphiquement, on n'a pas l'impression qu'il existe de tendance ou de rupture en moyenne ni même en variance au cours des 60 années. Plus précisément, on trouve une légère tendance linéaire positive égale à +0,0042/an sur 60 ans. Par contre, si l'on ne regarde que les valeurs de g_t des trente dernières années, on y trouve une tendance linéaire négative égale à -0,0059/an. Autant dire que la dérive climatique, si elle existe, ne semble pas très marquée. Son étude mériterait toutefois des investigations complémentaires, en particulier des comparaisons avec d'autres données climatologiques. Quant aux effets de mémoire, ils semblent exister un peu puisque l'on a parfois plusieurs valeurs proches de g_t pour des années successives, comme par exemple entre g_4 et g_9 , c'est-à-dire de 1949 à 1954. Mais l'on a également souvent des fluctuations très brusques avec une valeur faible succédant à une valeur forte, comme par exemple entre 1987 et 1988.

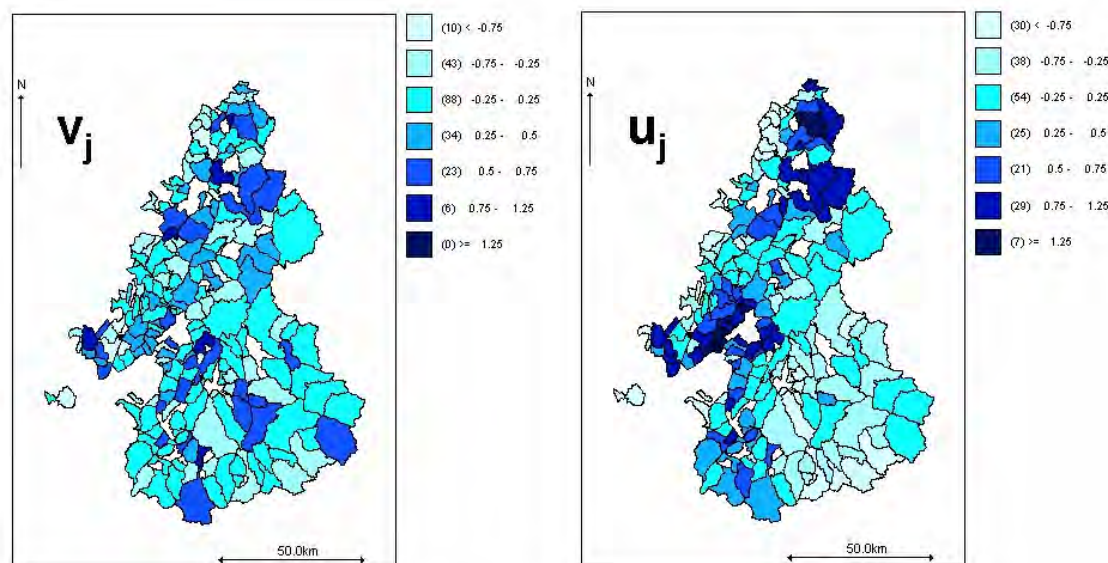
Par rapport au modèle précédent, la carte des u_j reste quasi inchangée car c'est toujours la même structure spatiale qui est capturée. Néanmoins, le fait de retirer la variabilité interannuelle et de travailler avec 60 répétitions plutôt que seulement 12 permet de la visualiser plus nettement en accroissant les contrastes entre les communes à fort u_j et les communes à faible u_j (Figure 5.23, droite). La variabilité capturée par les g_t l'est au détriment du bruit local, de sorte que l'on retrouve des valeurs v_j bien moins importantes qu'avec le modèle précédent (Figure 5.23, gauche). En

conséquence, le rapport spatial remonte très nettement puisqu’il vaut à nouveau 0.65 (Tableau 5.5). La variabilité interannuelle est du même ordre de grandeur que le bruit résiduel et presque trois fois moins forte que la variabilité spatiale (Tableau 5.6)

Tableau 5.6 : Paramètres du modèle spatio-temporel de base (Savoie-Haute-Savoie)

	Moy.	SD	q _{2.5}	q _{97.5}
σ_g	0.56	0.06	0.46	0.68
σ_u	1.51	0.24	1.02	1.97
σ_v	0.66	0.12	0.41	0.9

Figure 5.23 : Composante locale et composante spatiale avec le modèle spatio-temporel de base



Modèle spatio-temporel de base appliqué à l’ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

L’ensemble de ces résultats assez satisfaisants s’explique par le fait qu’il s’agit d’un modèle qui s’ajuste nettement mieux aux données que les modèles précédents grâce à la prise en compte explicite de la variabilité interannuelle. Néanmoins, on reste toujours face à quelques excès locaux indiquant que la composante locale n’a pas disparu (Figure 5.23, gauche). A partir de ce modèle spatio-temporel de base, on va donc essayer d’améliorer la modélisation spatio-temporelle pour tenter de trouver des structures de dépendance mieux adaptées à la variabilité des données qu’un modèle CAR intrinsèque pour la structure spatiale et qu’un modèle de bruit blanc pour la structure temporelle.

5.4. Améliorer la modélisation spatio-temporelle

5.4.1. Différents modèles de dépendance spatiale

5.4.1.1. CAR propre contre CAR intrinsèque

- *Modèle*

Le modèle CAR utilisé jusqu’à présent est impropre au sens où la loi jointe n’existe pas car la matrice de précision est singulière. Il ne peut donc être utilisé que sous forme conditionnelle. On le remplace donc ici par un modèle CAR « propre » pouvant être déclaré indifféremment sous sa forme conditionnelle ou jointe. Un tel modèle possède par rapport au CAR intrinsèque un paramètre γ_{CAR} supplémentaire représentant le degré de dépendance spatiale. Avec le CAR intrinsèque, ce paramètre γ_{CAR} est en fait fixé à son maximum, de sorte qu’un CAR intrinsèque est capable de modéliser correctement uniquement un champ bien spatialisé et que, comme l’a montré l’exemple de la Savoie,

on est donc obligé d'utiliser également un terme de bruit pour modéliser les occurrences avalanches. Avec le CAR propre, γ_{CAR} peut fluctuer en fonction du rapport entre variabilité spatiale et bruit local, de sorte que l'on n'a plus besoin du terme local. On remplace donc la somme $u_j + v_j$ précédemment utilisée par un terme unique w_j représentant à la fois la structure spatiale et la composante locale. Toutefois, on est ainsi amené à considérer explicitement un terme de moyenne régionale α_{moy} afin de permettre éventuellement une légère différence entre les nombres totaux d'événements attendus et observés (eq. 5.20). Ce dernier s'identifie à $\frac{1}{M} \sum_{j=1}^M v_j$ dans le modèle

précédent. On espère trouver une valeur γ_{CAR} proche de son maximum, égal à 1 dans notre cas, parce qu'une valeur faible indiquerait une structure spatiale peu significative.

$$W_j / W_{-j} \sim N \left(\alpha_{moy} + \frac{1}{\omega_{j+}} \times \sum_{k=1}^M \omega_{jk} \times \gamma_{CAR} \times (w_k - \alpha_{moy}), \frac{\sigma_w^2}{\omega_{j+}} \right) \text{ (eq. 5.20)}$$

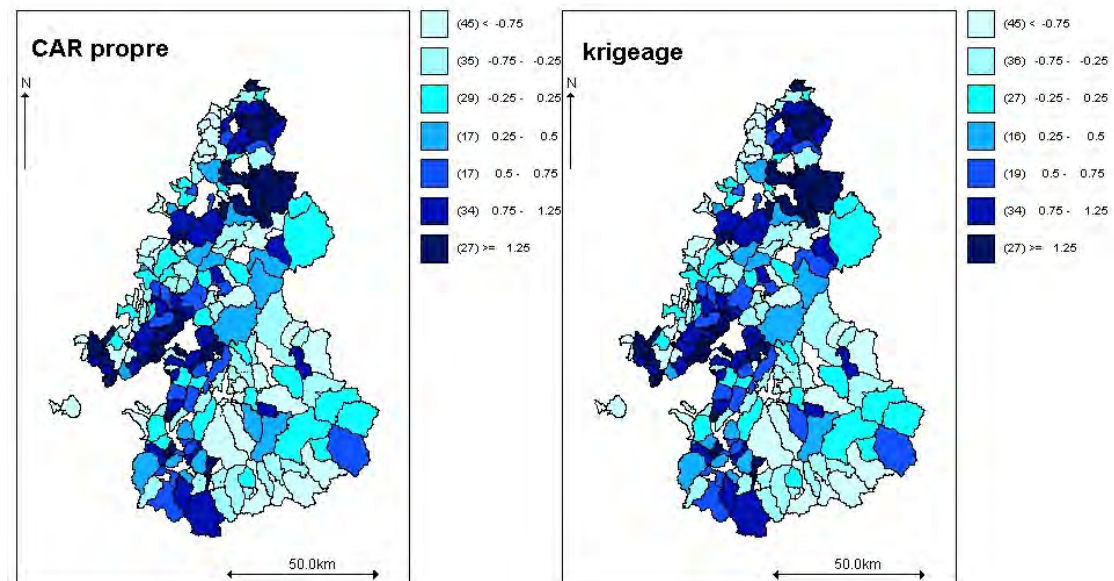
• *Application*

On vérifie que α_{moy} est très faible. La moyenne régionale reste donc proche de zéro avec ce modèle, ce qui est cohérent avec les résultats obtenus précédemment. γ_{CAR} vaut 0.87 et est donc proche de son maximum de 1 (Tableau 5.7). Cela constitue un résultat satisfaisant qui confirme que la structure spatiale reste significative même après l'introduction des covariables. Par contre, en termes de critère DIC, le modèle spatio-temporel de base était un peu meilleur (Tableau 5.10). De plus, avec ce modèle de CAR propre, le bruit n'est plus séparé de la composante spatiale. La Figure 5.24 (gauche) est ainsi une sorte de somme des deux cartes de la Figure 5.23, de sorte que la structure spatiale apparaît moins clairement. On préfère donc conserver la somme $u_j + v_j$ plutôt que le modèle de CAR propre.

Tableau 5.7 : Paramètres globaux de différents modèles (Savoie-Haute-Savoie)

	Moy.	SD	q _{2.5}	q _{97.5}
γ_{CAR} , CAR propre	0.87	0.07	0.7	0.98
ϕ_{exp} , krigeage bayésien	0.23	0.05	0.15	0.35
ρ_{auto} , AR1	0.01	0.13	-0.24	0.26

Figure 5.24 : Composante spatiale et bruit cumulés avec le CAR propre et le modèle de krigage



Modèles appliqués à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

5.4.1.2. Relation de voisinage

- *Modèle*

Jusqu'à présent, on a donné un poids égal à 1 aux communes voisines directes et 0 aux autres. On teste à présent une dépendance d'ordre 2 pour voir si une dépendance spatiale à plus longue portée est susceptible d'améliorer le modèle en facilitant le partage de l'information. On affecte ainsi un poids égal à 2 aux voisins directs dans la matrice ω_{jk} , un poids égal à 1 aux voisins d'ordre 2 et finalement un poids 0 à toutes les autres communes.

- *Application*

En visualisant les cartes des u_j , on ne remarque que de toutes petites différences avec la carte obtenue avec une dépendance à l'ordre 1. La structure spatiale des fréquences d'avalanches n'apparaît donc pas plus nettement que précédemment. En termes de DIC (Tableau 5.10) et de rapport spatial (Tableau 5.5), on ne voit également aucun gain. L'introduction d'une dépendance spatiale de portée plus longue ne semble donc pas améliorer le modèle, ce qui laisse à penser que la structure spatiale est à courte portée. On préfère donc conserver le voisinage d'ordre 1 précédemment utilisé. Pour progresser, il faudrait sans doute essayer une pondération plus physique faisant intervenir les effets topographiques, par exemple imposer une corrélation plus forte entre les communes d'une même vallée et au contraire imposer une corrélation nulle entre des communes certes voisines mais séparées par une ligne de crête les soumettant à des forçages climatiques totalement différents.

5.4.1.3. Un modèle de krigeage

- *Modèle*

On essaye à présent un modèle de krigeage bayésien permettant, comme le modèle de CAR propre, la modélisation à la fois de la structure spatiale et du bruit local. On le note donc également w_j . La structure spatiale y est modélisée par un variogramme exponentiel avec un paramètre ϕ_{exp} qui représente l'inverse de la portée, cette dernière étant la distance limite jusqu'à laquelle la proximité géographique influe sur la corrélation. On remplace ainsi une dépendance spatiale basée sur le voisinage par une dépendance spatiale basée sur la distance inter communes. La matrice de voisinage est donc remplacée dans le modèle par la matrice des distances inter centroïdes.

- *Application*

L'inférence sur ϕ_{exp} est difficile avec un *prior* très peu informatif. La stratégie préconisée par la littérature (Diggle et Ribeiro, 2007) et qui a été employée est de choisir une distribution uniforme avec des bornes inférieure et supérieure permettant d'obtenir une corrélation de 0 à 1 pour toute distance comprise entre les distances intercentroïdes minimale et maximale présentes dans le jeu de données, soit $\phi_{\text{exp}} \sim \text{Unif}(0.032, 4.06)$. Le *posterior* obtenu présente un mode bien marqué très différent du *prior*. Sa moyenne *a posteriori* s'établit à 0.23 (Tableau 5.7), ce qui correspond à une portée d'environ $1/\phi_{\text{exp}} = 7$ kilomètres. Cette faible valeur indique que l'information n'est partagée qu'entre les communes dont les centroïdes sont proches. Elle confirme donc le résultat apporté par le modèle de dépendance d'ordre 2 d'une portée courte de la structure spatiale des occurrences avalancheuses.

Sur la carte des w_j (Figure 5.24, droite), on voit clairement, comme pour le modèle CAR propre, une combinaison des cartes u_j et v_j du modèle spatio temporel de base. La carte obtenue est même très similaire à celle obtenue avec le CAR propre, les mêmes massifs ressortant toujours. Cela indique que le passage d'une structure de dépendance de type voisinage à une structure basée sur les distances n'affecte pas beaucoup la structure spatiale. Le critère DIC augmente quant à lui de quelques points (Tableau 5.10), ce qui indique que l'ajustement à la variabilité des données se fait un peu plus difficilement. On conserve donc plutôt le modèle spatio-temporel de base, d'autant que, dans notre cas, l'utilisation d'un modèle de krigeage est un peu artificielle, puisqu'il est appliqué aux centroïdes des communes et non pas aux points où s'effectuent réellement les observations. Un tel modèle

pourrait cependant être intéressant pour, par la suite, travailler couloir par couloir plutôt qu'à l'échelle communale.

5.4.2. Trois régimes plutôt que des effets de mémoire

5.4.2.1. Modèle autorégressif d'ordre 1

- *Modèle*

L'analyse du box plot de g_t pour le modèle spatio-temporel de base (Figure 5.22) a laissé à penser que g_t pouvait dépendre de g_{t-1} , c'est-à-dire que la valeur du forçage climatique d'un hiver pouvait ressembler à celle de l'année précédente. En d'autres il semblerait que l'on puisse identifier des fluctuations du forçage climatique régional à une échelle de temps de quelques années. On teste donc un modèle autorégressif d'ordre 1, i.e. $G_t \sim N(\rho_{auto} \times G_{t-1}, \sigma_g^2)$. Par rapport au modèle spatio-temporel de base, ce modèle est caractérisé par l'apparition du paramètre d'autorégression ρ_{auto} .

Toutefois il n'introduit pas de paramètre supplémentaire car on impose $\sigma_g^2 = \frac{1}{1 - \rho_{auto}^2}$ afin de garantir la faisabilité de l'inférence.

- *Application*

Le résultat obtenu est très décevant puisque la valeur inférée de ρ_{auto} est très proche de zéro (Tableau 5.7). La raison en est qu'une valeur de coefficient de corrélation plus élevée contraint le bruit σ_g^2 à être très faible, ce qui ne permet pas d'ajuster les fluctuations parfois brusques d'une année sur l'autre de g_t . Il ne semble donc pas y avoir de mémoire d'un hiver à l'autre facilement identifiable au cours des 60 années d'observation mais plutôt des changements brusques d'une année à l'autre. Un prolongement de l'étude avec des modèles de séries temporelles plus élaborés serait cependant intéressant.

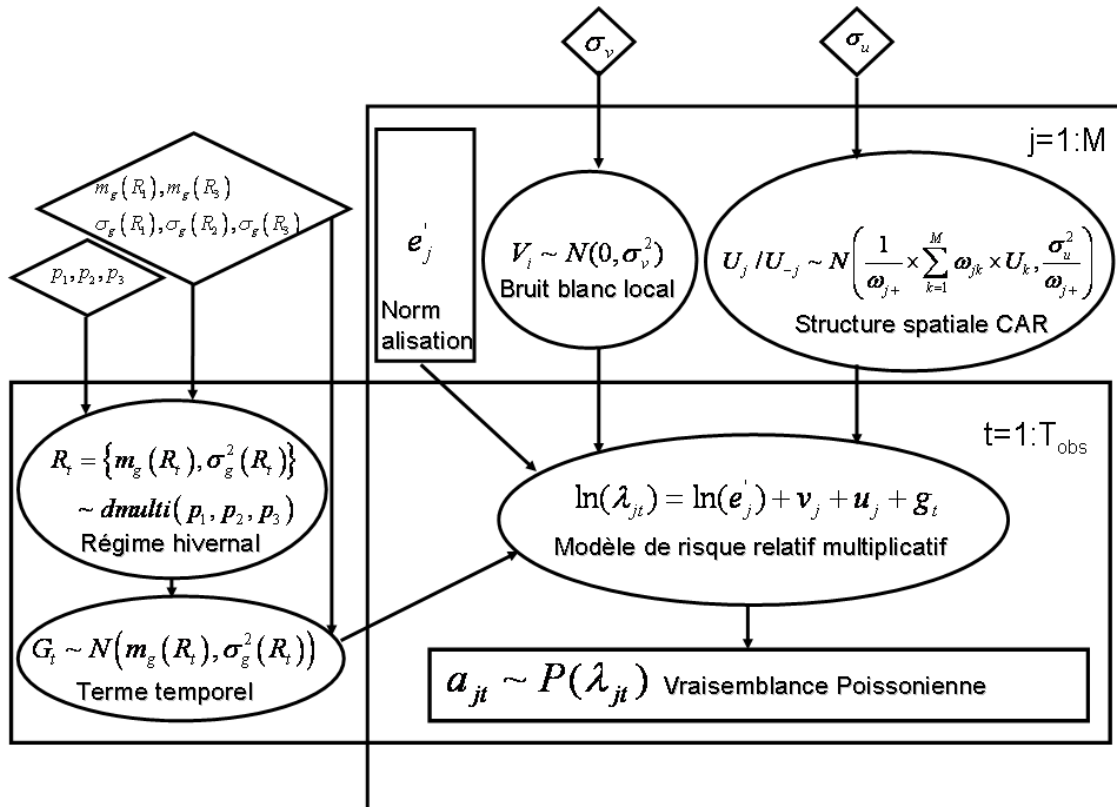
5.4.2.2. Modèle à trois régimes

- *Modèle*

Puisque des variations brutales d'une année sur l'autre sont observées, on suppose maintenant que g_t peut être issu de trois lois différentes représentant trois régimes de forçage climatique bas/moyen/haut. On se propose donc de distinguer les années d'activité exceptionnellement faible des années d'activité intense ou moyenne. On note R_t le régime bas, moyen ou haut de l'année t et l'on suppose $G_t \sim N(m_g(R_t), \sigma_g^2(R_t))$. Pour le régime bas, on impose une moyenne négative, pour le régime moyen, une moyenne nulle et pour le régime haut, une moyenne positive. Les variances correspondantes sont toutes trois estimées. Le modèle se ramène au modèle spatio-temporel de base si les moyennes des régimes bas et haut ne sont pas significativement différentes de zéro et si les trois variances sont identiques.

Pour chaque année, le régime R_t est supposé issu d'une loi multinomiale de paramètres (p_1, p_2, p_3) quantifiant les probabilités des 3 régimes au cours de la période d'observation. R_t est donc considéré comme une variable latente et (p_1, p_2, p_3) sont des paramètres du modèle qui sont estimés. La loi *a posteriori* de R_t quantifie la probabilité de l'année t d'appartenir à chacun des 3 régimes. On peut ainsi affecter un régime à chaque année en privilégiant à chaque fois le régime le plus probable. La Figure 5.25 présente le DAG du modèle obtenu.

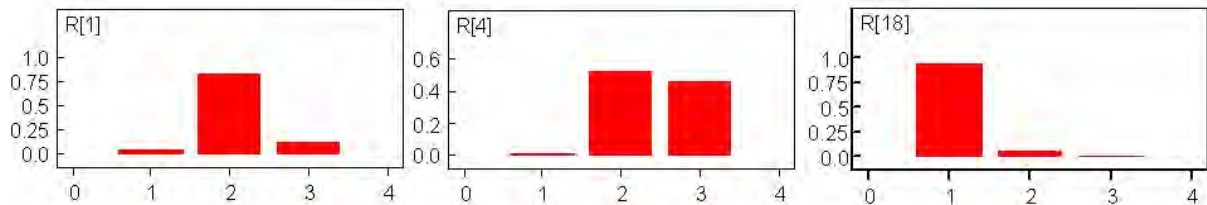
Figure 5.25 : DAG du modèle avec trois régimes



• Application

La Figure 5.26 présente la loi *a posteriori* de R_t pour trois années. L'année 1946 (R_1) est classifiée sans ambiguïté dans le régime moyen. De même, l'année 1963 (R_{18}) dont on avait déjà souligné le caractère exceptionnellement faible est classifiée très fermement dans le régime bas. L'année 1949 (R_4) est plus difficile à classifier avec un maximum de probabilité d'appartenir au régime moyen mais également une probabilité importante de se situer dans le régime haut. Plus généralement, si l'on choisit à chaque fois le régime le plus probable, toutes les années sont affectées au régime moyen sauf 1963. Mais, pour un nombre assez important d'années, le régime haut est, comme pour 1949, assez probable.

Figure 5.26 : Lois *a posteriori* des régimes annuels pour le modèle à trois régimes

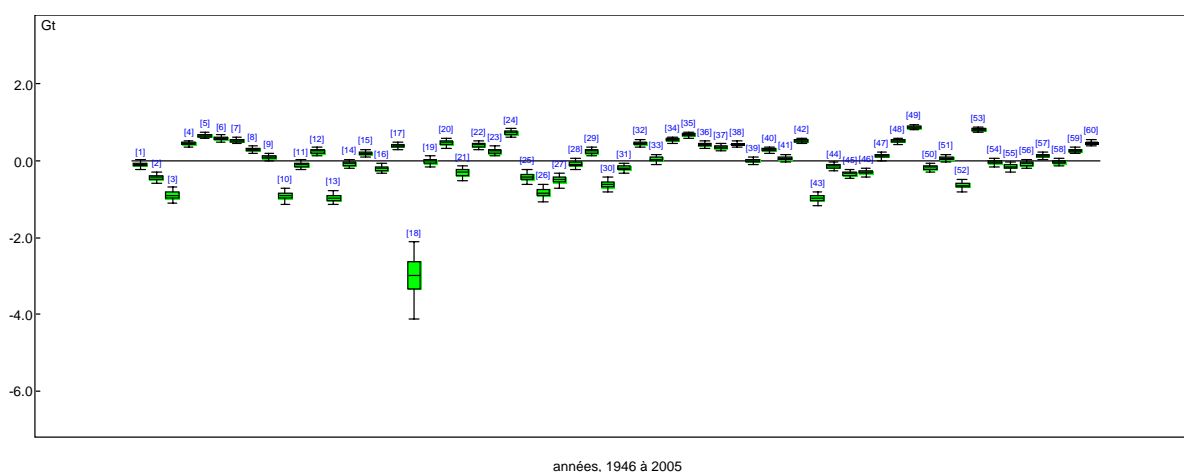


Modèle spatio-temporel à trois régimes appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

Les moyennes *a posteriori* des probabilités globales de chacun des trois régimes reflètent cet état de fait puisqu'elles montrent que 67 % des années appartiennent au régime moyen, 24% au régime haut et seulement 9% au régime bas. Néanmoins, ces probabilités sont difficiles à estimer précisément puisque les intervalles de crédibilité obtenus sont très larges (Tableau 5.8). Les moyennes des trois régimes sont significativement différentes, mais la moyenne du régime bas est beaucoup plus éloignée du régime moyen que celle du régime haut. Enfin, la variance du régime bas est nettement plus faible que celle des deux autres régimes. Ces effets sont dus à la valeur exceptionnelle de 1963 qui pousse la moyenne du régime bas vers le bas et empêche que plus d'une année ne soit classifiée dans ce régime.

L'intérêt d'avoir une variance différente pour chacun des régimes apparaît sur la Figure 5.27. Par rapport à la Figure 5.22 présentant les résultats du modèle spatio-temporel de base, les moyennes *a posteriori* des g_t ne sont pas beaucoup modifiées mais les intervalles de crédibilité sont plus faibles car trois paramètres de variance offrent une plus grande souplesse pour ajuster les données. Le Tableau 5.8 confirme d'ailleurs que les trois paramètres de variance sont plus faibles que le paramètre σ_g global du modèle spatio temporel de base. On peut donc avec ce modèle à trois régimes quantifier le forçage climatique correspondant à chaque année plus précisément.

Figure 5.27 : Box plot du terme temporel avec le modèle à 3 régimes



Modèle spatio-temporel à trois régimes appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

Bien sûr, en termes de structure spatiale et de carte des u_j , on n'enregistre aucun changement. Par contre, on capture un peu plus de variabilité dans les g_t en rendant le modèle plus souple, de sorte que par rapport au modèle spatio-temporel de base σ_v diminue un peu (Tableau 5.8). Le rapport entre variabilité locale et variabilité spatiale passe donc de 0.65 à 0.74 (Tableau 5.5). Enfin, au niveau du critère DIC, une amélioration très faible est observée (Tableau 5.10).

Tableau 5.8 : Paramètres du modèle à trois régimes (Savoie-Haute-Savoie)

	Moy.	SD	q _{2.5}	q _{97.5}
p_1	0.09	0.12	0	0.44
p_2	0.67	0.25	0.07	0.97
p_3	0.24	0.19	0.01	0.81
$m_g(R_1)$	-2.28	1.29	-4.21	-0.17
$m_g(R_3)$	0.49	0.67	0.11	1.3
$\sigma_g(R_1)$	0.42	0.5	0.08	1.43
$\sigma_g(R_2)$	0.46	0.26	0.13	0.96
$\sigma_g(R_3)$	0.23	0.15	0.08	0.52
σ_u	1.66	0.2	1.27	2.07
σ_v	0.57	0.12	0.34	0.79

Au final, les fluctuations interannuelles, bien que fortes, ne le semblent pas suffisamment pour clairement identifier un régime haut et l'année 1963 apparaît plus comme étant réellement exceptionnelle que comme une représentante d'un régime bas particulier. Néanmoins, ce modèle

permet de diminuer le bruit résiduel, et, ce faisant, d'affiner les résultats précédemment obtenus. De plus, il s'avérera intéressant pour traiter des jeux de données présentant différents régimes plus marqués. On le conserve donc pour la suite, tout en sachant qu'un modèle avec un seul régime, voire vraisemblablement deux, fournit également des résultats satisfaisants avec ce jeu de données.

5.4.3. Structure spatio-temporelle du second ordre

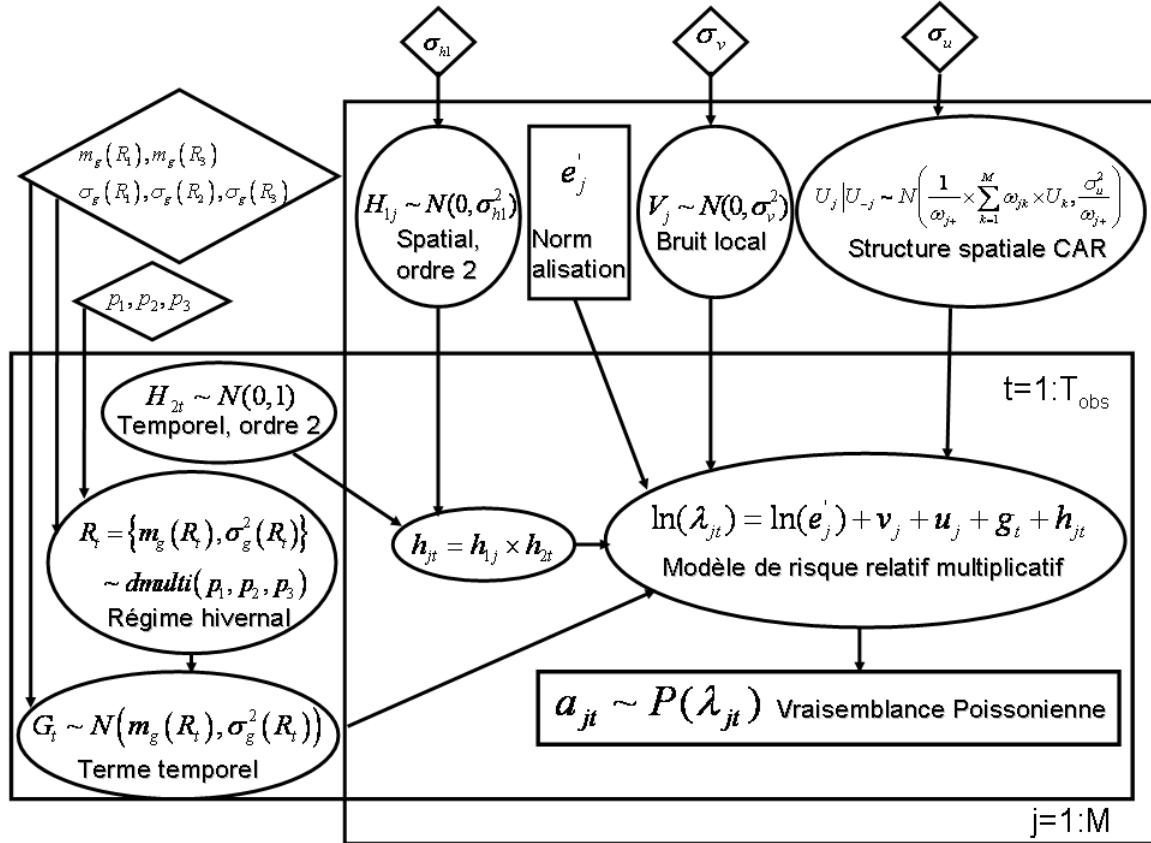
5.4.3.1. Modèle

Dans ce modèle, on ajoute au modèle précédent un terme spatio-temporel croisé h_{jt} dans le modèle de risque relatif. On le suppose, conformément au schéma proposé au chapitre 3.3, décomposable en un terme spatial du second ordre h_{1j} et un terme temporel du second ordre h_{2t} , ce qui s'apparente sous

les contraintes $\sum_{t=1}^{T_{obs}} h_{2t} = 0$ et $\sum_{j=1}^M h_{1j} = 0$ à un modèle d'Anova à deux facteurs (eq. 5.21).

$$\ln(\lambda_{jt}) = \ln(e'_j) + u_j + v_j + g_t + h_{jt} \quad (\text{eq. 5.21})$$

Figure 5.28 : DAG du modèle final avec trois régimes et structure du second ordre



On n'impose pas de structure sur les deux champs latents introduits, de sorte que l'on suppose les valeurs locales tirées dans des normales indépendantes. De plus, on se heurte à des problèmes d'identifiabilité, de sorte que l'on est obligé de fixer complètement les paramètres d'au moins un des deux champs latents introduits. On prend donc $H_{2t} \sim N(0, 1)$ et $H_{1j} \sim N(0, \sigma_h^2)$, avec un seul paramètre supplémentaire σ_h quantifiant la variabilité du champ spatial du second ordre à estimer. Sa valeur dépend toutefois de la variance fixe de h_{2t} , de sorte que la variabilité du second ordre est mieux

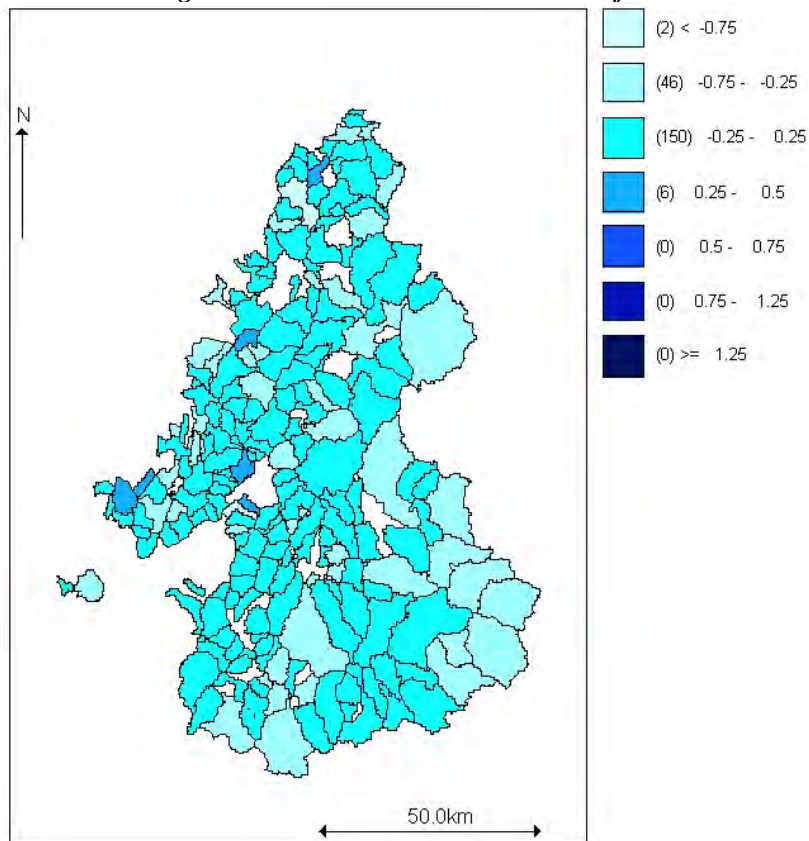
quantifiée par $\sigma_h = \sqrt{\text{Var}(H)}$ avec $\text{Var}(H) = \frac{1}{M \times T_{obs}} \times \sum_{j=1}^M \sum_{t=1}^{T_{obs}} \left(\left(h_{jt} - \frac{1}{M \times T_{obs}} \times \sum_{j=1}^M \sum_{t=1}^{T_{obs}} h_{jt} \right)^2 \right)$.

Notons également qu’une indétermination de signe irréductible existe, de sorte que l’on peut interpréter seulement la significativité des termes h_{1j} et h_{2t} , ainsi que leurs valeurs absolues. Le terme croisé h_{jt} est par contre quant à lui tout à fait interprétable comme un excès ou un déficit local dans le temps et l’espace. Par rapport aux notations introduites au chapitre 3.5, on a donc à présent $\theta_{F_{jt}} = (\lambda_{jt}, h_{jt})$, $\theta_{F_t} = (g_t, h_{2t}, R_t)$, $\theta_{F_j} = (u_j, v_j, h_{1j})$, $\theta_{FR} = (\sigma_u, \sigma_v, \sigma_{h1})$ et $\theta_{FT} = (m_g(R_1), m_g(R_3), \sigma_g(R_1), \sigma_g(R_2), \sigma_g(R_3))$, soit par rapport au modèle spatio-temporel de base, cinq paramètres, un champ spatial et deux champs temporels supplémentaires (Figure 5.28).

5.4.3.2. Application

L’ajout du terme croisé dans notre modèle est justifié, car un nombre très important de termes h_{jt} sont significativement différents de 0. De même, pour la composante spatiale du second ordre h_{1j} , 85 valeurs sur 204 sont significatives, soit un peu plus que 40 %. La conséquence en est la quasi disparition du bruit local v_j , avec des valeurs toutes inférieures à 0.5 et pour la plupart très proche de 0 (Figure 5.29). Le bruit local est donc efficacement capturé par le terme croisé, et ce alors que la structure spatiale du premier ordre u_j ne bouge pas (Figure 5.30, gauche). Le rapport entre variabilité locale et variabilité spatiale atteint donc plus de 80 % (Tableau 5.5). La variabilité du second ordre σ_h est du même ordre de grandeur que celle du bruit résiduel σ_v , soit presque quatre fois moins que la variabilité spatiale σ_u (Tableau 5.9). A noter que la variabilité du champ spatial du second ordre σ_{h1} est identique à σ_h car celle de la structure temporelle du second ordre est fixée à 1 (Tableau 5.9). Quant au DIC, il chute fortement (Tableau 5.10). Ces franches améliorations indiquent que le modèle plus souple proposé ajuste bien mieux la variabilité des données.

Figure 5.29 : Bruit local avec le modèle final



Modèle spatio-temporel à trois régimes et structure du second ordre appliqué à l’ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

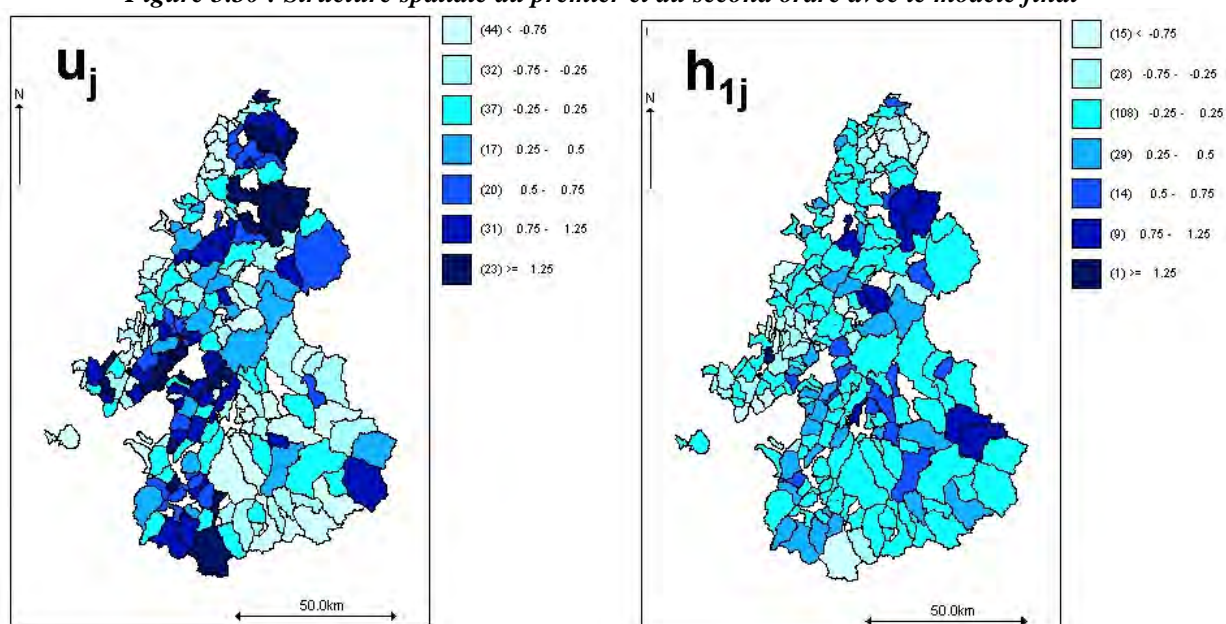
Tableau 5.9 : Paramètres du modèle final (Savoie-Haute-Savoie)

	Moy.	SD	q_{2.5}	q_{97.5}
p_1	0.18	0.14	0.01	0.53
p_2	0.53	0.26	0.05	0.94
p_3	0.29	0.21	0.02	0.81
$m_g(R_1)$	-1.14	0.81	-3.14	-0.16
$m_g(R_3)$	0.52	0.47	0.11	0.91
$\sigma_g(R_1)$	0.64	0.35	0.1	1.37
$\sigma_g(R_2)$	0.36	0.18	0.11	0.7
$\sigma_g(R_3)$	0.26	0.15	0.09	0.57
σ_u	2.05	0.18	1.7	2.4
σ_v	0.52	0.08	0.37	0.69
σ_{h1}	0.55	0.06	0.44	0.69
σ_h	0.55	0.03	0.49	0.60

Les termes croisés sont difficiles à interpréter parce qu'ils dépendent à la fois du temps et de l'espace et que l'on se heurte aux problèmes d'indétermination précédemment soulignés. Les relier à des variations spatiotemporelles de l'activité avalancheuse nécessiterait d'aller sur le terrain et rechercher pourquoi un déficit ou un excès local est observé une année donnée.

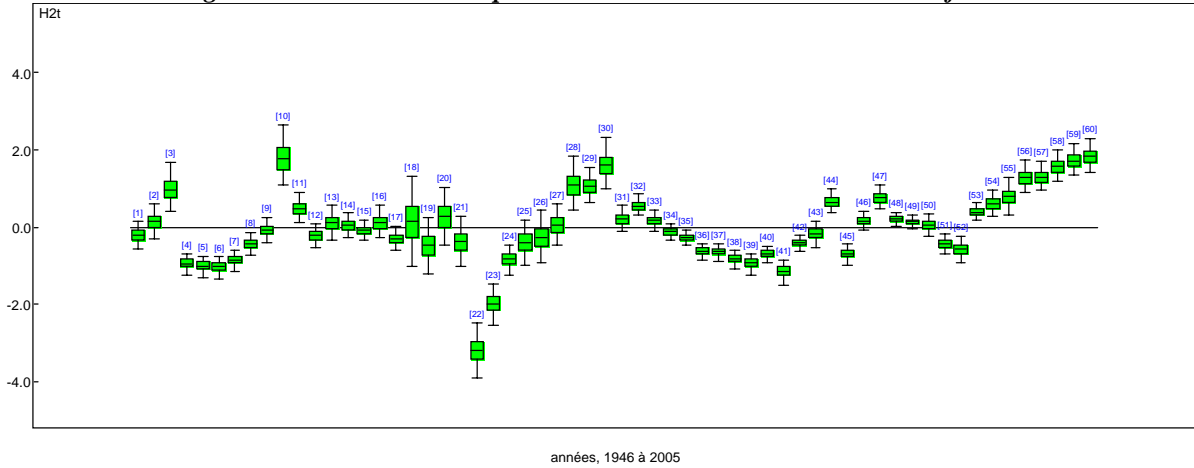
Concernant le terme temporel du second ordre (Figure 5.31), on a visiblement des effets de mémoire et on pourrait ajuster un modèle autorégressif d'ordre 1 plus facilement que sur le terme temporel g_t . Cette corrélation correspond vraisemblablement à l'effet d'observation non pris en compte dans le prétraitement des données. En effet, un observateur reste en place plusieurs années consécutives et on peut penser qu'il collecte l'information locale dont il a la charge de manière analogue au cours des années successives de son service. Quant à la structure spatiale du second ordre (Figure 5.30, droite), elle permet de visualiser les communes où cet effet d'observation s'écarte le plus de la moyenne régionale.

Figure 5.30 : Structure spatiale du premier et du second ordre avec le modèle final



Modèle spatio-temporel à trois régimes et structure du second ordre appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

Figure 5.31 : Structure temporelle du second ordre avec le modèle final

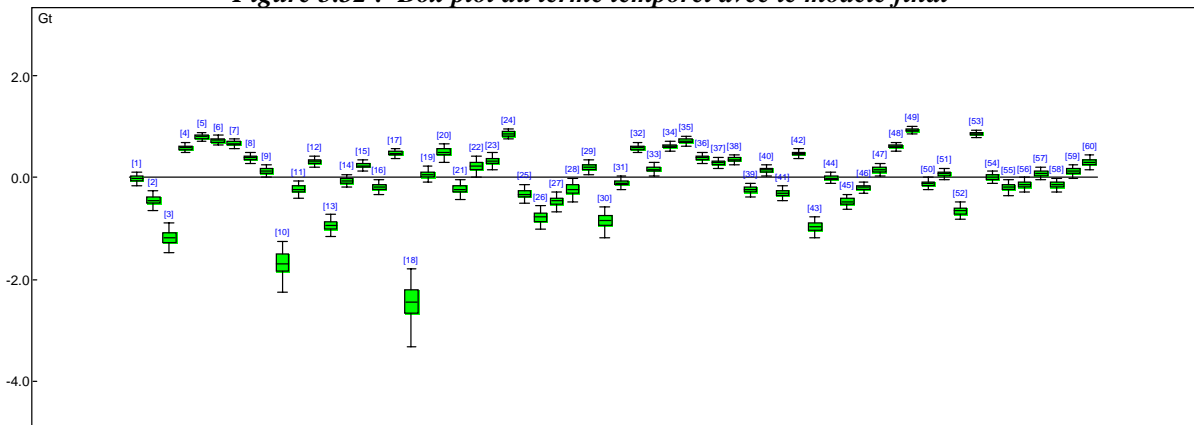


années, 1946 à 2005

Modèle spatio-temporel à trois régimes et structure du second ordre appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

Par rapport au modèle précédent, les g_t sont presque inchangés en moyenne comme en variance (Figure 5.32), ce qui est naturel puisque le modèle de structure temporelle du premier ordre décrivant le forçage climatique régional reste inchangé. Néanmoins, *a posteriori*, les probabilités (p_1, p_2, p_3) de chacun des trois régimes sont un peu modifiées, la possibilité qu'une année soit en régime bas étant en particulier doublée au détriment du régime moyen (Tableau 5.9). Si l'on affecte chaque année à son régime le plus probable, on classe à présent 7 années dans le régime bas, 39 années dans le régime moyen et 14 années dans le régime haut, principalement parce que par rapport au modèle précédent la moyenne du régime bas s'est rapprochée de zéro et que la variance du régime moyen a diminué (Tableau 5.9). Il semble donc que l'on distingue plus facilement 3 régimes avec le terme d'interaction du fait d'une meilleure séparation entre bruit local et forçage climatique global. Néanmoins, ces changements restent très peu significatifs. En particulier, les intervalles de crédibilité sur (p_1, p_2, p_3) ainsi que sur les moyennes et variances des trois régimes restent extrêmement larges, de sorte que les différences entre les estimateurs obtenus avec les deux modèles traduisent sans doute aussi la difficulté à identifier clairement trois régimes climatiques bien distincts sur la région étudiée. Notons donc qu'il est également possible d'utiliser un modèle avec le terme croisé h_{jt} mais un seul régime sans dégrader beaucoup son efficacité. Les résultats fournis par un tel modèle ne sont toutefois pas présentés afin d'alléger un peu le contenu du chapitre.

Figure 5.32 : Box plot du terme temporel avec le modèle final



années, 1946 à 2005

Modèle spatio-temporel à trois régimes et structure du second ordre appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

5.5. Bilan et application du modèle final

5.5.1. D'un modèle local rudimentaire à un modèle spatio temporel élaboré

A partir du modèle spatial initial dérivé de la littérature, le développement du modèle s'est fait par petits ajouts successifs : prétraitement des données selon une procédure automatisée, intégration de covariables topographiques et nivo-météorologiques et recalcul des nombres attendus, prise en compte de non stationnarités dans le temps, division du temps en trois régimes et prise en compte d'une structure spatiale du second ordre. Tout n'a cependant pas été « simple » et différentes structures de modèles ont été rejetées après avoir été testées. Le Tableau 5.10 récapitule l'évolution du critère DIC au cours du développement du modèle. Les améliorations les plus notables sont obtenues lors de l'introduction de la non stationnarité dans le temps et des effets spatio-temporels locaux, deux changements qui offrent au modèle davantage de souplesse pour ajuster la variabilité des données. A noter que pour obtenir le DIC associé au modèle spatial stationnaire, on l'a fait tourner sur les 60 années de données contrairement aux résultats précédemment présentés avec 12 périodes de 5 ans.

Tableau 5.10 : DIC pour les différents modèles spatio-temporels testés (Savoie-Haute-Savoie)

	DIC
Intégration de covariables dans le modèle	29063.4
Non-stationnarité dans le temps	25424.6
CAR propre	25426.5
Modèle avec dépendance spatiale d'ordre 2	25425.5
Modèle de krigeage	25428.2
Modèle avec autorégression	25421.7
Modèle à 3 régimes	25421.3
Modèle d'interaction + 3 régimes	23078.3

Le modèle final obtenu est capable de modéliser les occurrences avalanches comme un processus non homogène dans l'espace et le temps. La structure spatiale du premier ordre permet de visualiser les groupes de communes où il se produit anormalement beaucoup ou peu d'avalanches par rapport à une moyenne régionale pondérée par la climatologie et la topographie locale. Sur la région étudiée, on a établi qu'elle était significative mais que sa portée était faible et qu'une dépendance voisinage semblait meilleure qu'une dépendance basée uniquement sur la distance. La structure temporelle du premier ordre permet d'identifier le forçage climatique régional. Il a été difficile d'y identifier une structure car il semble y avoir des variations brutales d'une année sur l'autre tout au long de la période d'observation plutôt qu'une tendance ou des effets de mémoire nets. Enfin, les effets spatio-temporels locaux, bien que difficiles à interpréter, ont certainement un lien avec l'effort local d'observation. Pour clore ce chapitre, différents résultats obtenus non seulement en Savoie et Haute-Savoie mais également dans d'autres régions vont être à présent brièvement présentés. Ils permettent d'illustrer les utilisations possibles du modèle obtenu à la fois pour la prédétermination et la compréhension des variations dans le temps et l'espace de l'activité avalancheuse.

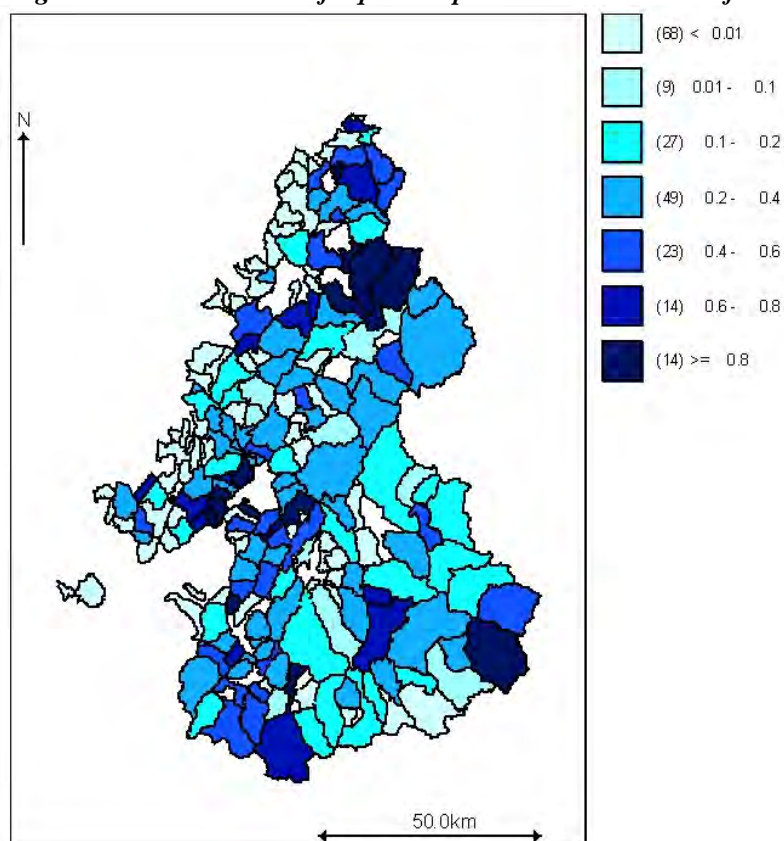
5.5.2. Prédiction avec le meilleur modèle

La prédiction avec le modèle final peut s'effectuer soit sachant un régime, soit au prorata des trois régimes selon leurs probabilités respectives. C'est cette deuxième option qui doit être privilégiée à des fins de prédétermination. Signalons que du fait des valeurs très élevées qui sont parfois générées en prédiction à cause du terme spatio-temporel local h_{jt} , il est plus prudent, si la loi prédictive dans la commune considérée est très dispersée, de choisir la médiane plutôt que la moyenne *a posteriori* comme estimateur ponctuel, cette dernière pouvant être considérablement surestimée à cause d'une valeur excessivement élevée « accidentelle ».

De façon un peu surprenante, les fréquences annuelles prédictives moyennes f_j^{pred} (Figure 5.33) et les intervalles de crédibilité à 95% correspondants ainsi obtenus sont proches de ceux obtenus avec le modèle spatial de base dès lors que les données sont prétraitées (Figure 5.16). En termes de prédiction, utiliser un modèle plus complexe n'apporte donc pas beaucoup plus d'informations. Néanmoins, pour

quelques communes, on a parfois des différences importantes entre les deux modèles, ce qui indique que l'on a quand même réussi à affiner un peu la prédiction en complexifiant beaucoup le modèle ! Enfin, afin d'illustrer l'utilisation possible de ce modèle pour la prédétermination sur sites peu ou pas documentés, signalons que la valeur prédite sur la commune de Bessans est 0.94 avalanches par couloir et par an avec un intervalle de crédibilité à 95% de [0.16, 2.2]. La prédiction communale moyenne correspond donc très bien dans ce cas à la valeur qui avait été inférée localement sur deux des couloirs de la commune au paragraphe 5.1. Cela indique que, sur ces sites, le calcul des aléas de référence avec une approche régionale pour la modélisation fréquentielle aurait été presque aussi précis qu'avec une approche purement locale.

Figure 5.33 : Médiane des fréquences prédites avec le modèle final



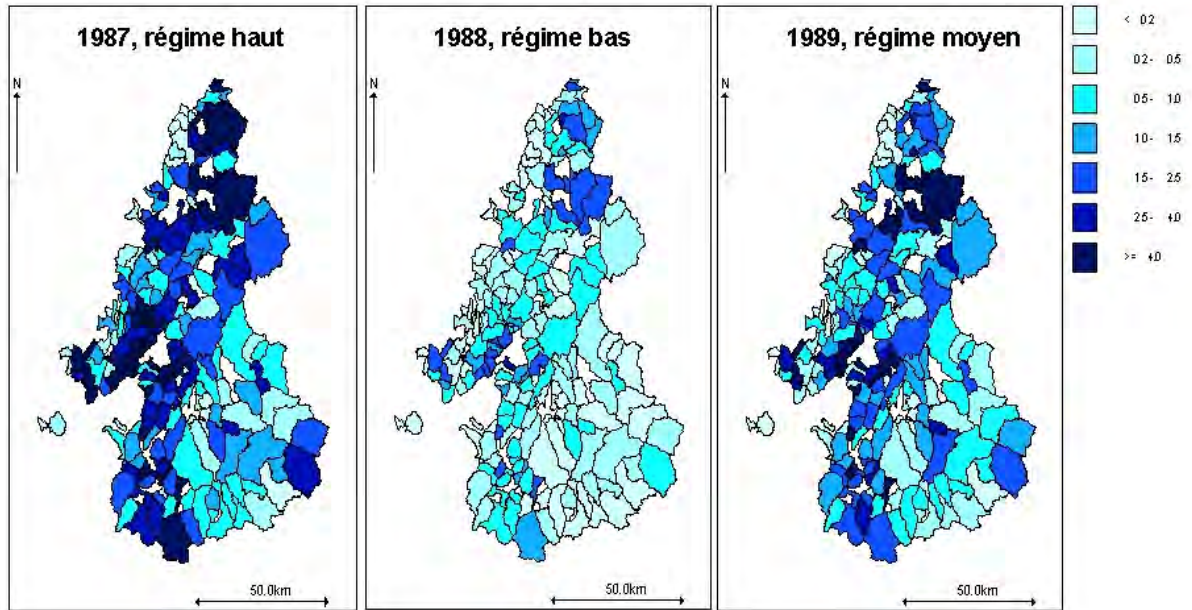
Modèle spatio-temporel à trois régimes et structure du second ordre appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

5.5.3. Visualisation des variations spatio-temporelles des occurrences avalancheuses

5.5.3.1. Variations temporelles des risques relatifs en Savoie et Haute-Savoie

Si du point de vue opérationnel de la prédétermination, la complexification du modèle peut sembler presque inutile, il n'en va pas de même d'un point de vue plus physique puisque les champs latents introduits permettent de visualiser en détail les variations spatio-temporelles des occurrences avalancheuses. Ainsi, la Figure 5.34 compare les cartes de risque relatif pour les années 1987, 1988 et 1989, respectivement classifiées dans les régimes haut, bas et moyen avec le modèle final. La structure spatiale reste toujours presque identique car c'est la carte sous jacente des u_j . Par contre, elle apparaît globalement plus ou moins contrastée en fonction de l'intensité du terme annuel g_t . De plus, des excès ou des déficits locaux dus au terme croisé h_{jt} sont observés pour quelques communes.

Figure 5.34 : Variation des risques relatifs en Savoie et Haute-Savoie

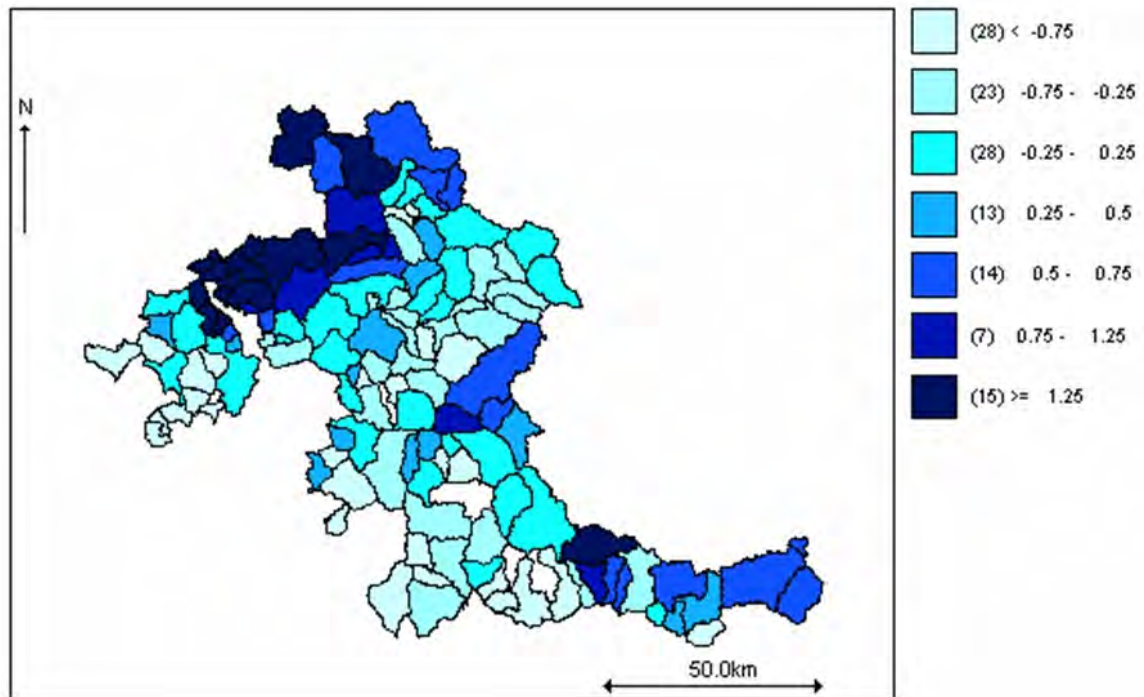


Modèle spatio-temporel à trois régimes et structure du second ordre appliqué à l'ensemble Savoie-Haute-Savoie après prétraitement des données et recalcul des nombres attendus.

5.5.3.2. Structure spatio-temporelle dans les Alpes du Sud

Les Alpes du Sud correspondent aux départements des Hautes-Alpes (05), des Alpes-Maritimes (06) et des Alpes de Haute Provence (04). Les données des Alpes-Maritimes et de Haute-Provence ne sont fiables qu'à partir de l'année 1960. Pour les communes de ces deux départements, beaucoup de données sont donc classifiées comme manquantes par la procédure de prétraitement avant 1960. Pour les Hautes-Alpes, ce problème se pose moins parce que, comme pour la Savoie et la Haute-Savoie, la qualité des données est satisfaisante dès la fin de la seconde guerre mondiale.

Figure 5.35 : Composante spatiale dans les Alpes du Sud

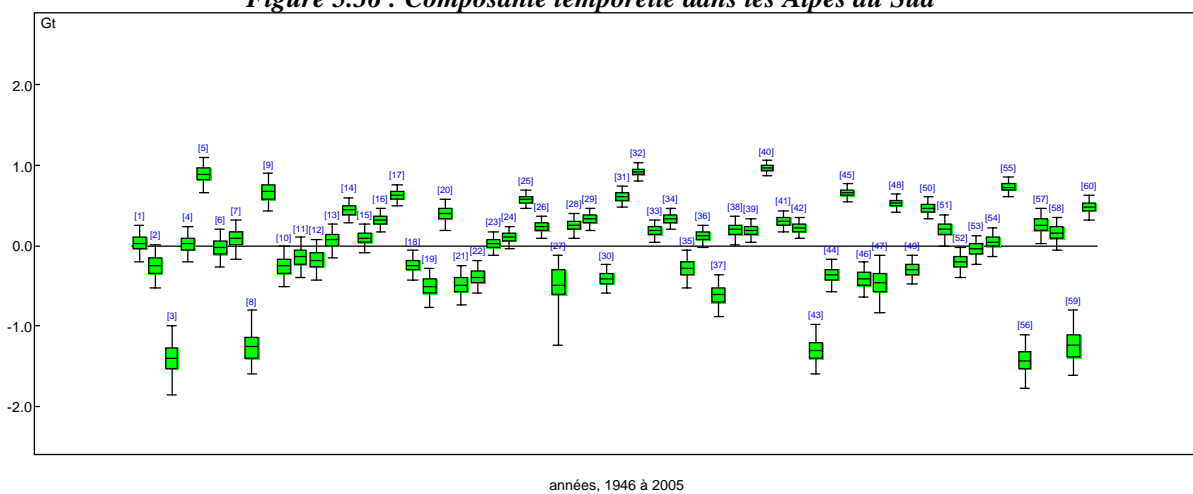


Modèle spatio-temporel à trois régimes et structure du second ordre appliqué aux Alpes du Sud après prétraitement des données et recalcul des nombres attendus.

Le modèle final a néanmoins été appliqué pour toutes les communes de ces trois départements sur la période 1946-2005 et il a donné de très bons résultats. La structure spatiale est ainsi fortement significative avec un rapport spatial de 0.71. Sur la carte des u_j (Figure 5.35), les massifs montagneux des Ecrins (au nord ouest) et du Mercantour (au sud est) ressortent clairement, de même qu'à un degré moindre le massif de la Haute Ubaye (au centre).

Quant aux variations de g_t pendant les 60 ans d'observation, on n'observe, comme pour la Savoie et la Haute-Savoie, ni tendance ni rupture manifeste mais des fluctuations assez brusques d'une année à l'autre. En outre, la classification en trois régimes donne des résultats plus satisfaisants que dans les Alpes du Nord avec des lois *a posteriori* des R_t privilégiant de façon plus marquée un régime qu'un autre et des intervalles de crédibilité sur les lois *a posteriori* des probabilités globales (p_1, p_2, p_3) un peu plus resserrés. Il semble donc qu'il y ait une variabilité interannuelle un peu plus forte dans les Alpes du Sud que dans les Alpes du Nord.

Figure 5.36 : Composante temporelle dans les Alpes du Sud



Modèle spatio-temporel à trois régimes et structure du second ordre appliqué aux Alpes du Sud après prétraitement des données et recalcul des nombres attendus.

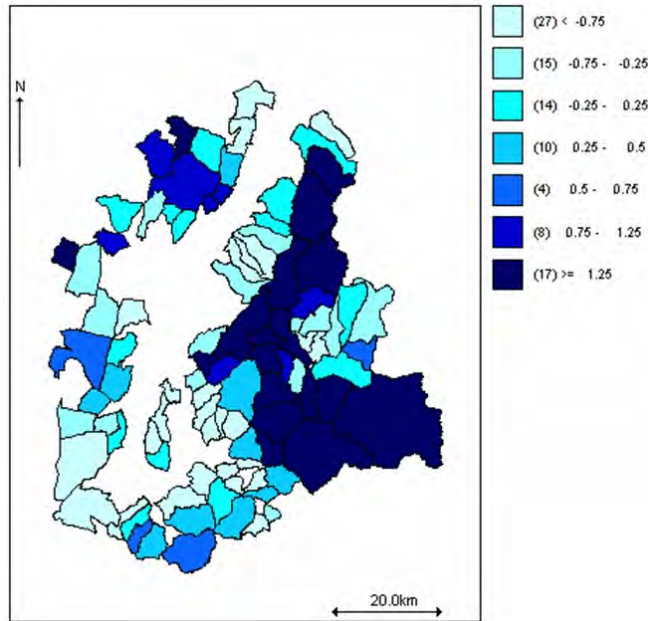
5.5.3.3. Structure spatiale en Isère

Pour le département de l'Isère, on ne dispose pas de la même qualité de données que pour les départements de la Savoie et de la Haute-Savoie. En effet, c'est seulement à partir de l'année 1960 que l'on est en possession de données relativement complètes et correctes. On a néanmoins appliqué à ces données le même prétraitement et le modèle spatio-temporel de base sur la période 1946-2003. Le modèle final n'a pas été utilisé car la quantité et la qualité des données ont été jugées insuffisantes pour distinguer 3 régimes climatiques et des structures spatio-temporelles du second ordre.

Sur la carte des u_j (Figure 5.37), on remarque que la structure spatiale est bien nette, ce qui, étant donné la qualité douteuse des données sur ce département, est assez remarquable. Ce sont les massifs montagneux élevés de Belledonne et de l'Oisans (est et est-sud est) qui ressortent le plus. A un degré moindre, le massif de Chartreuse apparaît également (nord ouest).

On a essayé de modéliser les données de l'Isère soit avec les données de la Savoie et de la Haute-Savoie, soit avec celles des trois départements des Alpes du Sud, mais les résultats obtenus n'ont pas été satisfaisants : l'ensemble du signal de l'Isère était alors classifié dans le bruit, empêchant l'identification d'une quelconque structure spatio-temporelle. L'Isère a donc un comportement naturel assez différent des deux autres régions.

Figure 5.37 : Composante spatiale en Isère



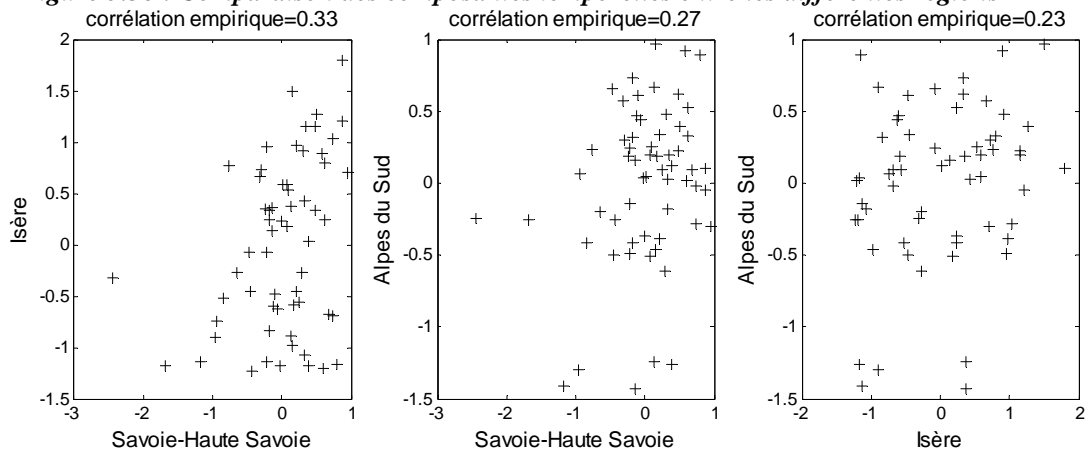
Modèle spatio-temporel de base appliqué aux Alpes du Sud après prétraitement des données et recalcul des nombres attendus.

5.5.3.4. Comparaison du fonctionnement des trois régions

Pour finir, les composantes temporelles g_t ont été comparées dans les trois régions afin de regarder si la composante climatique annuelle y était analogue. La Figure 5.38 montre ainsi une certaine corrélation entre les valeurs g_t obtenues dans les 3 régions au cours des 60 années d'étude, les trois coefficients de corrélation empiriques entre les estimateurs des g_t étant significativement non nuls au seuil des 95%. Néanmoins, ces derniers restent relativement faibles, ce qui indique que le fonctionnement des trois régions est malgré tout assez différent.

La corrélation est assez nettement plus forte entre l'Isère et l'ensemble Savoie-Haute-Savoie (coefficient de corrélation empirique de 0.33 entre les estimateurs ponctuels des g_t) qu'entre l'Isère et les Alpes du Sud (coefficient de corrélation empirique de 0.22 entre les estimateurs ponctuels des g_t), ce qui indique que l'Isère fonctionne davantage comme les autres départements des Alpes du Nord que comme ceux des Alpes du Sud. Plus curieusement étant donné que l'Isère est géographiquement située entre la Savoie et les Hautes-Alpes, la corrélation est plus forte entre les Alpes du Sud et l'ensemble Savoie-Haute-Savoie (coefficient de corrélation empirique de 0.27 entre les estimateurs ponctuels des g_t) qu'entre l'Isère et l'ensemble Savoie-Haute-Savoie. Ce résultat surprenant est peut-être lié à la qualité douteuse des données en Isère.

Figure 5.38 : Comparaison des composantes temporelles entre les différentes régions



5.6. Conclusion du chapitre 5

Dans ce chapitre, le calcul de la fréquence avalancheuse moyenne a été effectuée d'abord localement puis à l'échelle communale grâce à des modèles spatiaux et/ou spatio-temporels dérivés de l'épidémiologie statistique. Comme pour les modèles d'intensité, l'ensemble des difficultés techniques liées à l'inférence a pu être surmonté dans le cadre bayésien hiérarchique.

Les occurrences avalancheuses ont été modélisées localement grâce à un modèle d'observation poissonien très simple. L'adéquation obtenue entre données et modèle permet d'envisager le calcul des aléas de référence pour les cas d'étude avec confiance.

La construction d'un modèle spatio-temporel élaboré s'est faite ensuite à partir d'un modèle spatial dérivé de l'épidémiologie statistique qui a été adapté au problème traité puis progressivement complexifié par petites touches successives. L'hypothèse de modélisation principale a été que des communes topographiquement proches connaissent une activité avalancheuse similaire et peuvent donc partager leur information sur la base d'une matrice de distances. Ce postulat contredit quelque peu l'opinion des spécialistes des avalanches qui considèrent généralement le déclenchement comme la résultante de processus très locaux. Il a néanmoins été bien supporté par les données sur les régions étudiées avec une variabilité spatiale représentant par exemple 73% de la variabilité des risques relatifs en Savoie et Haute-Savoie une fois traitées les observations manquantes. Il a également permis d'effectuer une prédiction à l'échelle communale utile, par exemple, pour les sites pas ou peu documentés, et d'analyser la structure spatiale des occurrences avalancheuses sur plusieurs départements. De même, l'introduction d'un terme temporel a permis de mettre en évidence les fluctuations très fortes de l'activité avalancheuse d'une année sur l'autre.

Néanmoins l'hypothèse d'homogénéité communale reste critiquable et l'approche proposée pourra être utilement prolongée à l'échelle du couloir avec des modèles de krigeage bayésien du type de celui utilisé dans ce chapitre au niveau des centroïdes des communes une fois qu'un plus grand nombre de données aura été acquis sur des sites à la localisation sûre.

6. Calcul des aléas de référence

Préambule

Ce chapitre combine les modèles locaux d'intensité proposés au chapitre 4 avec la fréquence avalancheuse moyenne estimée au chapitre 5 afin d'obtenir les aléas de référence sur les cas d'étude choisis. Conformément au cadre défini au chapitre 3, l'incertitude liée à l'erreur d'estimation y est combinée avec la variabilité intrinsèque du phénomène pour obtenir la loi de probabilité prédictive de n'importe quelle période de retour. Les distributions des autres variables de sortie fournies par les différents modèles sont également étudiées et le calcul des pressions d'impact est mené. Une analyse de sensibilité approfondie est proposée afin de tenter d'évaluer quelles hypothèses de modélisation sont les plus lourdes de conséquence sur le zonage de l'aléa. Enfin, les différents modèles de prédétermination proposés sont comparés.

6.1. Moyenne *a posteriori* et période de retour

6.1.1. Avec le modèle de Coulomb

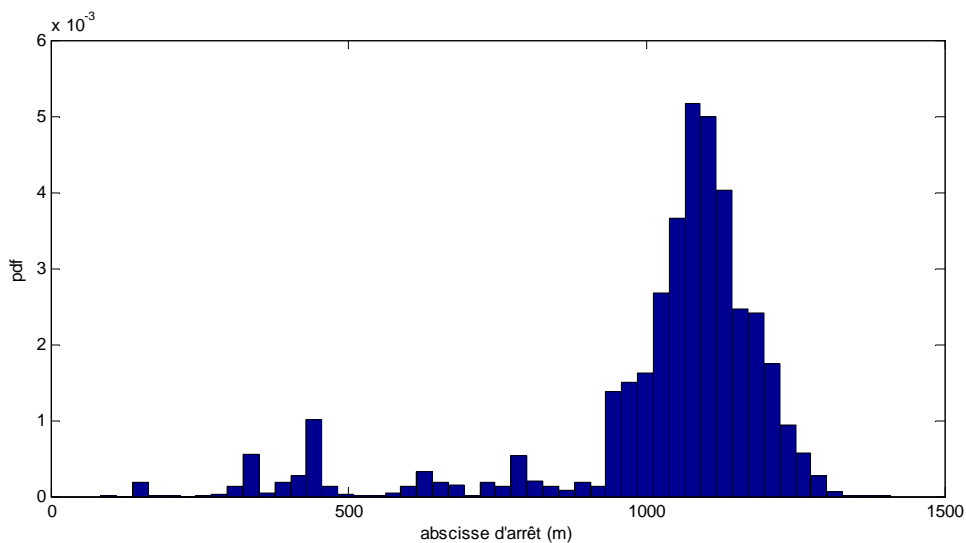
La moyenne *a posteriori* de la distribution des distances d'arrêt s'obtient en propageant celle du couple des variables d'entrée (x_{start}, μ_{coul}) selon l'approche de Monte Carlo détaillée au chapitre

3.4.1, de sorte que $p(x_{stop} | \hat{\theta}_M) = p\left(G(x_{start}, \mu_{coul} | \hat{\theta}_M)\right)$. La propagation numérique s'effectue sur

le profil topographique discrétisé et la distance d'arrêt associée à chaque doublet de valeurs d'entrée est calculée au pas de discrétisation près. Le temps de calcul total dépend de la longueur du couloir, du pas de discrétisation et du nombre de simulations effectuées.

Le cas d'étude est celui du chapitre 4.1.2, c'est à dire le couloir EPA 5 de Bessans. Le pas de discrétisation du profil topographique a été fixé à 5 m. Comme le couloir est assez court et que le modèle de Coulomb est rudimentaire, les calculs sont très rapides puisque 500 000 simulations n'ont nécessité qu'une heure de calcul sur un PC standard. Une telle taille d'échantillon est largement suffisante pour évaluer des quantiles de distance d'arrêt correspondant à des probabilités allant jusqu'à à 0.001 avec une excellente précision.

Figure 6.1 : Distribution des distances d'arrêt avec le modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

La Figure 6.1 montre que la distribution des distances d'arrêt obtenue est discrète. C'est en partie à cause de la discrétisation numérique du calcul de distance d'arrêt. Mais c'est également parce que

certaines portions du couloir caractérisées par des pentes locales trop fortes ne voient s’arrêter aucune avalanche, comme par exemple entre $x=520$ m et $x=580$ m. La Figure 6.1 montre également que la distribution des distances d’arrêt est multimodale, avec un mode principal très marqué autour de $x=1100$ m, et de nombreux petits modes secondaires entre $x=200$ m et $x=900$ m, ce qui est cohérent avec la topographie du site. En effet, le mode principal correspond à l’exutoire du couloir dans la vallée d’Avéroles où la pente devient très modérée, de sorte que la majeure partie des avalanches y est stoppée. De plus, les modes secondaires correspondent à de légers changements de la pente locale du couloir suffisants pour stopper quelques petites avalanches caractérisées par un coefficient de frottement élevé. La rivière, point le plus bas du profil ($x=1186.8$ m), correspond grossièrement au quantile 90 de la distribution. Cela indique qu’environ une avalanche sur 10 est capable d’aller sur la rive opposée tout d’abord plane puis présentant ensuite une pente négative.

Le calcul de la période de retour $T_{x_{stop}}$ s’effectue d’après l’équation 3.16. La probabilité de non

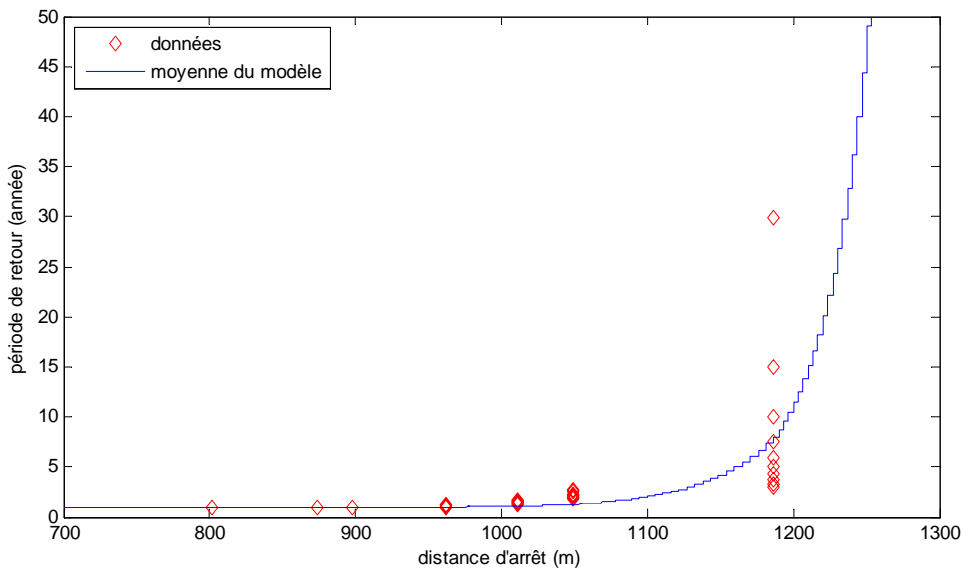
dépassement $\hat{F}(x_{stop}) = P\left(X_{stop} \leq x_{stop} \mid \hat{\theta}_M\right)$ est pondérée par la fréquence avalancheuse moyenne.

Avec le modèle poissonien local du chapitre 5.1, cette dernière est simplement le paramètre $\hat{\lambda}$, de sorte qu’une relation théoriquement univoque entre la distance d’arrêt à la période de retour est obtenue (eq. 6.1). Bien sûr, cette relation peut être inversée pour obtenir la distance d’arrêt x_{stop_T} correspondant à n’importe quelle période de retour T . Pour le cas d’étude, la relation n’est cependant pas tout à fait univoque du fait de la discrétisation du calcul de propagation.

$$T_{x_{stop}} = \frac{1}{\hat{\lambda} \times \left(1 - \hat{F}(x_{stop})\right)} \quad (\text{eq. 6.1})$$

La courbe obtenue permet de visualiser que la période de retour augmente très rapidement avec l’abscisse à partir du début de la zone d’arrêt. Cela indique que, même pour de très grandes périodes de retour, les avalanches ne vont pas très loin (Figure 6.2). Ainsi l’écart entre les abscisses décennale et centennale n’est que de 70 mètres (Tableau 6.1). L’explication est que le versant opposé de la vallée, situé au delà de la rivière, est suffisamment pentu pour empêcher les avalanches d’atteindre de très longues distances d’arrêt.

Figure 6.2 : Période de retour avec le modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

L’accord entre le modèle et les périodes de retour empiriques calculées avec les données est satisfaisant pour des distances d’arrêt faibles. Il est par contre difficile à mesurer pour des distances d’arrêt élevées car l’incertitude concernant les périodes de retour empiriques associées aux distances

d'arrêt observées les plus longues est très forte. En effet, dix données correspondent à la même distance d'arrêt c'est-à-dire à l'abscisse de la rivière ($x=1186.8$ m, abscisse correspondent à l'altitude du fond de vallée). Une telle distribution empirique n'est pas réaliste car la rivière n'est pas une barrière physique : parmi les événements enregistrés, certains ont en réalité dépassé la rivière, tandis que d'autres se sont arrêtées bien avant. Toutefois, dans l'EPA, ce sont les altitudes qui sont enregistrées et non les distances d'arrêt, ces dernières ayant été, comme dit au chapitre 3, recalculées à partir du profil topographique $z=f(x)$. Comme la topographie du fond de vallée est très plane, l'observateur a enregistré la même altitude pour toutes les avalanches l'ayant atteint, de sorte que l'on ne sait pas vraiment jusqu'où les différents événements sont allés. Le modèle, s'il doit certes être regardé avec prudence, reflète ainsi vraisemblablement mieux la réalité du terrain que les données. En tout cas, cet exemple illustre très bien la difficulté à travailler avec les données de l'EPA bruitées par une multitude de sources d'erreur, et plus encore la quasi impossibilité à valider des aléas de référence correspondant à des probabilités très faibles.

Tableau 6.1 : Distance d'arrêt et période de retour avec le modèle de Coulomb

Période de retour T (année)	Abscisse d'arrêt x_{stop_T} (m)
1	962.3
2	1099.7
5	1159.6
10	1196.8
30	1236.9
100	1267.9
300	1288.1

6.1.2. Avec le modèle de Coulomb bruité

Le cas d'étude est celui du chapitre 4.1.5, c'est-à-dire celui du couloir EPA 13 de Bessans. Rappelons que traiter ce couloir avec le modèle non bruité est impossible du fait des erreurs de propagation trop fortes qui sont générées tandis que traiter le couloir EPA 5 de Bessans avec le modèle bruité ne présente pas d'intérêt car sa topographie est si régulière que le terme d'erreur de propagation y est toujours quasi nul.

Le calcul de la distribution des distances d'arrêt s'effectue pour le modèle de Coulomb bruité exactement de la même manière que pour le modèle de Coulomb non bruité. Les seules différences sont l'utilisation d'une vitesse initiale non nulle $v_{zero} = 25$ m/s et l'addition d'une erreur de propagation aléatoire au résultat du calcul déterministe, de sorte que

$$p\left(x_{stop} \mid \hat{\theta}_M, v_{zero}\right) = p\left(G\left(x_{start}, \mu_{coul} \mid \hat{\theta}_M, v_{zero}\right) + \varepsilon \mid \hat{\theta}_M\right).$$

Grâce au terme d'erreur aléatoire, les abscisses inatteignables avec le modèle non bruité du fait du modèle de propagation et de la discrétisation numérique deviennent accessibles. En outre, grâce à la vitesse initiale non nulle, des géométries de couloir relativement complexes peuvent être traitées de manière relativement satisfaisante. Les conséquences immédiates en sont, par rapport au modèle non bruité, le lissage de la distribution de probabilité des distances d'arrêt et la possibilité de simuler des distributions de distance d'arrêt fortement multimodales.

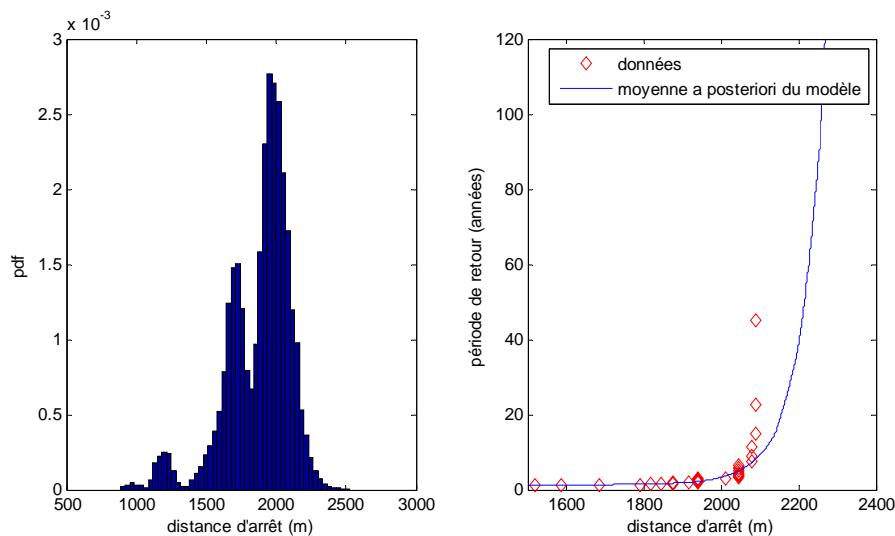
Tableau 6.2 : Comparaison entre données et moyenne a posteriori du modèle de Coulomb bruité

	Données	Distribution simulée
Moy. (m)	1909.3	1875.9
SD (m)	159	232.6
Skew.	-0.93	-1.2
q₁ (m)	1518.9	1131.8
q₁₀ (m)	1636.8	1609
q₂₅ (m)	1845.1	1732.4
q₅₀ (m)	1941	1936
q₇₅ (m)	2046.1	2030.7
q₉₀ (m)	2078.8	2113
q₉₉ (m)	2090	2261

Sur le cas d'étude, la distribution des distances d'arrêt présente un mode principal autour de $x=2000$ m, un mode secondaire autour de $x=1700$ m et un mode tertiaire autour de $x=1200$ m, ce qui traduit la topographie déjà raisonnablement complexe du couloir (Figure 6.3, gauche). Le mode principal correspond aux pentes faibles situées à l'exutoire du couloir dans la vallée et stoppant la majorité des avalanches. Le mode secondaire correspond à un segment du couloir où la pente locale décroît jusqu'à 23 degrés et stoppe ainsi un nombre significatif d'avalanches avant de croître à nouveau jusqu'à près de 40 degrés. Le petit mode tertiaire au milieu du couloir correspond à une légère diminution de la pente qui stoppe les avalanches caractérisées par les coefficients de frottement les plus élevés. L'incrément de distance d'arrêt avec le quantile, par exemple près de 150 mètres entre les quantiles décennal et centennal, est bien plus rapide que pour le cas d'étude précédent car le couloir se termine par des pentes positives faibles assez longues suivies d'une zone plane sans versant opposé permettant aux avalanches exceptionnelles de par leur intensité d'aller très loin (Tableau 6.2). Cette vision de l'activité avalancheuse correspond relativement bien aux données. En effet, un test classique de type Kolmogorov-Smirnov ne rejette pas l'hypothèse que les distributions empiriques et simulées sont issues d'une même population. Bien que la p-valeur du test soit faible, ce résultat est encourageant concernant la capacité du modèle à reproduire la variabilité des données sachant que l'estimation ne s'est pas faite directement sur les distances d'arrêt.

Si l'on compare plus précisément les distributions des distances d'arrêt observées et simulées, on constate que, dans l'échantillon utilisé pour l'inférence, figurent effectivement quelques avalanches s'arrêtant entre $x=1650$ m et $x=1750$ m, c'est à dire au niveau du mode secondaire de la distribution simulée, tandis que la plupart des autres se sont arrêtées entre $x=1900$ m et $x=2100$ m c'est à dire au niveau du mode principal de la distribution simulée. Le Tableau 6.2 montre ainsi qu'entre la médiane empirique et la médiane simulée, la différence n'est que de 10 mètres, tandis qu'elle est d'environ 30 mètres pour les quantiles d'ordre 10 et 90. Pour des quantiles plus rares, l'agrément est nettement moins bon. Une raison évidente est la faible taille de l'échantillon d'apprentissage. Néanmoins, la Figure 6.3 (droite) semble également suggérer un comportement pessimiste du modèle, c'est-à-dire une surestimation systématique des distances d'arrêt correspondant à des périodes de retour supérieures à 10 ans.

Figure 6.3 : Distances d'arrêt et période de retour avec le modèle de Coulomb bruité



Couloir EPA 13 de la commune de Bessans, Savoie, France.

6.1.3. Avec le modèle de Voellmy bloc glissant

Le calcul de la distribution moyenne des distances d'arrêt s'effectue comme pour le modèle de Coulomb, mais avec une variable d'entrée supplémentaire, la hauteur de bloc h . Soit

$$p\left(x_{stop}, v_x \mid \hat{\theta}_M\right) = p\left(G\left(x_{start}, h, \mu_{voel} \mid \hat{\theta}_M\right)\right).$$

Cette notation rappelle qu'à la différence des deux modèles de Coulomb, on considère le modèle de Voellmy bloc glissant suffisamment réaliste en

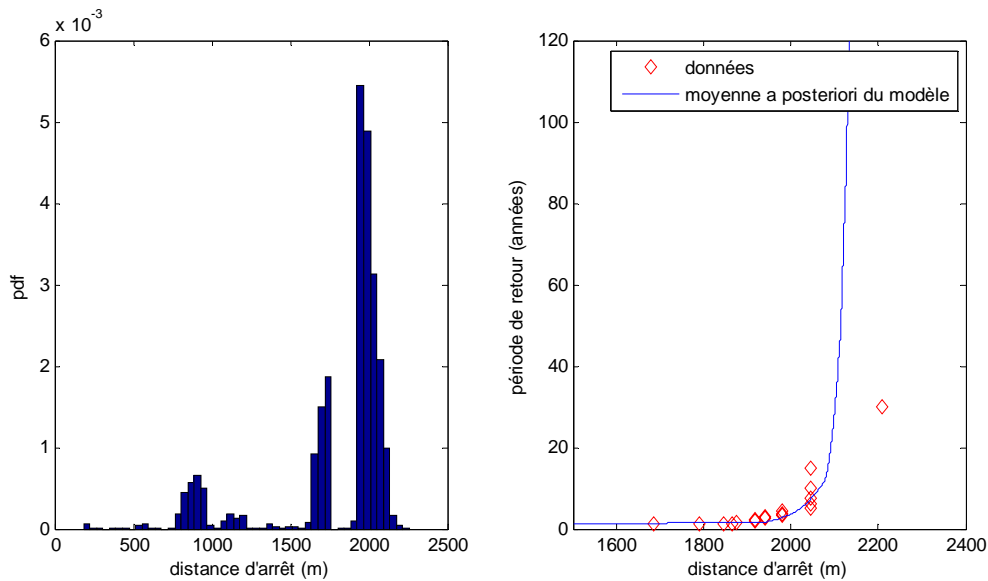
termes de calcul de vitesse pour que les distributions des vitesses en chaque point du couloir puissent être utilisées avec profit. Notons également que $\hat{\theta}_M$ intervient ici non seulement au travers de la distribution des variables d'entrée mais aussi sous la forme du paramètre de propagation $\hat{\theta}_G = \hat{\xi}$.

A la différence du modèle de Coulomb bruité, on considère le modèle de propagation suffisamment performant pour atteindre toutes les distances d'arrêt réalistes, de sorte que l'on n'ajoute pas le bruit aléatoire σ_{num} à la prédiction, ce dernier ayant été uniquement introduit pour permettre la mise en œuvre de l'inférence. Afin de ne pas obtenir une distribution des distances d'arrêt faussement discrète, on effectue comme pour l'inférence un calcul exact de la distance d'arrêt associé à chaque triplet de variable d'entrée grâce à l'équation 3.32. Comme pour les modèles de Coulomb, le temps de calcul total dépend du nombre de simulations, de la longueur du couloir et du pas de discrétisation. Il reste toutefois suffisamment supportable pour qu'un nombre important de simulations puisse être effectué sans peine de façon à pouvoir négliger l'erreur de Monte Carlo.

6.1.3.1. Distances d'arrêt

Le cas d'étude retenu est celui du chapitre 4.2, soit le même couloir que pour le modèle de Coulomb bruité, mais avec seulement 26 données du fait de l'absence de relevés des volumes de dépôt pour un nombre important d'événements. Malgré cela, une distribution des distances d'arrêt trimodale assez analogue à celle obtenue avec le modèle de Coulomb bruité est obtenue (Figure 6.4, gauche). Les modes sont cependant plus marqués qu'avec le modèle de Coulomb bruité et certaines parties du couloir, notamment la zone de pente forte précédant l'arrivée en fond de vallée, ne sont atteintes par aucune avalanche à cause de la non prise en compte d'une erreur aléatoire. Du fait de sa loi de frottement plus réaliste, on peut penser que ce modèle offre ainsi une vision plus exacte de la physique du phénomène, notamment en ne permettant pas l'arrêt sur des pentes trop fortes.

Figure 6.4 : Distances d'arrêt et période de retour avec le modèle de Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

En termes d'adéquation entre données et modèles, le résultat semble de prime abord moins satisfaisant qu'avec le modèle de Coulomb bruité, avec notamment une variance beaucoup plus forte de la distribution simulée et un écart en moyenne de plus de 100 m (Tableau 6.3). Cette faible adéquation est à relativiser du fait de la quantité réduite de données utilisée pour l'inférence, d'autant qu'elle est surtout due à l'absence dans l'échantillon d'apprentissage de distances d'arrêt courtes.

Si l'on s'intéresse plus spécifiquement aux distances d'arrêt présentant un intérêt pour le zonage (Figure 6.4, droite), on est au contraire amené à constater que le modèle s'ajuste bien sur les données dont les périodes de retour empiriques s'échelonnent de 5 à 20 ans. A la différence du modèle de Coulomb bruité, il ne semble donc pas y avoir de surestimation systématique par le modèle. Le modèle

sous estime même peut-être les distances d'arrêt correspondant aux périodes de retour les plus élevées. En effet, il propose, pour une période de retour de 35 ans, une distance d'arrêt à peine supérieure à 2100 m alors que la distance d'arrêt la plus élevée du jeu de données qui correspond à une période de retour empirique de 35 ans est supérieure à 2200 m.

Ce constat doit cependant être nuancé par la difficulté que pose la caractérisation fréquentielle de la distance d'arrêt enregistrée la plus longue. Tout d'abord, la période de retour empirique associée n'a pas grand sens étant donné qu'elle correspond nécessairement à la durée d'observation pondérée par la fréquence avalancheuse moyenne. Ensuite, cette distance d'arrêt exceptionnelle correspond à une zone plane, ce qui renvoie au problème déjà évoqué de correspondance abscisse-altitude, et ce d'autant plus que l'événement est ancien et que la donnée est donc vraisemblablement peu précise et non vérifiable... Enfin, du fait de son caractère exceptionnel, cette donnée influence fortement les valeurs élevées du modèle lorsqu'elle est utilisée pour l'inférence. Elle ne figurait pas dans l'échantillon utilisé avec le modèle de Coulomb bruité car c'est un événement très ancien et que pour ce modèle on avait limité la période d'étude à 1960-2003 afin de privilégier les données les plus fiables. La comparaison des résultats obtenus avec les deux modèles est donc difficile. On y reviendra au chapitre 6.3 sous la forme d'une analyse à quantité d'information équivalente pour l'inférence.

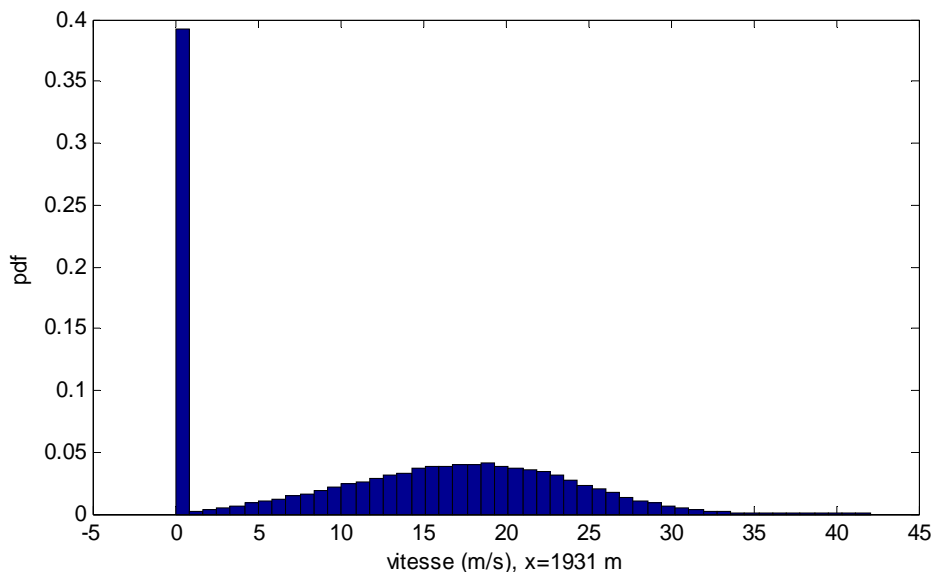
Tableau 6.3 : Comparaison entre données et moyenne du modèle de Voellmy bloc glissant

	Données (m)	Distribution simulée (m)
Moy. (m)	1933.1	1806.2
SD (m)	104.8	375.5
Skew.	0.19	-1.94
q₁ (m)	1685.5	770.4
q₁₀ (m)	1796.1	953.6
q₂₅ (m)	1876.1	1726
q₅₀ (m)	1915.6	1962.8
q₇₅ (m)	1979.8	2013.8
q₉₀ (m)	2046.1	2071.9
q₉₉ (m)	2207.8	2132.4

6.1.3.2. Vitesse et coefficient de frottement

La Figure 6.5 illustre la distribution des vitesses obtenue au point de la zone d'arrêt dont l'abscisse correspond à une période de retour quasi bisannuelle c'est-à-dire qu'une avalanche y passe en moyenne une année sur deux. Le pic observé en zéro en correspond aux avalanches n'atteignent pas cette abscisse. Conditionnellement à son atteinte, ce point enregistre une vitesse moyenne de 17.4 m/s et des vitesses maximales de l'ordre de 40 m/s.

Figure 6.5 : Distribution des vitesses avec le modèle de Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

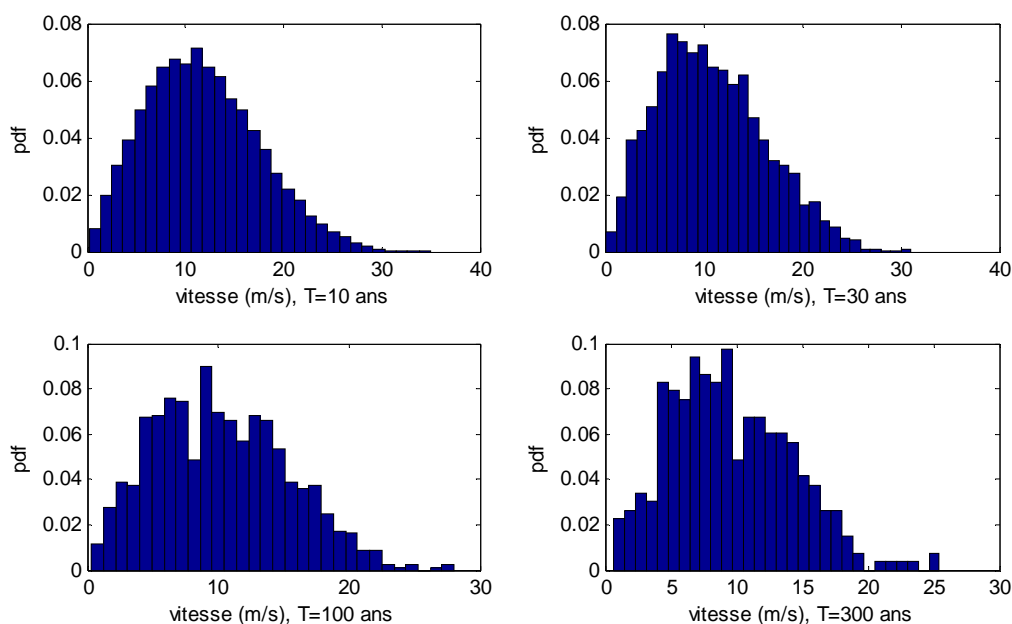
Les distances d'arrêt correspondant à des périodes de retour plus élevées sont évidemment atteintes par moins d'avalanches. Afin de ne pas « écraser » les valeurs non nulles, les distributions des vitesses aux abscisses correspondant à des périodes de retour de 10 à 300 ans sont donc proposées conditionnellement à leur atteinte (Figure 6.6). La probabilité d'atteinte qui correspond simplement à la fraction de valeurs non nulles diminue naturellement avec la période de retour (Tableau 6.4). Notons également que la régularité des distributions diminue avec la période de retour. Cela est dû à une diminution du nombre d'événements considérés qui entraîne une augmentation de l'erreur de Monte Carlo. Obtenir des distributions plus lisses pour des abscisses de période de retour très élevées nécessite un effort de simulation plus important sans modifier notablement l'interprétation des résultats.

Tableau 6.4 : Vitesses des événements de référence avec le modèle de Voellmy bloc glissant

	$P(x_{stop} \geq x_{stopT})$	Moy.	SD	Skew	q _{2.5}	q _{97.5}
$v_{T_{10}}$ (m/s)	0.121	11.8	5.7	0.43	2.2	24.2
$v_{T_{30}}$ (m/s)	0.028	10.8	5.4	0.46	2.2	22.3
$v_{T_{100}}$ (m/s)	0.009	10.2	5.1	0.38	1.9	20.5
$v_{T_{300}}$ (m/s)	0.003	9.5	4.7	0.5	1.6	18.6

A n'importe quelle abscisse, la distribution des vitesses est unimodale. Elle est également asymétrique, ce qui indique une décroissance lente de la probabilité de dépassement avec la vitesse (Tableau 6.4). Ainsi pour une avalanche atteignant une abscisse décennale a-t-on une vitesse moyenne de 11.8 m/s mais encore un pourcentage non négligeable d'avalanches présentant une vitesse supérieure à 20 m/s.

Figure 6.6 : Vitesses et période de retour avec le modèle de Voellmy bloc glissant

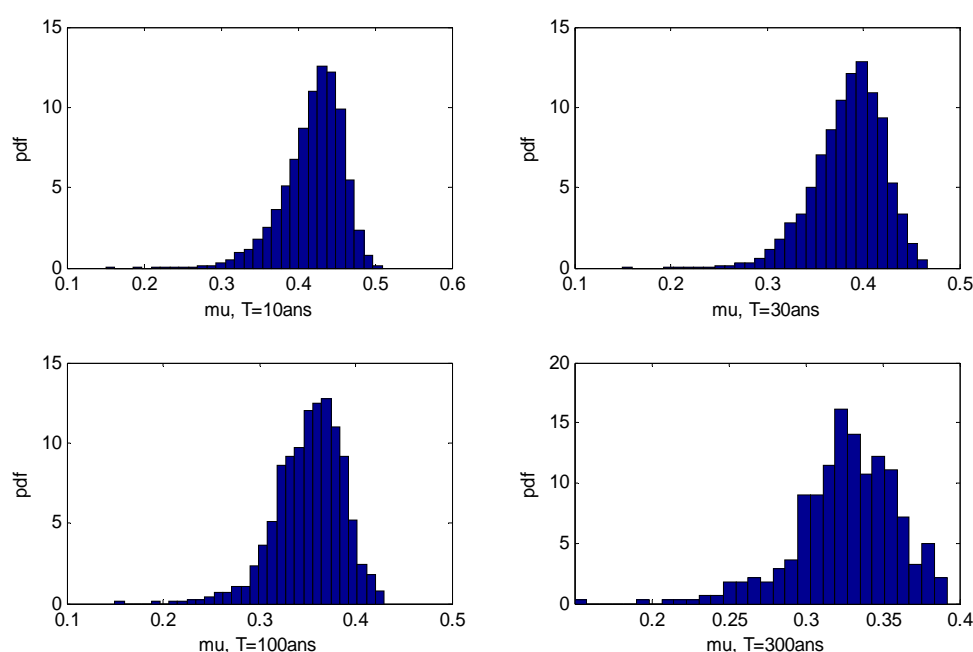


Couloir EPA 13 de la commune de Bessans, Savoie, France. Les distributions sont conditionnelles à l'atteinte de l'abscisse considérée. Ainsi pour une période de retour de 10 ans, les avalanches considérées ont une probabilité annuelle d'atteinte de 0.1.

Si l'on compare les distributions obtenues aux différentes abscisses, on constate une diminution naturelle des vitesses moyennes et maximales avec la période de retour. Néanmoins, elle n'est pas très rapide puisque la vitesse moyenne des avalanches qui atteignent une abscisse tricentennale est à peine inférieure à celles des avalanches qui atteignent une abscisse centennale. Quant aux vitesses maximales prédites, elles sont de l'ordre de 30 m/s à une abscisse décennale et encore de l'ordre de 20 à 25 m/s à une abscisse tricentennale. Ces décroissances assez lentes sont à relier à la topographie de

la zone d'arrêt et à la loi de frottement utilisée. En effet, la zone d'arrêt est, sur ce couloir, caractérisée par une pente faible mais non nulle. Celle-ci est suffisante pour ne pas freiner les avalanches caractérisées par les valeurs de coefficient de frottement μ_{voel} les plus faibles. Pour ces avalanches, $\mu_{voel} < \tan(\phi_d)$, de sorte que la vitesse limite prédite par le modèle de Voellmy est atteinte et qu'il n'y a pas de ralentissement. Au fur et à mesure que l'abscisse augmente, il se produit donc un « filtrage » des avalanches caractérisées par les coefficients de frottement les plus forts, tandis que les autres continuent à avancer sans freiner jusqu'à proximité immédiate de leur point d'arrêt. Cela est confirmé si l'on s'intéresse aux distributions des coefficients de frottement μ_{voel} aux mêmes abscisses. Celles-ci sont en moyenne plus faibles et beaucoup plus asymétriques que la distribution moyenne *a posteriori* (Figure 4. 29), ce qui indique que presque toutes les avalanches avec de petites valeurs de μ_{voel} arrivent aux abscisses considérées, tandis que presque aucune avalanche avec de fortes valeurs de μ_{voel} n'y arrive (Figure 6.7).

Figure 6.7 : Frottement latent et période de retour avec le modèle de Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les distributions sont conditionnelles à l'atteinte de l'abscisse considérée.

En outre, on a bien une diminution du μ_{voel} moyen avec la période de retour, celui-ci passant de 0.42 à 0.326 lorsque la période de retour passe de 10 à 300 ans (Tableau 6.5). Les valeurs observées les plus faibles restent, mêmes pour les périodes de retour les plus élevées, supérieures à la valeur de 0.16 généralement retenue par la littérature pour les avalanches les plus extrêmes. L'écart type des distributions aux différentes abscisses reste par contre remarquablement constante et égal à environ la moitié de celui de la distribution moyenne *a posteriori*, ce dernier valant 0.079.

Tableau 6.5 : Frottement μ_{voel} des événements de référence, modèle Voellmy bloc glissant

	Moy.	SD	Skew	q _{2.5}	q _{97.5}
$\mu_{voel_{10}}$	0.420	0.038	-0.937	0.329	0.478
$\mu_{voel_{30}}$	0.385	0.036	-0.782	0.305	0.444
$\mu_{voel_{100}}$	0.352	0.034	-0.837	0.272	0.410
$\mu_{voel_{300}}$	0.326	0.034	-1.031	0.247	0.382

6.1.4. Avec Saint-Venant :

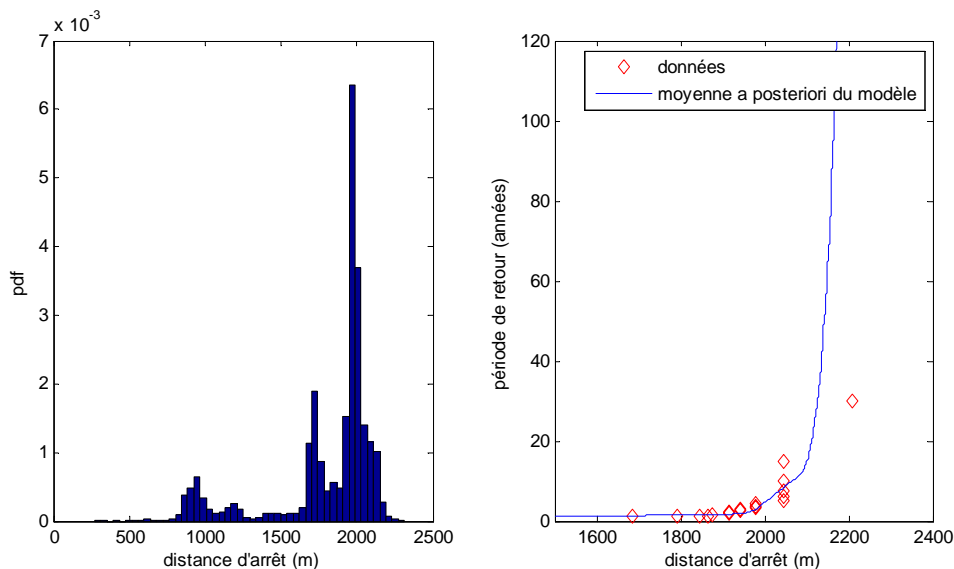
Pour le modèle de Saint-Venant, la distribution moyenne des sorties sachant celle des entrées est simulée selon $p\left(x_{stop}, v_{xt}, h_{xt} \mid \hat{\theta}_M\right) = p\left(G\left(x_{start}, h_{plaque}, L_{plaque}, \mu_{voel} \mid \hat{\theta}_M\right)\right)$. La longueur de départ L_{plaque} est obtenue de façon déterministe selon l'eq. 4.18 en transformant à chaque simulation la hauteur de départ h_{plaque} . Plutôt que les champs de hauteur et de vitesse, on retient uniquement les valeurs maximales à chaque abscisse. A noter également que, du fait de la condition d'arrêt définie en termes de débit seuil utilisée, les distances d'arrêt sont à nouveau calculées au pas de 5 m. Contrairement au modèle de Coulomb bruité où l'erreur de propagation avait lissé le résultat et au modèle de bloc glissant pour lequel un calcul exact de l'abscisse où la vitesse du bloc devient nulle avait été effectué, on retrouve donc une distribution des distances d'arrêt « pseudo discrète ». Enfin, signalons que, tout comme pour l'inférence, on est confronté à des problèmes de temps calcul, de sorte que seulement 20 000 simulations ont été effectuées, soit un temps de calcul total d'environ 5 jours sur un PC récent.

6.1.4.1. Distances d'arrêt

Le cas d'étude reste toujours le couloir 13 de Bessans avec 26 données. La distribution des distances d'arrêt obtenue est effectivement discrète (Figure 6.8, gauche). Mais elle est par ailleurs plus lisse qu'avec le modèle de bloc glissant puisqu'on n'observe plus de sections du profil ou aucune avalanche n'est capable de s'arrêter. Cela traduit la plus grande aptitude du modèle à reproduire une grande variété de distances d'arrêt. Celle-ci, déjà évoquée au chapitre 4.3, est due à la capacité du modèle à adapter les caractéristiques de l'écoulement à la topographie, en particulier au niveau des zones de ressaut.

D'un point de vue plus quantitatif, la distribution des distances d'arrêt obtenue avec le modèle de Saint-Venant est toutefois assez analogue à celle obtenue avec le modèle de bloc glissant, ce qui n'est pas choquant puisqu'une information quasi similaire a été utilisée pour l'inférence des deux modèles. En termes d'adéquation aux données, le comportement du modèle est donc assez similaire à celui du bloc glissant (Tableau 6.6). On constate néanmoins que le modèle de Saint-Venant est un peu plus pessimiste pour des périodes de retour élevées que le modèle de bloc glissant, c'est-à-dire que pour une période de retour donnée, l'abscisse obtenue avec le modèle de Saint-Venant est un peu supérieure à celle obtenue avec le modèle de Voellmy bloc glissant. Cela conduit le modèle de Saint-Venant à être un peu plus proche de la donnée correspondant à la période de retour empirique la plus élevée que le modèle de Voellmy (Figure 6.8, droite), ce qui, étant donné les réserves exprimées précédemment à propos de cette avalanche, n'est toutefois pas très significatif.

Figure 6.8 : Distance d'arrêt et période de retour avec le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 6.6 : Comparaison entre données et moyenne du modèle de Saint-Venant

	Données (m)	Distribution simulée (m)
Moy. (m)	1933.1	1805.8
SD (m)	104.8	343.3
Skew.	0.19	-1.69
q₁ (m)	1685.5	854
q₁₀ (m)	1796.1	1140.9
q₂₅ (m)	1876.1	1724
q₅₀ (m)	1915.6	1960.6
q₇₅ (m)	1979.8	1995.7
q₉₀ (m)	2046.1	2074
q₉₉ (m)	2207.8	2167.4

Comme un nombre limité de simulations a pu être effectué, des intervalles de confiance ont été construits sur les distances d'arrêt associées à différentes périodes de retour. La largeur de l'intervalle de confiance de Monte Carlo associé à chaque probabilité annuelle de non dépassement a ainsi été calculée conformément à l'éq. 3.18. Ensuite, les distances d'arrêt correspondant aux bornes de l'intervalle ont été « récupérées » à partir de l'échantillon ordonné. Notons que cette méthode néglige l'erreur de Monte Carlo sur l'estimation de la fréquence avalancheuse moyenne, celle-ci étant supposée connue sans erreur grâce au modèle conjugué utilisé.

Tableau 6.7 : Erreur de Monte Carlo avec le modèle de Saint-Venant

Période de retour T (année)	Abscisse d'arrêt x_{stop_T} (m)	q _{2.5}	q _{97.5}
2	1953.7	1951.1	1954.3
5	2004	2003.7	2006.5
10	2064	2059	2069
30	2125.2	2123.6	2129
100	2164	2160	2169
300	2203.9	2198	2213.1

L'analyse du Tableau 6.7 montre que les intervalles de confiance obtenus restent largement supportables puisque la distance d'arrêt centennale est donnée grossièrement à plus ou moins 5 m. 20 000 itérations semblent donc suffisantes pour déterminer les valeurs de projet recherchées avec une précision suffisante. De plus, on observe globalement bien un accroissement de la largeur de l'intervalle de confiance à 95% avec la période de retour, celui-ci passant de 3 m pour une période de retour bisannuelle à 15 m pour une période de retour tricentennale. Cela traduit simplement la difficulté croissante à estimer précisément par simulation des probabilités de dépassement de plus en plus petites.

Toutefois, la largeur de l'intervalle n'est pas strictement croissante avec la période de retour puisqu'elle est de 10 m pour l'abscisse décennale tandis qu'elle n'est que de 6 m pour l'abscisse centennale. Cela indique un effet de la propagation numérique et de la topographie. En effet, lorsque la pente locale est « favorable » à l'arrêt du fait d'une diminution nette de la pente, la probabilité de non dépassement est estimée précisément par simulation, de sorte que l'intervalle de confiance est réduit. Quand, au contraire, la pente locale favorise un étalement des distances d'arrêt, la probabilité de non dépassement est plus difficile à estimer par simulation et l'intervalle de confiance est plus large.

6.1.4.2. Hauteurs et vitesses maximales

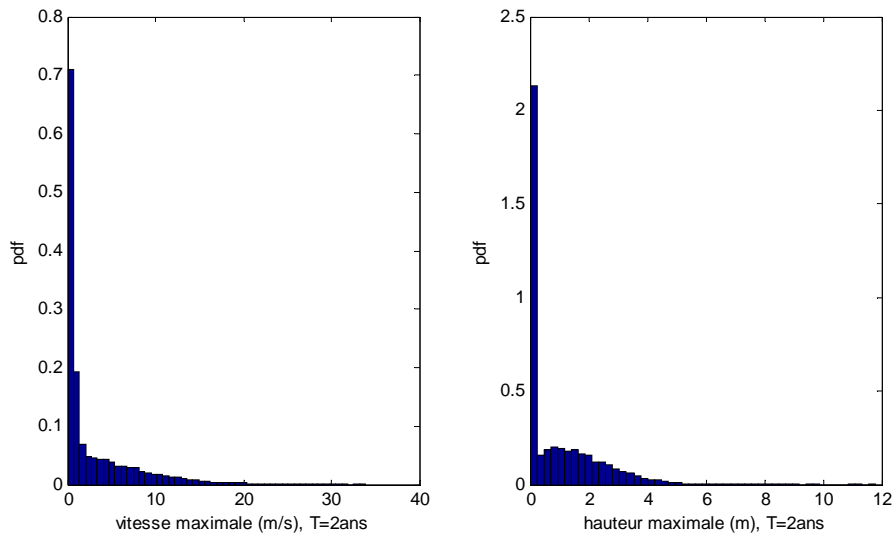
La Figure 6.9 présente les distributions des vitesses et des hauteurs maximales pour l'abscisse correspondant à une période de retour de 2 ans. Sur les deux distributions, le pic en zéro correspond aux avalanches n'atteignant pas cette abscisse, soit un peu moins d'une avalanche sur deux.

Les vitesses maximales des avalanches atteignant cette abscisse sont majoritairement inférieures à 15 m/s. Néanmoins, la distribution étant très asymétrique, des valeurs allant jusqu'à plus de 30 m/s sont enregistrées. De façon un peu surprenante, le mode de la distribution est situé très proche de zéro et la densité décroît avec la vitesse. C'est une différence marquante avec le modèle de bloc glissant pour

lequel un mode bien marqué autour de 20 m/s avait été observé pour les avalanches atteignant une abscisse bisannuelle (Figure 6.5).

La distribution des hauteurs (Figure 6.9, droite) est moins surprenante avec pour les avalanches atteignant l'abscisse bisannuelle des hauteurs maximales s'échelonnant très majoritairement de 0 à 5 m. Le mode est distinct de zéro mais assez faible et les valeurs élevées sont très dispersées puisque des hauteurs allant jusqu'à 12 mètres sont simulées. De telles hauteurs n'ont toutefois rien d'irréaliste puisque des hauteurs plus importantes sont parfois observées sur le terrain pour les événements les plus extrêmes.

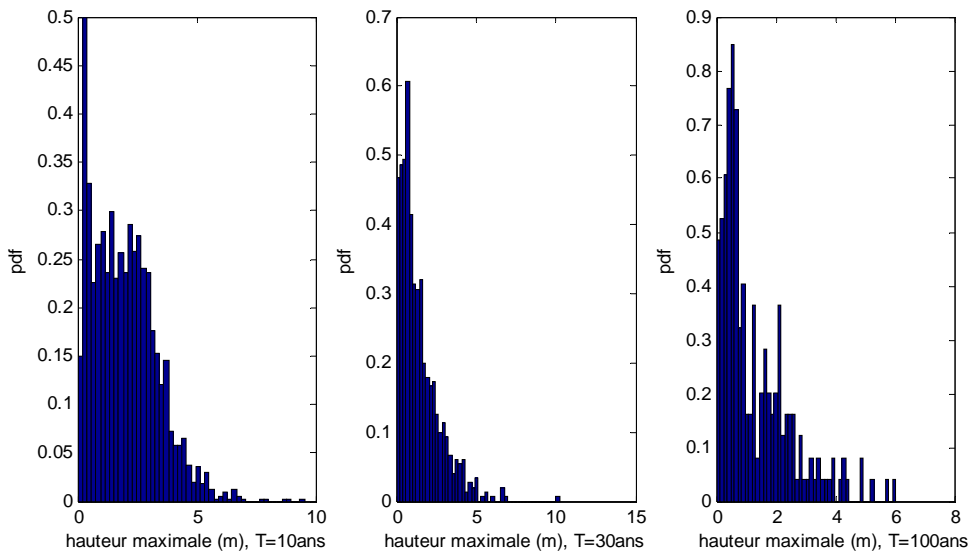
Figure 6.9 : Distribution des vitesses et hauteurs maximales avec le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Pour des périodes de retour plus élevées, seules les vitesses et hauteurs maximales des avalanches atteignant l'abscisse considérée sont représentées. Evidemment, comme pour le modèle de Voellmy, plus la période de retour augmente et moins on a d'événements, de sorte que les distributions obtenues sont de plus en plus irrégulières. C'est d'ailleurs d'autant plus visible que seules 20 000 avalanches au total ont été simulées, soit un peu plus de 200 avalanches centennales puisque la fréquence avalancheuse moyenne est légèrement inférieure à 1.

Figure 6.10 : Hauteurs maximales et période de retour avec le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les distributions sont conditionnelles à l'atteinte de l'abscisse considérée.

Pour les avalanches décennales, les hauteurs prédites s'échelonnent de façon assez équiprobables de 0 à 5 m (Figure 6.10, gauche), et les vitesses correspondantes de 0 à 20 m/s, avec un mode bien marqué autour de 5 m/s (Figure 6.11, gauche). Globalement, vitesses et hauteurs maximales décroissent ensuite avec la période de retour du fait de la propagation dans une zone d'arrêt de plus en plus plane conduisant les avalanches à ralentir puis s'arrêter les unes après les autres.

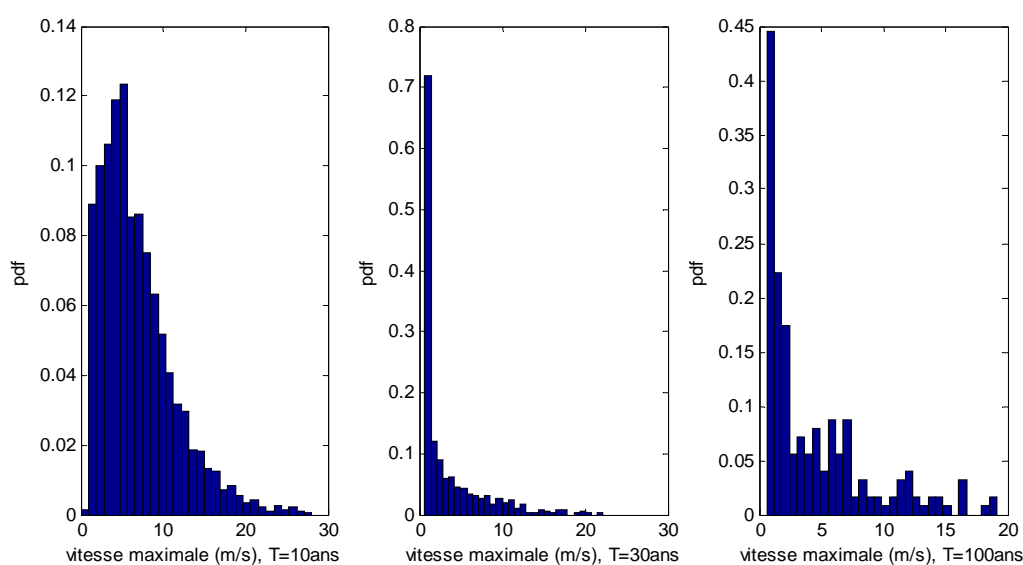
Tableau 6.8 : Hauteurs maximales des événements de référence avec le modèle de Saint-Venant

	$P(x_{stop} \geq x_{stop_T})$	Moy.	SD	q _{2.5}	q _{97.5}
$h_{T_{10}}$ (m)	0.107	1.97	1.35	0.18	5.09
$h_{T_{30}}$ (m)	0.037	1.43	1.26	0.04	4.63
$h_{T_{100}}$ (m)	0.011	1.31	1.20	0.09	4.50

Pour des périodes de retour inférieures ou égales à 100 ans, le mode de la distribution des hauteurs maximales est distinct de zéro. Cependant, sa valeur diminue assez rapidement avec la période de retour puisqu'il est d'environ 0.8 m pour les avalanches centennales, les valeurs maximales prédites étant alors de l'ordre de 4.5 m (Tableau 6.8). Ainsi, même si quelques avalanches de hauteurs assez importantes sont susceptibles d'atteindre des distances d'arrêt élevées, la très grande majorité des événements les plus rares possèdent une épaisseur faible, voire très faible. C'est lié au fait que les avalanches atteignant des distances d'arrêt importantes sont presque toujours caractérisées par un frottement faible qui favorise un écoulement rapide, un étalement de la neige, et ainsi une diminution de la hauteur.

Pour les vitesses maximales, on retrouve pour des périodes de retour de 30 et 100 ans des allures de densités proches de celle observée pour une période de retour de 2 ans, c'est-à-dire strictement décroissante avec la vitesse (Figure 6.11). Les distributions des vitesses des avalanches de période de retour élevées sont donc bien très différentes avec le modèle de Saint-Venant de celles obtenues avec le modèle de Voellmy : beaucoup plus asymétriques, de moyennes nettement moins élevées et caractérisées par des valeurs maximales légèrement moins fortes (Tableau 6.9). Le modèle de Saint-Venant est donc moins pessimiste pour les vitesses susceptibles d'affecter des enjeux situés très loin en zone d'arrêt que le modèle de Voellmy bloc glissant, et ce alors qu'il s'agit en outre avec le modèle de Saint-Venant de vitesses maximales. D'un point de vue opérationnel, c'est plutôt une bonne nouvelle puisque le modèle de Saint-Venant est supposé plus réaliste.

Figure 6.11 : Vitesses maximales et période de retour avec le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les distributions sont conditionnelles à l'atteinte de l'abscisse considérée.

Tableau 6.9 : Vitesses maximales des événements de référence avec le modèle de Saint-Venant

	Moy.	SD	q_{2.5}	q_{97.5}
$v_{T_{10}}$ (m/s)	6.8	4.5	1.4	18.4
$v_{T_{30}}$ (m/s)	3.6	4.0	0.8	15.7
$v_{T_{100}}$ (m/s)	4.4	4.3	0.7	16.1

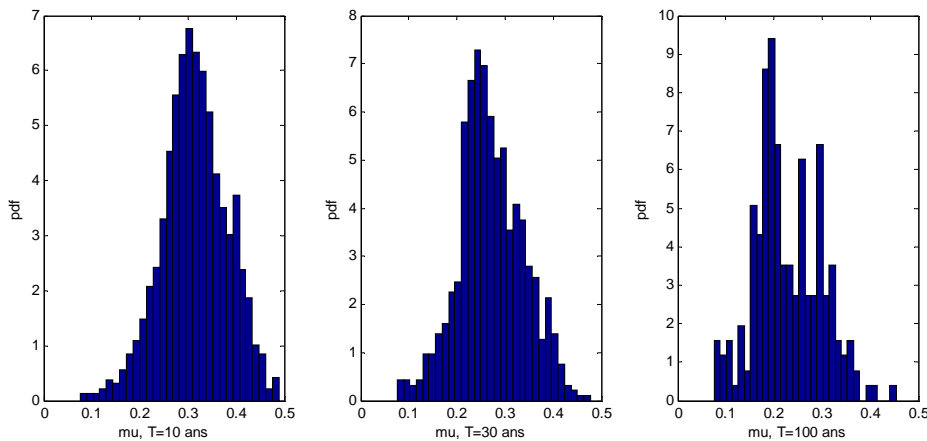
6.1.4.3. Coefficient de frottement

Comme pour le modèle de Voellmy, les valeurs de coefficient de frottement correspondant aux événements de référence peuvent être analysées. On retrouve que la moyenne des distributions obtenues diminue avec la période de retour, tandis que l'écart type reste constant (Tableau 6.10). Par contre, les moyennes obtenues sont plus faibles qu'avec le modèle de Voellmy. Ces valeurs faibles sont en bonne concordance avec les hauteurs maximales faibles observées puisque l'étalement est d'autant plus facile que la neige est fluide et μ_{voel} faible. Elles sont également plus proches des valeurs généralement attribuées par la littérature aux avalanches extrêmes puisque la valeur de 0.16 se situe nettement dans l'intervalle de crédibilité à 95% du coefficient de frottement centennal et se situe encore dans l'intervalle de crédibilité à 95% du coefficient trentennal (Tableau 6.10). Par rapport aux distributions obtenues avec le modèle de bloc glissant, on note également que celles-ci sont plus dispersées, sans doute en partie à cause de la moindre quantité de simulations effectuées, et plus symétriques (Figure 6.12).

Tableau 6.10 : Frottement μ_{voel} des événements de référence avec le modèle de Saint-Venant

	Moy.	SD	q_{2.5}	q_{97.5}
$\mu_{voel_{10}}$	0.314	0.068	0.172	0.441
$\mu_{voel_{30}}$	0.271	0.068	0.134	0.403
$\mu_{voel_{100}}$	0.228	0.070	0.092	0.363

Figure 6.12 : Frottement latent et période de retour avec le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les distributions sont conditionnelles à l'atteinte de l'abscisse considérée.

6.1.4.4. Nombres de Froude

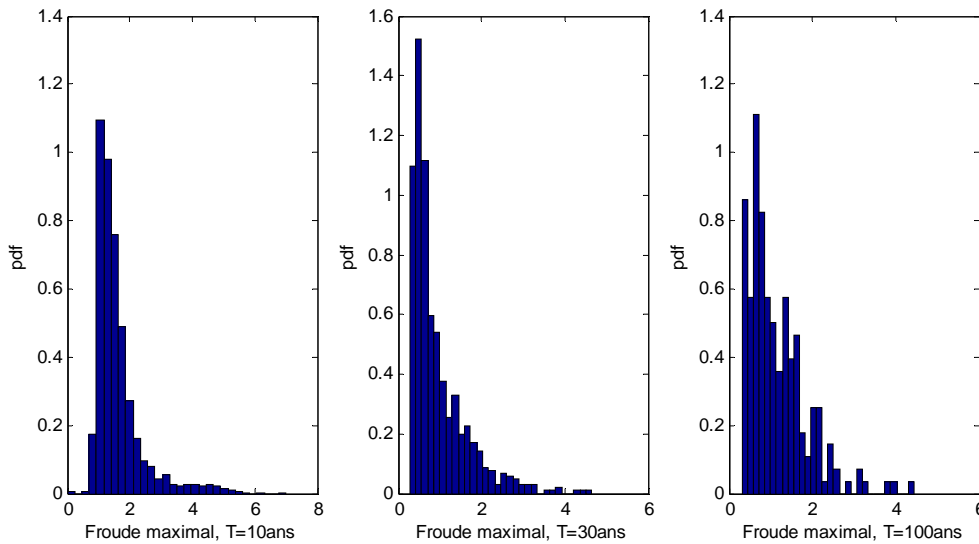
Enfin, hauteurs maximales et vitesses maximales peuvent être combinées pour calculer la distribution des nombres de Froude maximaux. Ceux-ci caractérisent la dynamique des écoulements avalancheux et présentent de ce fait un intérêt pour la construction d'un éventuel ouvrage de protection. Les distributions obtenues montrent, conformément aux distributions de vitesses maximales, une diminution faible du nombre de Froude avec la période de retour (Figure 6.13). Pour les avalanches décennales, des nombres de Froude de 1 à 2.5 sont les plus probables, mais des écoulements encore

très dynamiques avec des nombres de Froude supérieurs à 4 sont possibles. Ceci est à relier à la pente encore significative existant au niveau de l'abscisse décennale qui permet aux avalanches caractérisées par un frottement faible de présenter un écoulement assez dynamique. Pour des périodes de retour plus élevées, la distribution des nombres de Froude se stabilise avec une moyenne autour de 1 traduisant des écoulements nettement moins dynamiques. Néanmoins, des nombres de Froude plus importants de l'ordre de 3 ne sont pas à exclure, même pour des avalanches centennales (Tableau 6.11). A noter que la distribution des nombres de Froude centennaux est même un peu plus élevée en termes de moyenne et de valeurs élevées que celle des nombres de Froude trentennaux du fait du profil de pente, ce qui n'est pas sans conséquences sur les pressions d'impact. On y reviendra au chapitre 6.5.

Tableau 6.11 : Nombres de Froude des événements de référence avec le modèle de Saint-Venant

	Moy.	SD	q _{2.5}	q _{97.5}
$Fr_{T_{10}}$	1.61	0.78	0.88	4.18
$Fr_{T_{30}}$	0.94	0.68	0.29	2.80
$Fr_{T_{100}}$	1.16	0.71	0.33	3.10

Figure 6.13 : Nombre de Froude maximal et période de retour avec le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les distributions sont conditionnelles à l'atteinte de l'abscisse considérée.

6.2. Loix prédictives

6.2.1. Variables du modèle

Au chapitre 5, la loi prédictive du modèle fréquentiel a été calculée pour prédire les fréquences avalanches à l'échelle communale, mais elle n'est pas utile pour le calcul des intervalles de crédibilité associés aux distances d'arrêt de référence. Pour ce dernier, la loi prédictive du modèle d'intensité est par contre nécessaire. Elle s'obtient, conformément à l'équation 3.20, en moyennant le modèle $l(x, y | \theta_M)$ sur la loi *a posteriori* des paramètres $p(\theta_M | data_2)$. Numériquement, comme dit au chapitre 3, le calcul s'effectue en deux étapes. D'abord la simulation de la loi prédictive des variables d'entrée $p(x^{pred} | \theta_M)$ puis celle de la loi prédictive des variables de sortie $p(y^{pred} | x^{pred}, \theta_M) = p(G(x^{pred}, \theta_G))$.

Ces considérations théoriques peuvent être illustrées avec n'importe lequel des modèles de propagation proposés, et l'on privilégie ici le plus simple d'entre eux, le modèle de Coulomb. La loi prédictive des deux variables d'entrée, toute deux observables, peut être simulée simplement en même

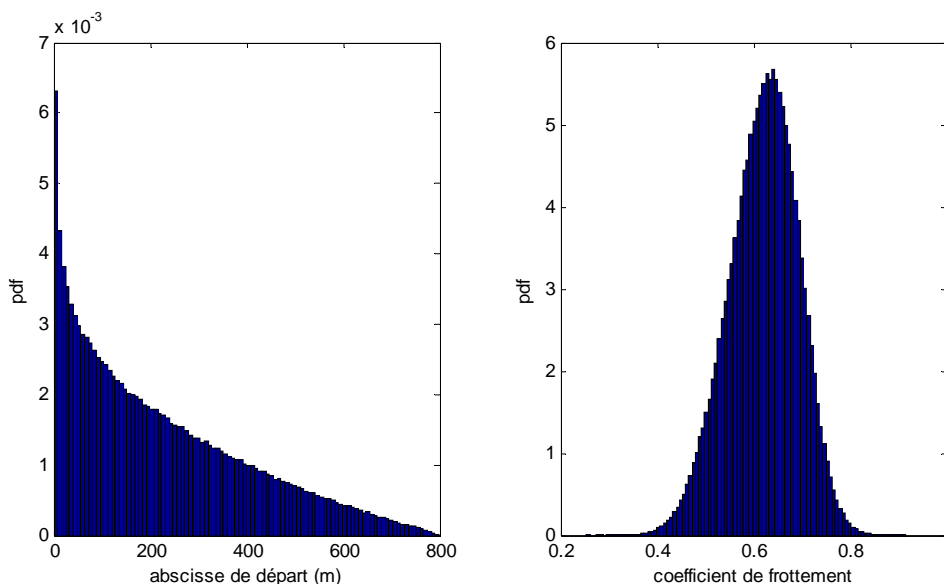
temps que l'inférence est effectuée en échantillonnant le modèle à chaque itération k , c'est-à-dire sachant $\theta_M = \theta_M^{(k)}$. Ensuite, la loi prédictive des distances d'arrêt est calculée de façon analogue à la moyenne *a posteriori* grâce à un nombre suffisant de tirages de Monte Carlo.

Tableau 6.12 : Lois prédictives des variables du modèle de Coulomb

	Moy.	SD	q_{2.5}	q_{97.5}
μ_{coul}	0.631	0.083	0.587	0.692
μ_{coul}^{pred}	0.62	0.072	0.471	0.75
x_{start} (m)	213.9	174.7	28.5	343.1
x_{start}^{pred} (m)	221.5	182.7	3.2	647.3
x_{stop} (m)	1027.6	205.1	401.3	1240.2
x_{stop}^{pred} (m)	1026.9	208	386.5	1246.9

Pour le cas d'étude, les différences entre la moyenne *a posteriori* du modèle et sa distribution prédictive sont très faibles, de sorte qu'il est difficile de les distinguer visuellement. La Figure 6.14 montre par exemple que les distributions prédictives marginales des variables d'entrée ressemblent beaucoup aux distributions moyennes présentées en Figure 4.5. La raison en est que la loi *a posteriori* des paramètres est assez peu dispersée du fait de la quantité de données assez importante utilisée pour l'inférence. Ainsi l'extra variabilité due à l'erreur d'estimation est assez réduite par rapport à la variabilité naturelle des entrées modélisées. Le Tableau 6.12 montre néanmoins que les distributions prédictives marginales sont quand même légèrement plus dispersées que les moyennes *a posteriori* calculées au chapitre 4, et ce à la fois pour les variables d'entrée et de sortie, de sorte que l'effet de l'erreur d'estimation est malgré tout significatif.

Figure 6.14 : Lois prédictives des variables d'entrée du modèle de Coulomb



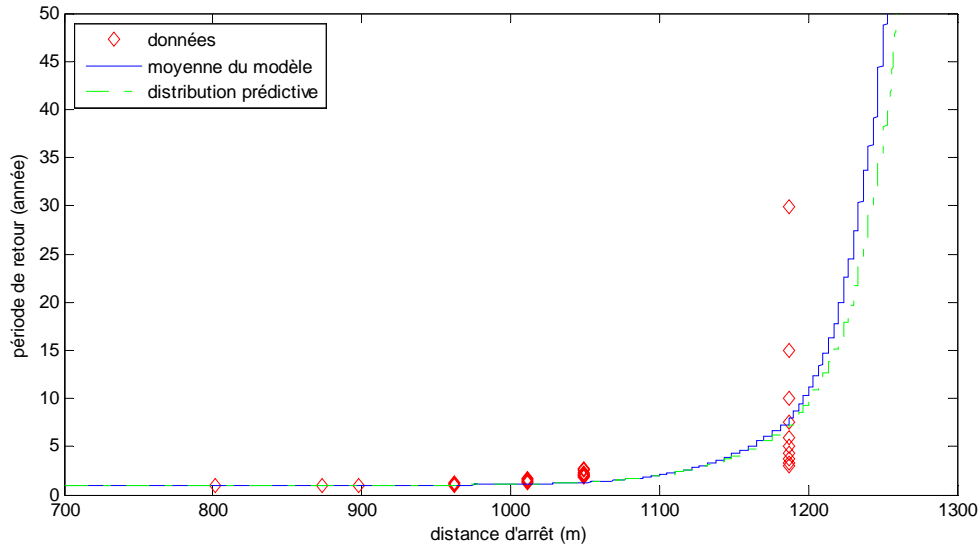
Couloir EPA 5 de la commune de Bessans, Savoie, France.

A titre d'exemple, et même si les périodes de retour prédictives sont d'un usage peu courant dans le domaine des risques naturels, les périodes de retour associées à la loi prédictive des distances d'arrêt ont également été calculées avec le modèle de Coulomb. Du fait de l'erreur d'estimation, celles-ci sont un peu plus pessimistes que les périodes de retour moyennes. Pour ce cas d'étude, la Figure 6.15 montre ainsi que, dès que la période de retour considérée est supérieure à deux ans, la distance d'arrêt de projet est plus importante avec la loi prédictive qu'avec la moyenne du modèle. La différence entre les deux distances d'arrêt reste toutefois modérée même pour de grandes périodes de retour puisqu'elle ne dépasse pas 10 m pour une période de retour de 300 ans (Tableau 6.13).

Tableau 6.13 : Comparaison entre moyenne et loi prédictive avec le modèle de Coulomb

Période de retour T (année)	x_{stop_T} (m), moyenne <i>a posteriori</i>	x_{stop_T} (m), loi prédictive
1	962.3	960.9
2	1099.7	1099.7
5	1159.6	1165
10	1196.8	1200.2
30	1236.9	1243.6
100	1267.9	1277.7
300	1288.1	1296.9

Figure 6.15 : Périodes de retour moyennes et prédictives



Couloir EPA 5 de la commune de Bessans, Savoie, France.

6.2.2. Aléas de référence

6.2.2.1. Avec le modèle de Coulomb

Si les lois prédictives du modèle sont d'un intérêt opérationnel discutable, il n'en va pas de même pour celles des distances d'arrêt de projet. Celles-ci sont en effet indispensables pour juger de la confiance à accorder aux valeurs de projet retenues. Comme détaillé au chapitre 3.4.4, les calculs s'effectuent numériquement en simulant successivement les distributions des distances d'arrêt correspondant à chaque valeur de la loi *a posteriori* et en combinant les distributions obtenues avec la loi *a posteriori* de la fréquence avalancheuse moyenne. Les quantiles d'ordre élevés s'obtiennent conformément à l'équation 3.21. Les quantiles annuels sont obtenus après pondération avec la fréquence avalancheuse moyenne selon l'équation 3.22. Enfin les périodes de retour sont calculées conformément à l'équation 3.23.

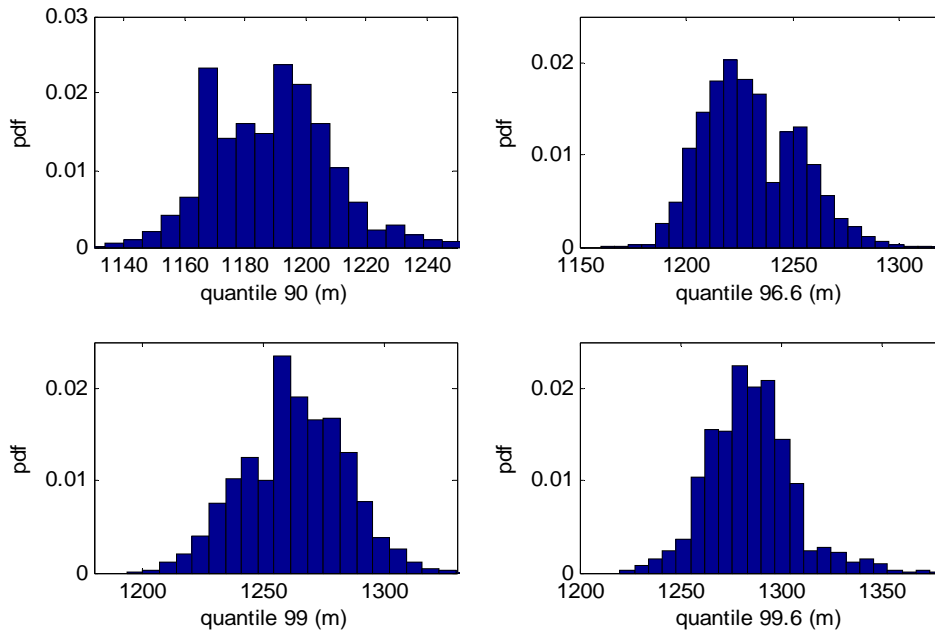
Tableau 6.14 : Quantiles de distance d'arrêt avec le modèle de Coulomb

	Moy.	SD	q _{2.5}	q _{97.5}
$x_{stop_{90}}$ (m)	1189	19.9	1154.1	1230.2
$x_{stop_{96.6}}$ (m)	1230.6	21.9	1193.5	1277
$x_{stop_{99}}$ (m)	1263.3	21.1	1223.5	1303.9
$x_{stop_{99.6}}$ (m)	1284.7	21	1243.6	1334.5

Pour le cas d'étude du modèle de Coulomb, 5000 valeurs de la loi *a posteriori* θ_M ont ainsi été choisies. Et pour chacune d'entre elles, on a simulé un échantillon de 5000 valeurs de $p(x_{stop} | \theta_M^{(k)})$,

soit un total de 25 millions d’itérations du modèle de propagation pour un temps de calcul total d’environ deux jours. La Figure 6.16 présente la distribution prédictive de quatre quantiles d’ordre élevé de distance d’arrêt. Ils présentent des formes en cloche avec toutefois une légère asymétrie à droite. L’irrégularité des densités obtenues est due au calcul discrétisé au pas de 5 m des distances d’arrêt. Bien sûr, la moyenne des distributions augmente avec le quantile. L’écart type reste par contre constant autour de la valeur de 21 m, de sorte que les quatre intervalles de crédibilité à 95 % présentent chacun une largeur d’environ 80 m (Tableau 6.14).

Figure 6.16 : Lois prédictives de différents quantiles avec le modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

La Figure 6.17 présente les distributions prédictives des quantiles annuels les plus fréquemment utilisés pour le zonage, c’est à dire ceux correspondant à des distances d’arrêt de 10, 30, 100 et 300 ans. Ces distributions présentent une moyenne supérieure d’environ 6 m et une variance supérieure d’environ 1 m aux quatre quantiles de la Figure 6.16. (Tableau 6.15).

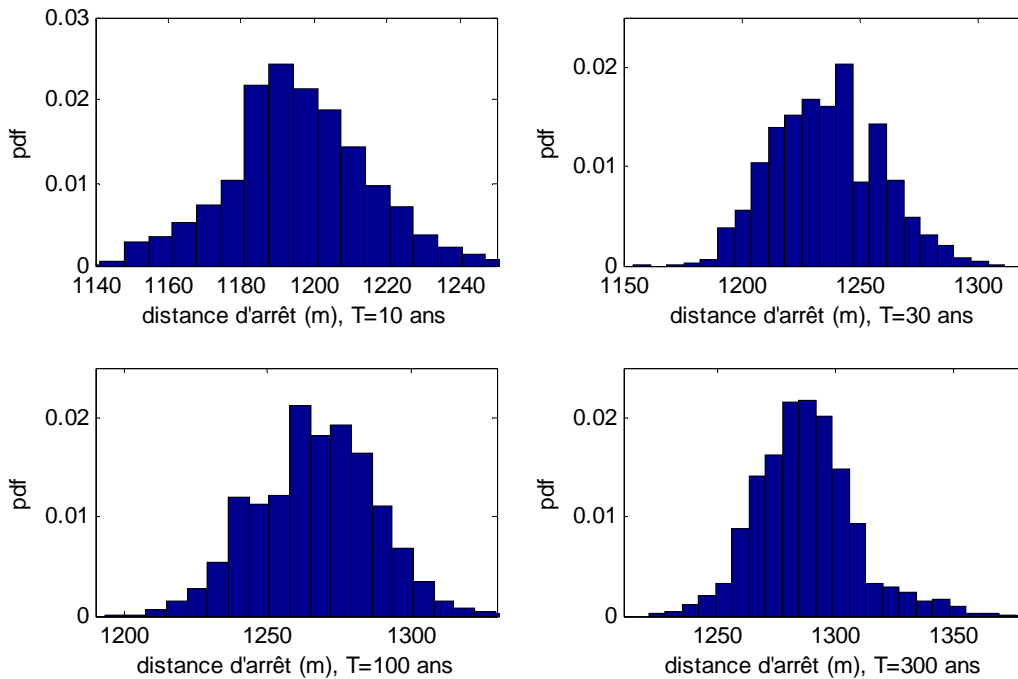
Ces différences entre quantiles et quantiles annuels sont dues à l’influence de la fréquence avalancheuse. En effet, la fréquence avalancheuse moyenne est de 1.21 avalanches par an, de sorte que la distance d’arrêt correspondant à la période de retour T est un peu plus longue que le quantile d’ordre $1 - \frac{1}{T}$. De plus, l’incertitude concernant la fréquence avalancheuse moyenne quantifiée par la

loi *a posteriori* de λ ajoute une source de variabilité supplémentaire accroissant la dispersion de la distribution autour de sa moyenne. Il peut enfin être noté que les moyennes des distributions prédictives correspondent bien aux périodes de retour moyennes précédemment proposées (Tableau 6.1). Cet agrément traduit la forme assez symétrique des distributions des quantiles annuels. Il constitue un argument supplémentaire pour dire que le pari moyen *a posteriori* constitue un estimateur robuste de la valeur de projet recherchée.

Tableau 6.15 : Quantiles annuels avec le modèle de Coulomb

	Moy.	SD	q _{2.5}	q _{97.5}
$x_{stop_{10}}$ (m)	1196.6	21	1159.6	1242.3
$x_{stop_{30}}$ (m)	1236.1	22.5	1196.8	1281.8
$x_{stop_{100}}$ (m)	1267.1	21.2	1226.9	1307.2
$x_{stop_{300}}$ (m)	1287.4	21.5	1247.7	1340.5

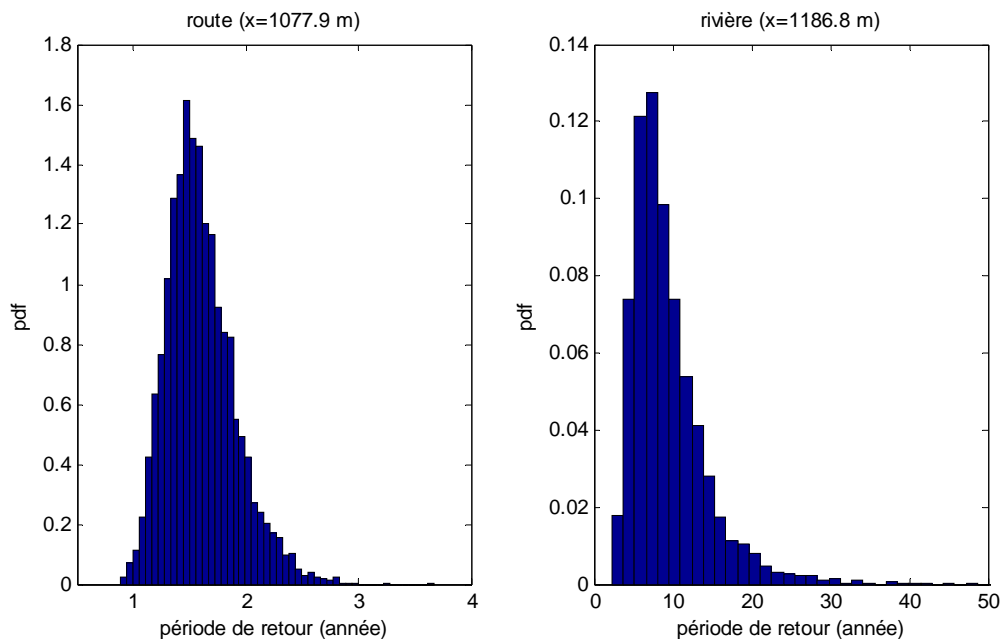
Figure 6.17 : Lois prédictives de différents quantiles annuels avec le modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

Enfin, la Figure 6.18 présente la distribution prédictive de la période de retour associée à deux abscisses remarquables du couloir, celle de la route et celle de la rivière. Les deux distributions obtenues présentent un mode très marqué, soit un an et demi pour la route et une dizaine d'années pour la rivière. Elles présentent également une forte asymétrie à droite. Néanmoins, pour une fois, seule la partie gauche de la distribution présente un intérêt pour le zonage puisqu'il convient avant tout de ne pas surestimer la période de retour associée à une abscisse donnée. Ainsi l'abscisse de la rivière (1187 m) est-elle atteinte par des avalanches de période de retour minimale de 3.7 ans avec un niveau de confiance supérieur à 95% (Tableau 6.16).

Figure 6.18 : Lois prédictives de différentes périodes de retour avec le modèle de Coulomb



Couloir EPA 5 de la commune de Bessans, Savoie, France.

Tableau 6.16 : Périodes de retour avec le modèle de Coulomb

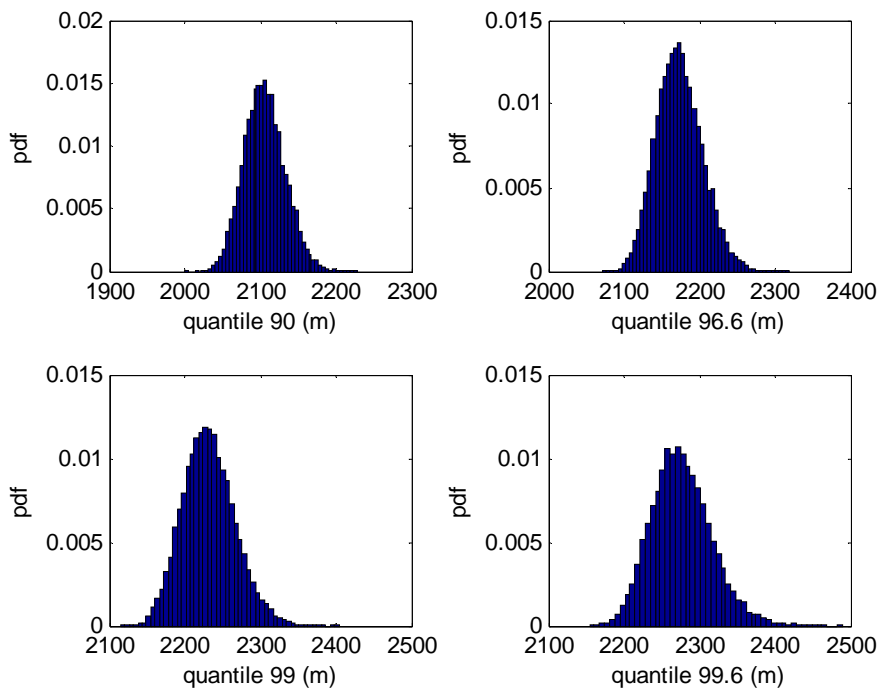
	Moy.	SD	q _{2.5}	q _{97.5}
$T_{x_{route}}$ (années)	1.6	0.3	1.1	2.3
$T_{x_{rivière}}$ (années)	9.4	4.9	3.7	21.8

6.2.2.2. Avec les autres modèles :

Le formalisme général proposé au chapitre 3.4.3 s’applique bien entendu à n’importe lequel des modèles stochastiques proposé et permet d’obtenir toutes les lois prédictives de toutes les variables. On retrouve toujours que, dès lors que la quantité de données est suffisante, celles-ci sont peu différentes des moyennes *a posteriori* tout en étant légèrement plus dispersées. On n’en présentera donc pas d’exemples supplémentaires.

En ce qui concerne les lois prédictives des valeurs de projet, les calculs sont également théoriquement faisables quel que soit le modèle. Avec les modèles de Coulomb bruité et de Voellmy bloc glissant, ils ne posent pas de problèmes particuliers mais les résultats obtenus méritent en revanche d’être discutés. Avec le modèle de Saint-Venant, les calculs n’ont pas été effectués parce que, du fait du nombre très important d’itérations nécessaires, les calculs auraient été extrêmement longs et qu’ils n’auraient vraisemblablement apporté que peu d’information supplémentaire par rapport aux mêmes calculs effectués avec le modèle de Voellmy bloc glissant du fait de la proximité entre ces deux modèles du point de vue de la modélisation stochastique.

Figure 6.19 : Lois prédictives des quantiles de distance d’arrêt avec le modèle de Coulomb bruité



Couloir EPA 13 de la commune de Bessans, Savoie, France.

La Figure 6.19 présente les quatre quantiles d’ordre élevé désormais « traditionnels » obtenus avec le modèle de Coulomb bruité sur le couloir 13 de Bessans. Tout comme pour le couloir 5 de Bessans avec le modèle de Coulomb non bruité, les 4 distributions présentent une forme globale en cloche. Elles sont toutefois nettement plus régulières du fait du terme d’erreur de propagation qui lisse fortement les résultats de la propagation numérique en permettant à n’importe quelle abscisse d’être atteinte. On observe, comme sur le couloir 5 de Bessans, un accroissement logique de la distance d’arrêt avec l’ordre du quantile. De façon plus intéressante, on observe également une augmentation de la variance et de l’asymétrie avec l’ordre du quantile alors qu’elles restaient constantes dans le cas précédent. Ainsi, l’intervalle de crédibilité à 95% est presque centré sur sa moyenne et large de 90 m

pour le quantile d'ordre 90 tandis qu'il est large de 130 m avec 60 m sous la moyenne et 70 m au dessus pour le quantile d'ordre 99.6 (Tableau 6.17).

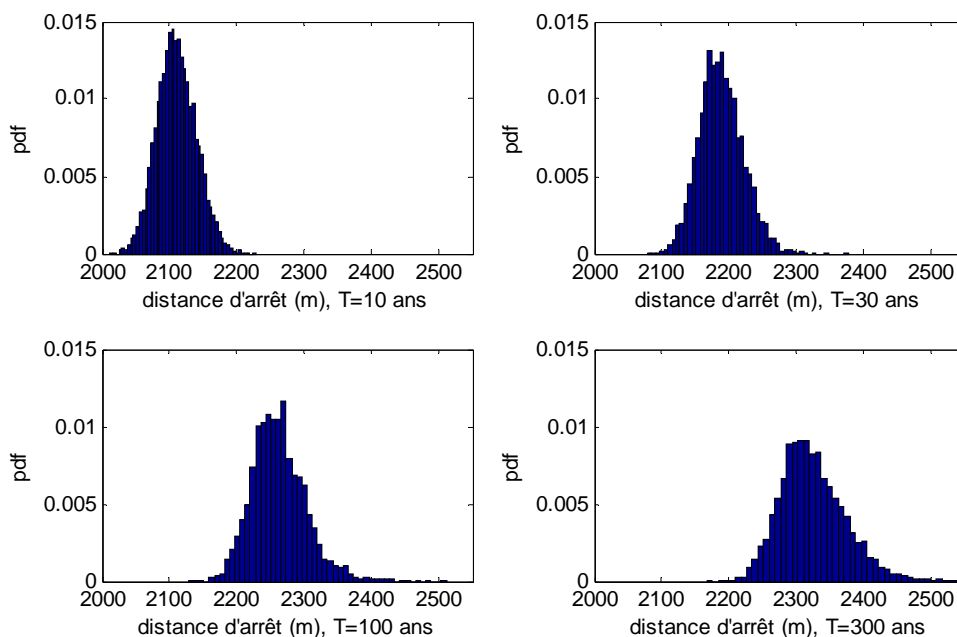
Ces résultats sont cohérents avec ceux obtenus dans l'approche statistique-dynamique implémentée par Barbolini et Savi (2001) : plus la probabilité de non dépassement est faible, plus l'incertitude liée à l'erreur d'estimation est forte et plus l'intervalle de crédibilité associé à la prédiction moyenne est asymétrique. Ce résultat questionne fortement le bien fondé des méthodes de zonage purement déterministes. En outre, il suggère même que des méthodes de zonages basées sur des approximations normales, c'est-à-dire définissant la borne supérieure de l'intervalle de confiance à 95% comme la moyenne de l'estimateur plus deux écarts types, peuvent s'avérer insuffisantes en négligeant la surdispersion de la distribution de l'estimateur vers les fortes valeurs.

Tableau 6.17 : Quantiles de distance d'arrêt avec le modèle de Coulomb bruité

	Moy.	SD	Skew.	q _{2.5}	q _{97.5}
$x_{stop_{90}}$ (m)	2105.2	27	0.28	2062.4	2150.8
$x_{stop_{96.6}}$ (m)	2173.6	30.8	0.32	2125.9	2226.7
$x_{stop_{99}}$ (m)	2231.8	35	0.34	2177.3	2291.9
$x_{stop_{99.6}}$ (m)	2276.2	39.8	0.43	2215.9	2345.6

Pour ce qui est des quantiles annuels avec le modèle de Coulomb bruité, leurs caractéristiques sont très proches des quantiles d'ordre élevés : distributions lisses, forme en cloche, légère asymétrie à droite augmentant avec la période de retour...etc. (Figure 6.20). Notons que l'on n'observe pas de changement en moyenne entre les quantiles et les quantiles annuels car, à la différence du couloir 5 de Bessans, la fréquence avalancheuse moyenne est très proche d'une avalanche par an sur le couloir 13 de Bessans. De plus, on retrouve bien la variance légèrement plus forte des quantiles annuels due à la variabilité additionnelle apportée par le modèle fréquentiel. Enfin, à la différence des résultats obtenus sur le couloir 5 de Bessans, on constate que les périodes de retour moyennes ne coïncident pas avec les moyennes des distributions prédictives des quantiles annuels, la différence s'accroissant même un peu avec la période de retour. C'est un effet de l'asymétrie plus marquée que sur le couloir numéro 5 de Bessans de la distribution des quantiles annuels. Mais l'écart reste suffisamment faible, 4 m pour la distance d'arrêt centennale, pour ne pas remettre en cause l'utilisation de la moyenne *a posteriori* comme aléa de référence (Tableau 6.18).

Figure 6.20 : Lois prédictives des quantiles annuels avec le modèle de Coulomb bruité



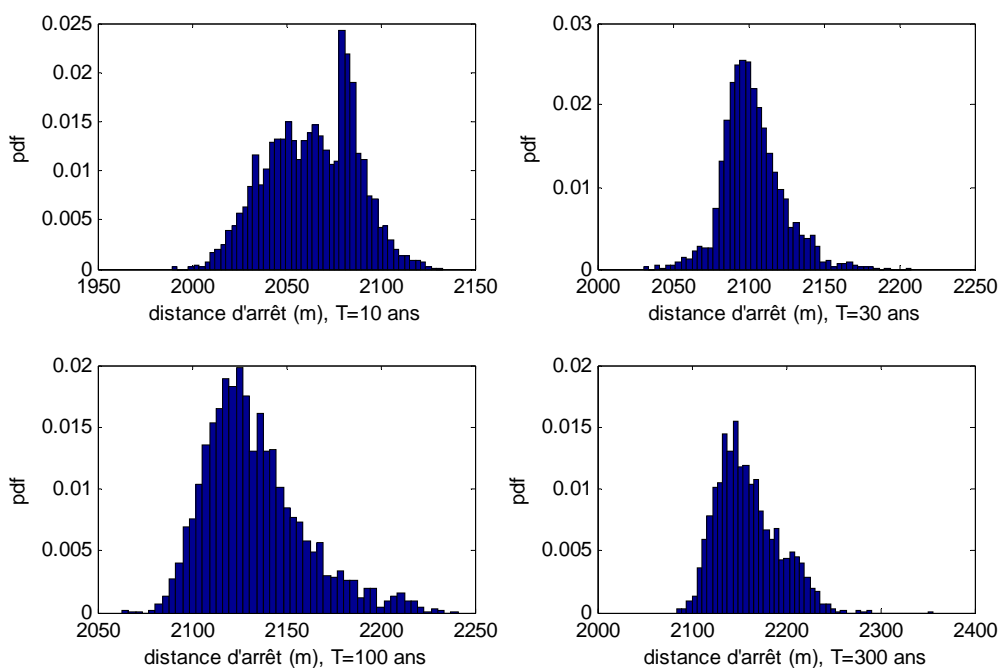
Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 6.18 : Quantiles annuels avec le modèle de Coulomb bruité

	Moyenne a posteriori	Moyenne de la prédictive	SD	Skew.	q_{2.5}	q_{97.5}
$x_{stop_{10}}$ (m)	2106.9	2110.4	29.3	0.17	2055.3	2169.6
$x_{stop_{30}}$ (m)	2181.7	2188.4	33	0.37	2127.9	2256.8
$x_{stop_{100}}$ (m)	2258	2262.3	40.3	0.71	2191.8	2353.8
$x_{stop_{300}}$ (m)	2320.3	2325	49	0.72	2243.3	2434.3

Enfin, pour ce qui est des quantiles annuels sur le même couloir avec le modèle de Voellmy, ils présentent des formes un peu plus irrégulières, notamment la distance d'arrêt décennale qui présente un pic bien marqué vers $x=2080$ m. Cela traduit l'influence visiblement plus forte de la topographie sur les valeurs de projet avec ce modèle de propagation plus physique qu'avec le modèle purement empirique de Coulomb. Le pic dans la distribution de la distance d'arrêt décennale correspond en particulier au début de la zone plane correspondant au fond de la vallée de l'Arc. Mis à part ces effets topographiques, les distributions sont toutefois relativement régulières. Cette régularité est due non pas à un terme d'erreur comme pour le modèle de Coulomb bruité mais au calcul exact de distance d'arrêt permis par le modèle (Figure 6.21).

Figure 6.21 : Lois prédictives des quantiles annuels avec le modèle de Voellmy



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 6.19 : Quantiles annuels avec le modèle de Voellmy bloc glissant

	Moyenne a posteriori	Moyenne de la prédictive	SD	Skew.	q_{2.5}	q_{97.5}
$x_{stop_{10}}$ (m)	2067.4	2063.9	23.6	-0.14	2019	2106.2
$x_{stop_{30}}$ (m)	2101.4	2102.8	19.8	0.57	2065.5	2146.2
$x_{stop_{100}}$ (m)	2130.5	2133.1	26.5	0.91	2093.3	2199.4
$x_{stop_{300}}$ (m)	2158	2159.2	32.8	0.68	2108.9	2229.2

Par rapport au modèle de Coulomb bruité sur le même couloir, on constate, en outre, que les quantiles annuels sont significativement plus faibles et de variance plus réduite. Cela va bien dans le sens déjà

noté lors de l'analyse de la moyenne *a posteriori* des distances d'arrêt d'un comportement nettement plus « optimiste » du modèle de Voellmy bloc glissant. En revanche, on constate, comme avec le modèle de Coulomb, une augmentation de la variance et de l'asymétrie avec la période de retour. Cette dernière est toutefois insuffisante pour qu'une différence significative entre moyenne *a posteriori* et moyenne de la prédictive ne soit notable (Tableau 6.19).

6.3. Analyse de sensibilité

Une analyse de sensibilité complète des différents modèles de prédétermination proposés nécessite l'investigation de leur sensibilité à l'information *a priori*, au modèle de propagation, aux opérateurs stochastiques décrivant la variabilité des entrées/sorties, aux différentes variables, à la quantité de données utilisée pour l'inférence, etc. Tous ces aspects ne peuvent être explorés et présentés exhaustivement. Seuls quelques points marquants permettant de comprendre le fonctionnement des différents modèles de prédétermination proposés et de les comparer sont donc discutés. L'analyse est essentiellement effectuée en termes de prédiction, mais aussi en termes d'inférence lorsque celle-ci est concernée.

6.3.1. Lois *a priori*

6.3.1.1. Priors sur le modèle de mélange

Le transfert d'information aux couloirs peu documentés (Sigurdsson et al., 1998) est souvent effectué par les spécialistes des avalanches en mélangeant des données issues de différents sites (McClung et al., 1989) ou par l'intermédiaire de valeurs tabulées des coefficients de frottement (Salm et al., 1990). L'utilisation du modèle de Coulomb sur le couloir EPA 5 de Bessans a été l'occasion de proposer une méthode de transfert d'information moins brutale basée sur la mise à jour d'un *prior* construit sur un couloir d'apprentissage. Les différences nettes observées entre *priors* et *posteriors* marginaux, notamment pour les paramètres c et a_1 (Figure 4.4), ont cependant montré que le transfert d'information peut être délicat du fait de la difficulté à choisir un couloir d'apprentissage approprié sur la base de critères purement géographiques et topographiques.

D'un point de vue opérationnel, la robustesse du résultat au choix du *prior* peut être investiguée en refaisant les calculs avec un *prior* très peu informatif comme celui utilisé pour le couloir d'apprentissage (Tableau 4.2, gauche). La comparaison des Tableaux 4.3 et 6.20 montre qu'utiliser un *prior* vague plutôt qu'un *prior* informatif augmente logiquement la variance *a posteriori*. Néanmoins, l'augmentation reste modérée du fait de la quantité déjà assez importante de données disponibles. La différence entre les lois *a posteriori* marginales obtenues avec les deux *priors* est, ainsi qu'on pouvait s'y attendre, surtout notable pour c , et dans une moindre mesure pour a_1 .

Tableau 6.20 : Statistiques descriptives des lois *a posteriori*, *prior* vague, modèle de Coulomb

	Moy.	SD	q _{2.5}	q _{97.5}
a_1	0.73	0.15	0.47	1.05
a_2	2.20	0.54	1.28	3.36
c	0.720	0.014	0.691	0.747
d	-0.51	0.07	-0.64	-0.38
σ	0.05	0.01	0.04	0.06

Concernant les quantiles annuels, les variances sont seulement légèrement supérieures avec le *prior* vague (Tableau 6.21) qu'avec le *prior* informatif (Tableau 6.15). De façon plus surprenante, les quantiles annuels moyens sont plus élevés avec le *prior* informatif qu'avec le *prior* vague, avec par exemple une différence de 7 m sur le quantile centennal. Cela indique que l'information apportée par le couloir d'apprentissage plaide pour des valeurs de projets plus pessimistes que les données. L'apport d'information se manifeste essentiellement au niveau du paramètre c , c'est-à-dire du terme constant du modèle de mélange gaussien. Avec le *prior* vague, son estimateur ponctuel est 0.72. Il décroît à 0.683 avec le *prior* informatif, permettant ainsi aux avalanches d'aller en moyenne plus loin du fait de la réduction du frottement.

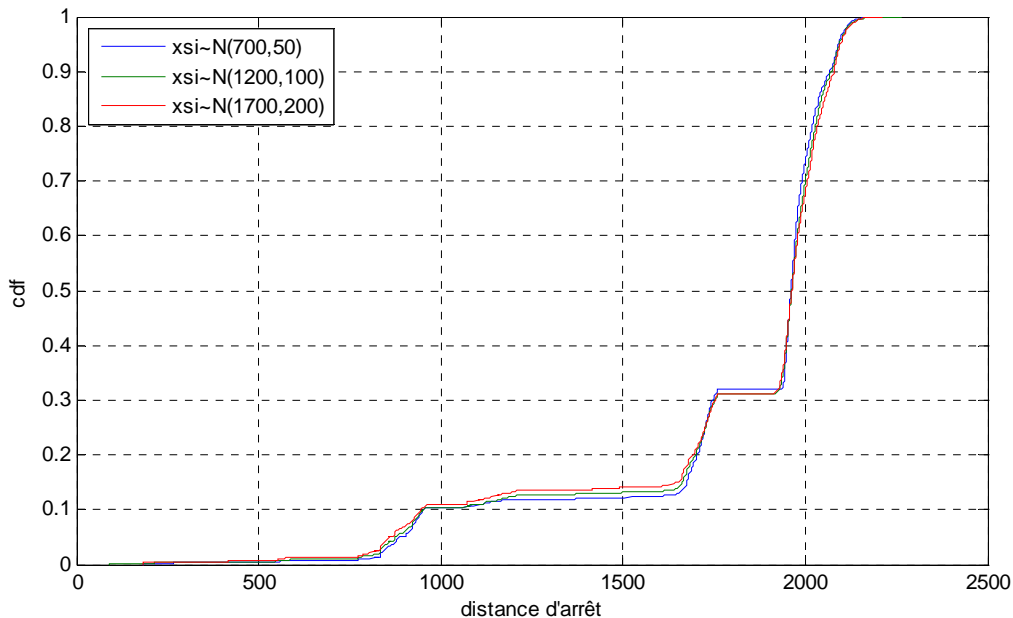
Tableau 6.21 : Quantiles annuels avec un prior vague pour le modèle de Coulomb

	Moy.	SD	q _{2.5}	q _{97.5}
$x_{stop_{10}}$ (m)	1185.6	21.2	1143.3	1226.9
$x_{stop_{30}}$ (m)	1225.5	21.7	1190.1	1269.7
$x_{stop_{100}}$ (m)	1256.6	21.2	1216.2	1295.3
$x_{stop_{300}}$ (m)	1276.4	20.4	1236.8	1317

6.3.1.2. Prior sur le coefficient de frottement turbulent du modèle de Voellmy bloc glissant

Pour le modèle de mélange, la sensibilité au *prior* est plutôt une chance d'un point de vue opérationnel car l'inférence est possible avec des *priors* informatifs ou non informatifs. L'introduction de connaissance extérieure aux données est donc possible sans être obligatoire, ce qui offre un grand nombre de possibilités à l'ingénieur comme au décideur. Pour le coefficient de frottement turbulent, les choses sont un peu différentes dans la mesure où un *prior* assez informatif semble en pratique indispensable pour assurer la convergence dans un temps raisonnable. La sensibilité au *prior* doit donc être investiguée afin de vérifier la robustesse des résultats obtenus au choix du *prior*.

Figure 6.22 : Sensibilité au prior sur ξ_{voel} de la moyenne a posteriori des distances d'arrêt



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant.

On termine ici les calculs entamés au chapitre 4.2, c'est-à-dire que l'on calcule la distribution moyenne *a posteriori* des distances d'arrêt sur le couloir 13 de Bessans avec les trois différents *priors* sur ξ_{voel} . Graphiquement, la sensibilité de la distribution des distances d'arrêt au *prior* sur ξ_{voel} semble faible puisque les trois distributions moyennes *a posteriori* obtenues sont quasiment confondues (Figure 6.22).

L'analyse des quantiles annuels montre qu'une certaine sensibilité existe tout de même, et qu'en particulier les quantiles annuels moyens semblent croître avec la moyenne *a priori* de ξ_{voel} . Ainsi la distance d'arrêt centennale est de 2126.6 m avec un *prior* centré sur 700 (Tableau 6.22), de 2130.5 m avec un *prior* centré sur 1200 (Tableau 6.19), et de 2137.7 m avec un *prior* centré sur 1700 (Tableau 6.23). Cette augmentation est compatible avec le fait que des valeurs plus élevées de ξ_{voel} se traduisent par une réduction du frottement. Néanmoins, l'accroissement reste modéré, 10 m tout au plus, de sorte que la sensibilité au *prior* sur ξ_{voel} reste peu marquée, d'autant que les différences entre les intervalles de crédibilité obtenus avec les trois *priors* sont très nettement non significatives au seuil des 95%. Au

final, la nécessité d'utiliser un *prior* assez informatif sur ξ_{voel} n'est donc pas un obstacle rédhibitoire à l'utilisation d'un modèle avec une loi de frottement de Voellmy, même quand l'information extérieure aux données disponible est peu sûre.

Tableau 6.22 : Quantiles annuels avec un prior $N(700,50)$ sur ξ_{voel}

	Moy.	SD	q _{2.5}	q _{97.5}
$x_{stop_{10}}$ (m)	2058	25.2	2012.1	2100
$x_{stop_{30}}$ (m)	2098.2	18.9	2058.7	2140.4
$x_{stop_{100}}$ (m)	2126.6	25.9	2089.3	2197.5
$x_{stop_{300}}$ (m)	2152.4	33.4	2103.1	2226.7

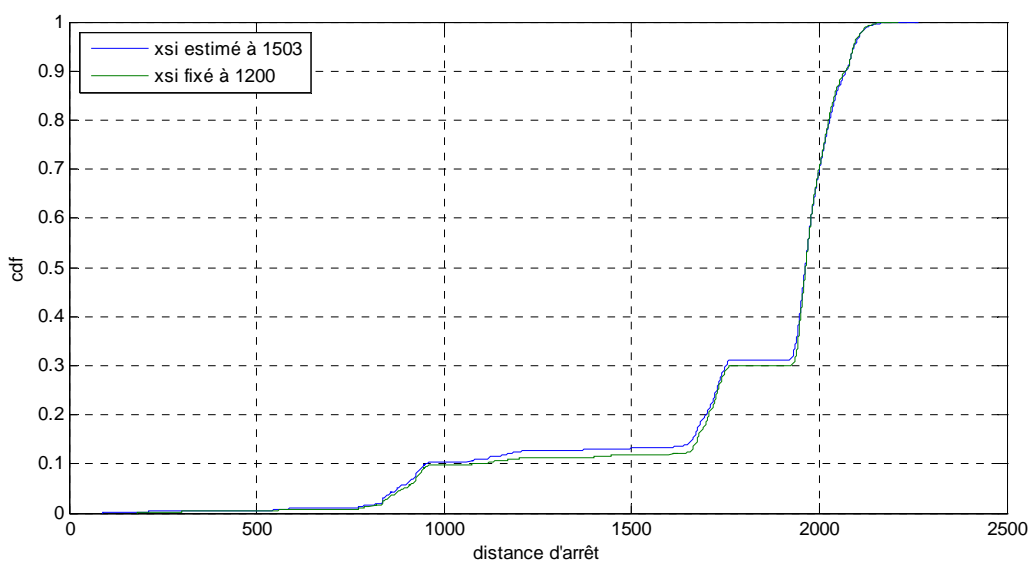
Tableau 6.23 : Quantiles annuels avec un prior $N(1700,200)$ sur ξ_{voel}

	Moy.	SD	q _{2.5}	q _{97.5}
$x_{stop_{10}}$ (m)	2068.6	23.2	2023.1	2109.6
$x_{stop_{30}}$ (m)	2106.8	20.4	2071.2	2154.2
$x_{stop_{100}}$ (m)	2137.7	27.2	2095.6	2204.8
$x_{stop_{300}}$ (m)	2164.3	32.7	2111.4	2232.6

6.3.1.3. Et si on fixe ξ_{voel} ?

Au chapitre 3.2., on avait suggéré en réponse aux difficultés d'inférence sur ξ_{voel} de fixer ce paramètre. On avait constaté sur le cas d'étude retenu que cela n'affectait pas l'inférence des autres paramètres. On examine ici les conséquences de la fixation de ξ_{voel} sur la prédiction moyenne des variables de sortie. La Figure 6.23 montre qu'en termes de distribution des distances d'arrêt, la différence est quasi nulle selon que l'on estime ξ_{voel} ou que l'on le fixe, et ce tout particulièrement pour des distances d'arrêt correspondant à des probabilités de non dépassement supérieures à 0.4.

Figure 6.23 : Fonction de répartition des distances d'arrêt avec ξ_{voel} estimé ou fixé



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant.

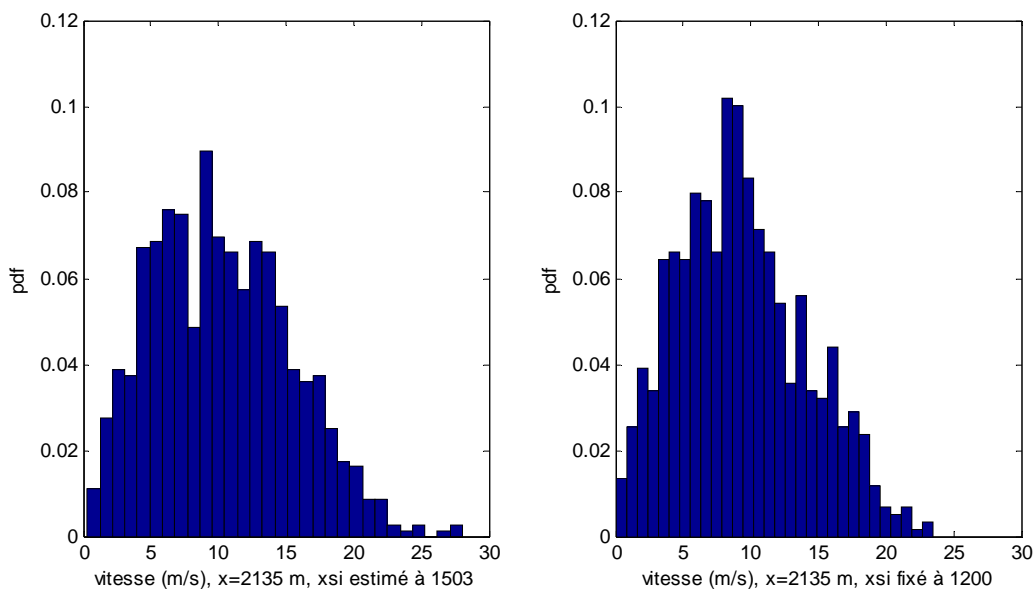
De façon encore plus remarquable, fixer ξ_{voel} plutôt que l'estimer ne semble donc pas affecter la distribution des aléas de référence, la comparaison des Tableaux et 6.19 et 6.24 ne permettant de déceler aucune différence ni en moyenne ni même en variance pour les quatre quantiles annuels proposés.

Tableau 6.24 : Quantiles annuels avec ξ_{voel} fixé à 1200

	Moy.	SD	Q _{2.5}	Q _{97.5}
$x_{stop_{10}}$ (m)	2060.7	24.7	2015.6	2104.3
$x_{stop_{30}}$ (m)	2100.2	20.1	2059.4	2145
$x_{stop_{100}}$ (m)	2129.6	26.9	2089.9	2200.2
$x_{stop_{300}}$ (m)	2155.5	33.7	2103.9	2227.6

La seule différence notable si l'on fixe ξ_{voel} plutôt que de l'estimer apparaît lorsqu'on analyse dans les deux cas les distributions des vitesses. La Figure 6.24 montre en effet que fixer ξ_{voel} conduit à légèrement sous-estimer les vitesses des avalanches au niveau de l'abscisse centennale. Cette sous-estimation s'explique par la différence entre la valeur fixée (1200) et estimée (1503) de ξ_{voel} , une valeur élevée de ce paramètre favorisant des vitesses plus élevées en réduisant le frottement. D'un point de vue opérationnel, fixer ξ_{voel} pour gagner du temps de calcul et réduire les difficultés de réglage de l'algorithme semble donc possible à condition d'être très confiant dans la valeur de ξ_{voel} utilisée et prudent par rapport aux distributions de vitesses obtenues.

Figure 6.24 : Distribution des vitesses avec ξ_{voel} estimé ou fixé



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant.

6.3.2. Modèle de vraisemblance

Au chapitre 4, la corrélation entre erreur de propagation ε et coefficient de frottement μ_{coul} obtenue après inférence du modèle de Coulomb bruité n'a pas été significative. De même, les corrélations entre variables d'entrée calculées avec les deux autres modèles se sont révélées soit peu significatives (avec le modèle de bloc glissant) soit pas du tout significatives (avec le modèle de Saint-Venant) tout en posant des problèmes de biais durant l'inférence. On peut donc s'interroger sur la nécessité d'utiliser un modèle permettant la prise en compte explicite de toutes les corrélations entre variables d'entrée.

L'inférence du modèle de Coulomb bruité a de ce fait également été effectuée avec 3 modèles de vraisemblance concurrents proposant une description plus simple de la variabilité des entrées. On note M_0 le modèle « complet » utilisé jusqu'ici. Le modèle M_1 diffère du modèle M_0 par la non prise en compte des corrélations avec l'abscisse de départ. Le couple μ_{coul}, ε est alors simplement modélisé par une loi binormale de moyenne (c, e) et de variance covariance Σ (eq. 6.2). Le modèle M_2 se distingue du précédent par la non prise en compte de la covariance entre coefficient de frottement et erreur de propagation. On est alors ramené à un modèle avec trois variables d'entrée indépendantes (eq. 6.3). Le modèle M_3 se distingue du précédent par le remplacement du modèle normal pour μ_{coul} par un modèle Beta. L'idée était qu'un modèle borné et si nécessaire asymétrique pouvait être remplacé avantageusement un modèle normal assez rigide pour capturer la variabilité des coefficients de frottement « observés ». Une écriture des paramètres de la loi Beta en fonction de sa moyenne c et de sa variance Σ_{11} est possible (eq. 6.4).

$$(\mu_{coul_i}, \varepsilon_i | c, e, \Sigma) \sim N_2 \left(\begin{pmatrix} c \\ e \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix} \right) \text{ (eq. 6.2)}$$

$$(\mu_{coul_i}, \varepsilon_i | c, e, \Sigma) \sim N_2 \left(\begin{pmatrix} c \\ e \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \right) \text{ (eq. 6.3)}$$

$$(\mu_{coul_i}, \varepsilon_i | c, e, \Sigma) \sim Beta \left(\frac{c}{\Sigma_{11}} (c - c^2 - \Sigma_{11}), \frac{1}{\Sigma_{11}} (c - c^2 - \Sigma_{11}) (1 - c) \right) \times N(e, \Sigma_{22}) \text{ (eq. 6.4)}$$

Les facteurs de Bayes entre les quatre modèles concurrents calculés selon la méthode de Kaas et Raftery présentée au chapitre 2 montrent que le modèle M_0 est significativement le meilleur. Les non prises en compte des corrélations intervariables dégradent donc toutes le modèle en lui retirant de la souplesse. Enfin, le modèle Beta se révèle entièrement inadapté puisqu'il est très nettement battu par n'importe lequel des trois autres modèles (Tableau 6.25).

Tableau 6.25 : Facteurs de Bayes pour le choix du modèle de Coulomb bruité

Facteur de Bayes	Modèle M_1	Modèle M_2	Modèle M_3
Modèle M_0	4.9	19.8	211.1
Modèle M_1	/	4	42.6
Modèle M_2	/	/	10.7

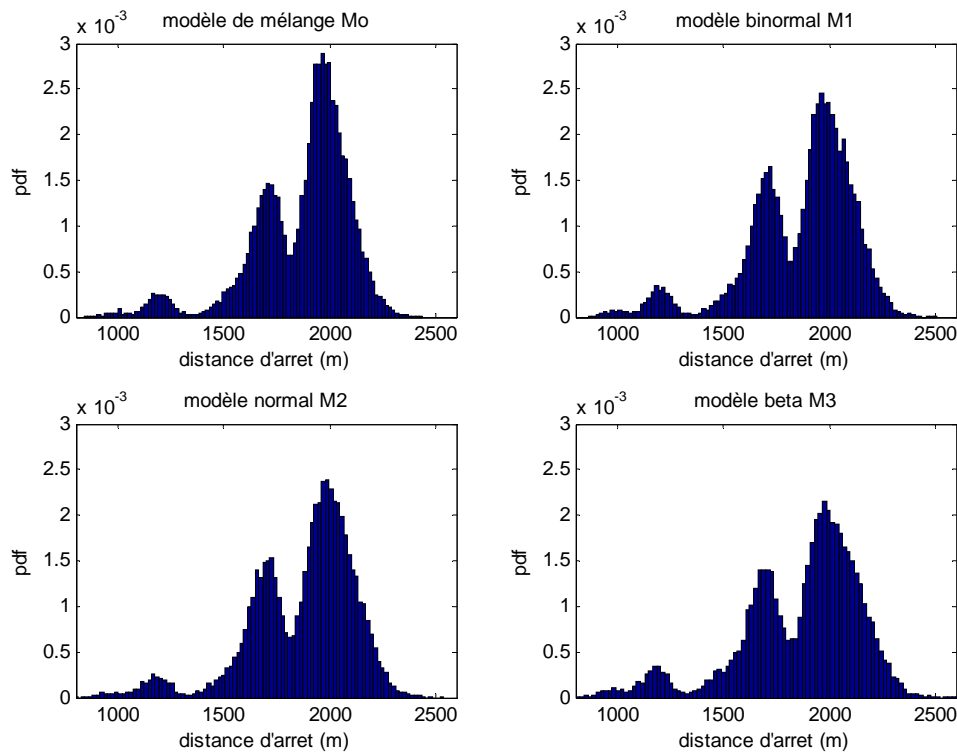
En plus du critère d'adéquation en inférence, les distributions prédictives des distances d'arrêt obtenues avec les quatre modèles ont été comparées (Figure 6.25). Il apparaît que la variance de la distribution est la plus faible avec le modèle de mélange M_0 , de sorte que c'est avec ce modèle que, pour un quantile d'ordre donné, la distance d'arrêt est la plus faible. C'est aussi avec le modèle de mélange M_0 que l'asymétrie de la distribution des distances d'arrêt est la plus forte. Ces deux points suggèrent que la souplesse du modèle de mélange lui permet de mieux s'adapter aux données et au terrain. Ainsi on évite de proposer des valeurs de projet trop fortes. Notons que la surdispersion des distances d'arrêt est particulièrement marquée avec le modèle Beta M_3 qui propose un quantile d'ordre 99 de distance d'arrêt supérieur de 76 m au modèle de mélange M_0 (Tableau 6.26).

Tableau 6.26 : Loi prédictive des distances d'arrêt et modèle de mélange

	Mélange M_0	Binormal M_1	Normal M_2	Beta M_3
Moy. (m)	1870.2	1864.9	1871.3	1863.1
SD (m)	231.4	252.9	250.1	282
Skew.	-1.25	-1.13	-1.1	-1.01
q₁ (m)	1121.7	1076.2	1078.3	1008.3
q₁₀ (m)	1603.4	1581.1	1587.9	1520.6
q₂₅ (m)	1732.8	1715.5	1715.7	1697.1
q₅₀ (m)	1931.4	1929.1	1931.3	1932.3
q₇₅ (m)	2023.2	2038.4	2044	2060.3
q₉₀ (m)	2105.6	2130.1	2136.8	2163.8
q₉₉ (m)	2246.3	2273.1	2279.3	2322.5

Cette brève étude a été menée dans une perspective fermée. De plus, elle a été effectuée avec un seul des modèles stochastiques proposés et elle aurait pu être moins concluante avec les autres modèles, notamment avec celui de Saint-Venant pour lequel les corrélations intervariables sont apparues moins significatives qu’avec le modèle de Coulomb. Il n’en reste pas moins qu’un modèle de mélange semble effectivement offrir davantage de souplesse dans l’adéquation aux données et éviter ainsi une surdispersion des résultats en prédiction. En ce sens, il semble donc bien préférable de privilégier une prise en compte des corrélations plutôt que d’imposer un formalisme indépendant, quitte à rejeter si nécessaire la corrélation après inférence.

Figure 6.25 : Sensibilité de la loi prédictive des distances d’arrêt au modèle de mélange



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle de Coulomb bruité.

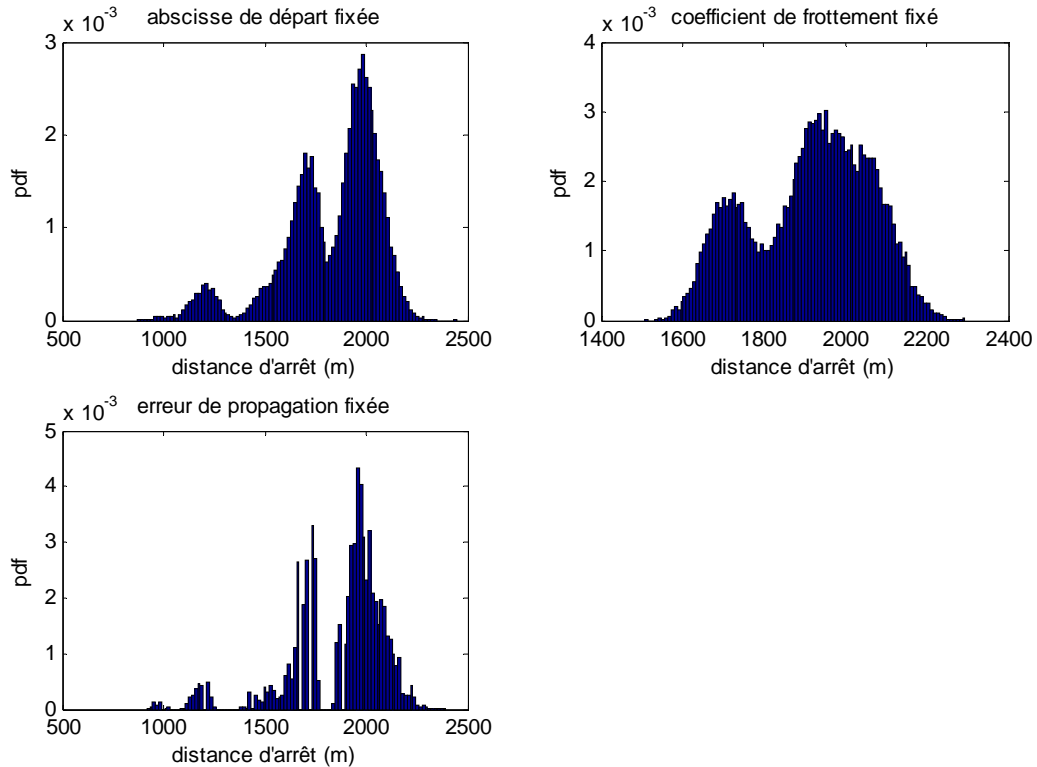
6.3.3. Variables d’entrée

6.3.3.1. Avec le modèle de Coulomb bruité

Pour étudier la sensibilité des modèles aux différentes variables d’entrée, le même exemple a été repris, c’est-à-dire celui du modèle de Coulomb bruité appliqué au couloir EPA 13 de Bessans. Successivement chacune des trois variables du triplet $x_{start}, \mu_{coul}, \varepsilon$ a été fixée à sa moyenne *a posteriori* et la distribution prédictive des distances d’arrêt a été recalculée. Il apparaît en premier lieu que l’allure générale bimodale de la distribution est toujours conservée, de sorte qu’elle est attribuable à la topographie du couloir bien plus qu’à une variable d’entrée en particulier (Figure 6.26).

Chacune des trois variables a cependant un effet différent sur la loi prédictive des distances d’arrêt. Avec un coefficient de frottement constant, la variance de la distribution des distances d’arrêt est considérablement diminuée : l’écart type passe de 241.4 m à 147 m et, en conséquence, le quantile d’ordre 99 passe de 2246.3 m à 2188 m. Avec une erreur de propagation constante, la distribution des distances d’arrêt devient discontinue du fait de la discrétisation numérique du calcul de propagation et de l’imperfection du modèle de propagation l’empêchant d’atteindre certaines abscisses. Néanmoins, la distribution n’est pas modifiée de façon significative à l’exception des quantiles d’ordres les plus élevés qui sont également diminués, le quantile d’ordre 99 passant par exemple de 2246.3 m à 2230.8 m. Enfin, à abscisse de départ constante, la distribution des distances d’arrêt est encore moins modifiée : seuls les quantiles d’ordres les plus élevés sont diminués, le quantile d’ordre 99 passant de 2246.3 m à 2189.2 m (Tableau 6.27).

Figure 6.26 : Sensibilité aux variables d'entrée avec le modèle de Coulomb bruité



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Tableau 6.27 : Sensibilité aux variables d'entrée du modèle de Coulomb bruité

	Moy. (m)	SD (m)	q ₉₀ (m)	q _{96.6} (m)	q ₉₉ (m)	q _{99.6} (m)
x_{start} fixé	1828.7	241.4	2076	2135.6	2189.2	2228.7
μ_{coul} fixé	1913.9	147	2095.8	2147.4	2188	2218
ϵ fixé	1870.2	232.3	2103.6	2159.7	2230.8	2278.6
Loi prédictive	1870.2	231.4	2105.6	2178.9	2246.3	2294.8

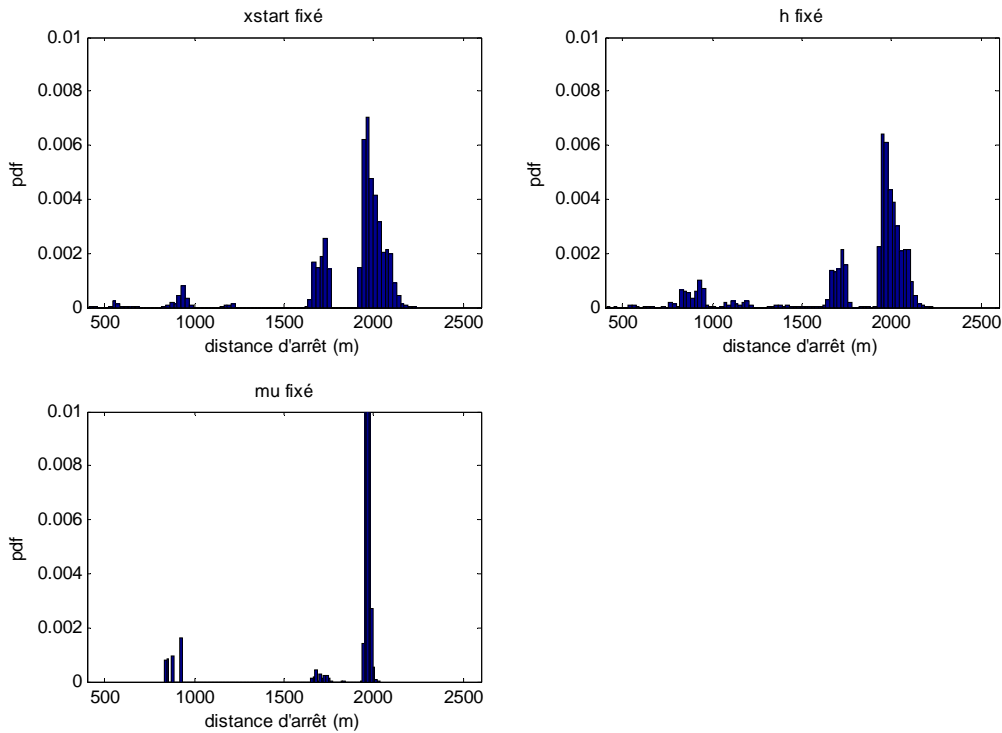
Cette brève analyse montre que, dès lors qu'une source de variabilité, même peu importante pour le résultat final, est négligée, les quantiles d'ordre élevés de distances d'arrêt sont sous estimés, ce qui évidemment doit être évité pour le zonage de l'aléa. La modélisation se doit donc de prendre en compte toutes les sources de variabilité, même celles qui semblent avoir un impact mineur sur la distribution des distances d'arrêt.

Il apparaît néanmoins que le coefficient de frottement joue un rôle primordial par rapport aux deux autres variables. Le modèle de Coulomb bruité proposé s'étant révélé très bien adapté aux données de coefficient de frottement et moins efficace pour capturer la variabilité des abscisses de départ et des erreurs de propagation, ce résultat justifie *a posteriori* son choix et suggère qu'il est vraisemblablement suffisant pour obtenir des résultats opérationnels satisfaisants malgré l'imperfection de la modélisation de deux des trois variables.

6.3.3.2. Avec le modèle de Voellmy bloc glissant

Le rôle crucial du coefficient de frottement est sans doute d'autant plus accentué avec le modèle de Coulomb bruité qu'il n'y en a qu'un seul dans la loi de frottement. La même analyse de sensibilité a donc été effectuée avec le modèle de Voellmy bloc glissant possédant deux coefficients de frottement mais un seul d'entre eux, μ_{voel} , variant d'une avalanche à l'autre. Les distributions moyennes *a posteriori* des distances d'arrêt ont donc été recalculées en fixant tour à tour chacune des trois variables d'entrée du triplet x_{start}, h, μ_{voel} .

Figure 6.27 : Sensibilité aux variables d'entrée avec le modèle de Voellmy bloc glissant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Globalement, seule la fixation du coefficient de frottement modifie fortement l'allure de la distribution des distances d'arrêt en en réduisant considérablement la variance, ce qui conduit à sous estimer les distances d'arrêt de périodes de retour élevées (Figure 6.27). Le rôle prépondérant du coefficient de frottement variant d'une avalanche à l'autre, qu'il soit observable avec le modèle de Coulomb ou latent avec le modèle de Voellmy bloc glissant, semble donc indéniable quel que soit le modèle de propagation utilisé. De même, on retrouve que fixer l'abscisse de départ ne modifie pas beaucoup la distribution des sorties du modèle. Cela suggère que la propagation fait « oublier » les conditions initiales. Néanmoins la variabilité des distances d'arrêt s'en trouve à nouveau quelque peu sous estimée. Enfin, il apparaît au contraire que fixer la hauteur de bloc a très peu d'influence sur la distribution des distances d'arrêt, celle-ci restant quasi inchangée (Tableau 6.28).

Tableau 6.28 : Sensibilité aux variables d'entrée du modèle Voellmy bloc glissant

	Moy. (m)	SD (m)	q _{2.5} (m)	q _{97.5} (m)
x_{start} fixé	1862.8	317.1	891.4	2113.9
h fixée	1803.6	380.3	843.7	2111.3
μ_{voel} fixé	1905.4	238.8	872	1988.2
Moyenne du modèle	1806.2	375.5	834.7	2109.9

6.3.4. Erreur d'estimation et quantité de données

La sensibilité du modèle à la quantité de données utilisée pour l'inférence est un point critique pour une utilisation pratique du fait du nombre important de couloirs pour lesquels le nombre d'événements enregistrés n'est pas très important. La conséquence naturelle de la diminution de la quantité de données est une augmentation des intervalles de crédibilité lors de l'inférence. Celle-ci apparaît très bien dans la différence déjà présentée entre moyenne *a posteriori* et loi prédictive du modèle. On s'intéresse ici plutôt à la sensibilité des aléas de référence proposés à la quantité de données utilisées. On cherche à évaluer grossièrement dans quelle mesure les résultats obtenus avec les différents modèles de prédétermination sont robustes face à une diminution du nombre de données et s'il existe, pour chacun des modèles proposés, une quantité de donnée minimale avant d'observer de gros dysfonctionnements.

6.3.4.1. Avec le modèle de Coulomb bruité

L'effet de la quantité de données sur le modèle de Coulomb bruité a été investigué en remplaçant les 40 données utilisées jusqu'ici par les 26 données utilisées avec les modèles de Voellmy bloc glissant et Saint-Venant. Comme signalé précédemment, les 26 données ne sont pas totalement incluses dans les 40 données, avec notamment la plus forte valeur de distance d'arrêt de l'échantillon historique qui appartient à l'échantillon de 26 données mais pas à celui de 40 données.

Les lois *a posteriori* ainsi obtenues (Tableau 6.29) sont de prime abord assez similaires à celles obtenues avec 40 données (Tableau 4.8). On observe néanmoins que les moyennes *a posteriori* des paramètres e et f caractérisant la distribution des erreurs de propagation sont nettement modifiées, de même que les trois valeurs de la matrice de covariance Σ . Au niveau de la variance *a posteriori*, on observe également, comme attendu, une augmentation plus générale liée à la diminution de la taille de l'échantillon d'apprentissage. Cette augmentation est particulièrement marquée pour les paramètres c, d, Σ_{11} qui caractérisent la distribution du coefficient de frottement μ_{coul} et dont les écarts types *a posteriori* sont quasiment doublés.

Tableau 6.29 : Statistiques descriptives des lois *a posteriori*, modèle bruité, 26 données

	Moy.	SD	q _{2.5}	q _{97.5}
a_1	1.56	0.38	0.99	1.38
a_2	2.82	0.72	1.6	4.39
c	0.748	0.018	0.713	0.782
d	-0.184	0.059	-0.305	-0.066
e	-0.6	15.5	-30.8	29.6
f	8.6	36.4	-63	80.1
Σ_{11}	0.0022	0.0007	0.0012	0.0039
Σ_{12}	0.08	0.44	-0.78	0.98
Σ_{22}	1850	580	1042	3266

Concernant la moyenne *a posteriori* des distances d'arrêt, les changements marquants interviennent surtout au niveau des quantiles d'ordre élevé. La différence est spectaculaire pour le quantile d'ordre 99 puisqu'il était estimé à 2261 m après apprentissage sur les 40 données et qu'il est maintenant estimé à 2474 m, soit bien plus que la plus forte valeur observée (Tableau 6.30). La faible modification des estimateurs enregistrée lors du passage de 40 à 26 données a donc un effet dispersif sur la distribution des distances d'arrêt suffisant pour affecter considérablement les valeurs de projet. Plus précisément, c'est surtout l'augmentation de Σ_{11} qui se traduit en simulation par la génération de bien plus de faibles valeurs de μ_{coul} permettant d'obtenir de grandes distances d'arrêt. Il convient toutefois de noter que l'on observe alors davantage un effet de changement d'échantillon d'apprentissage que de quantité de données puisque l'on est en prédiction moyenne.

Tableau 6.30 : Distribution moyenne *a posteriori* des distances d'arrêt, modèle bruité, 26 données

	Données (m)	Distribution simulée (m)
Moy. (m)	1933.1	1871.5
SD (m)	104.8	312.2
Skew.	0.19	-0.98
q₂₅ (m)	1876.1	1701.4
q₅₀ (m)	1915.6	1954.4
q₇₅ (m)	1979.8	2070.2
q₉₀ (m)	2046.1	2173
q₉₉ (m)	2207.8	2474

Si l'on effectue les calculs des intervalles de crédibilité associés aux valeurs de projet, l'erreur d'estimation s'ajoute à l'effet de dispersion dû au changement de l'échantillon d'apprentissage. Les aléas de référence obtenus sont alors très élevés en moyenne (Tableau 6.31), avec par exemple une différence de plus de 200 m par rapport au quantile centennal précédemment obtenu avec apprentissage sur 40 données. Ces résultats peuvent difficilement être considérés comme réalistes étant donné les résultats obtenus avec les modèles de Voellmy bloc glissant et Saint-Venant sur le même couloir. De plus, ils présentent des intervalles de crédibilité tellement larges, par exemple près

de 400 m pour le quantile centennal, qu'ils sont de toute façon inutilisables pour la pratique en ingénierie.

Au final, cette étude indique que l'utilisation du modèle de Coulomb bruité avec moins d'une trentaine de données est à proscrire. Elle confirme donc la sensibilité très forte de ce modèle à l'information utilisée pour l'inférence tant en qualité qu'en quantité. Cette forte sensibilité est pour partie attribuable au caractère purement empirique du modèle ne permettant pas de compenser une information défaillante par la modélisation physique. Elle est néanmoins également attribuable au *prior* très peu informatif qui a été utilisé pour l'inférence et qui ne permet pas de compenser une faible quantité de données par de l'information experte. Le modèle de Coulomb bruité pourrait donc éventuellement malgré tout être utilisé avec peu de données à condition de choisir des *priors* bien plus informatifs.

Tableau 6.31 : Périodes de retour avec le modèle bruité et 26 données

	Moy .	SD	q_{2.5}	q_{97.5}
$x_{stop_{10}}$ (m)	2176.3	53	2088.1	2298
$x_{stop_{30}}$ (m)	2314.9	87	2181	2516.8
$x_{stop_{100}}$ (m)	2480	104.9	2291.1	2702.5
$x_{stop_{300}}$ (m)	2587.6	118.9	2392.6	2857.2

6.3.4.2. Avec le modèle de Saint-Venant

Afin de voir si un modèle plus physique pouvait se révéler moins sensible à la quantité de données, on a également divisé par deux le nombre de données utilisées pour l'application du modèle de Saint-Venant sur la base d'une séparation aléatoire de l'échantillon de départ en deux sous échantillons de taille égale.

Tableau 6.32 : Lois a posteriori avec le modèle de Saint-Venant et 13 données

	Moy.	SD	q_{2.5}	q_{97.5}
x_{start} (m)	886.1	390.7	188.4	1618.8
h_{plaque} (m)	1.43	0.38	0.79	2.26
μ_{voel}	0.479	0.12	0.244	0.713
x_{stop} (m)	1824.6	334.5	884	2164

Les résultats obtenus en termes de prédiction semblent bien plus stables qu'avec le modèle de Coulomb bruité. En effet, même si la comparaison des moyennes *a posteriori* obtenues avec 13 données (Tableau 6.32) et 26 données (Tableau 4.23) montre certaines différences du fait de la modification de l'échantillon d'apprentissage, on n'observe pas d'« explosion » des quantiles annuels comme avec le modèle de Coulomb bruité. Le Tableau 6.33 montre par exemple que la distance d'arrêt centennale n'est augmentée que de 36 m par la division par deux de la taille de l'échantillon d'apprentissage. La Figure 6.28 montre quant à elle que la forme des deux distributions des distances d'arrêt est assez similaire.

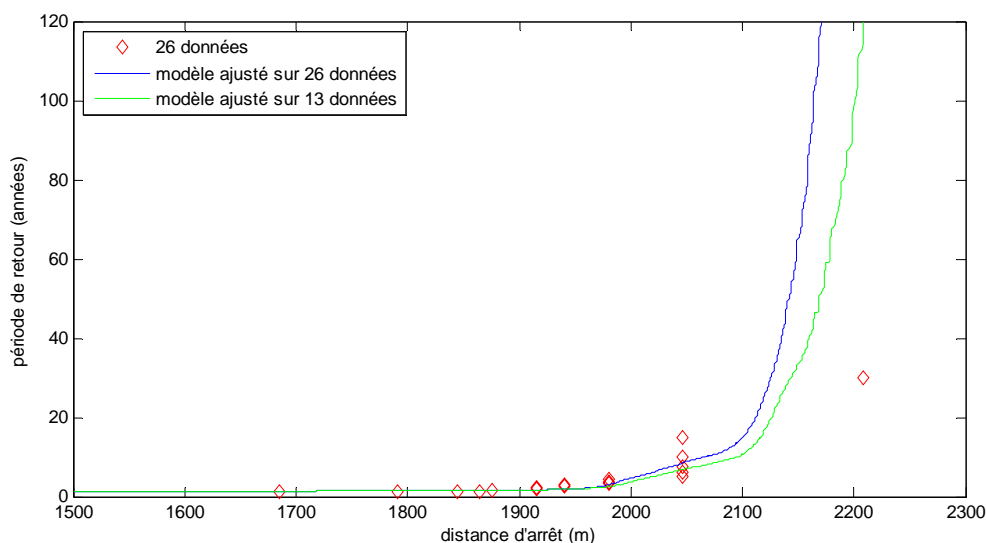
Tableau 6.33 : Période de retour avec le modèle de Saint-Venant et 13 données

Période de retour T (année)	Abcisse d'arrêt x_{stop_T} (m),	Abcisse d'arrêt x_{stop_T} (m),
	26 données	13 données
2	1953.7	1957.4
5	2004	2019
10	2064	2099
30	2125.2	2144
100	2164	2200.8
300	2203.9	2233.6

Bien sûr, si l'on calculait les intervalles de crédibilité sur les valeurs de projet avec 13 et 26 données, ce qui n'a pas été fait pour des raisons de temps calcul, on constaterait vraisemblablement qu'ils sont

significativement plus larges avec 13 données qu’avec 26 données. Néanmoins, il semble indéniable qu’une baisse importante du nombre de données affecte moins fortement le fonctionnement du modèle de Saint-Venant que celui de Coulomb bruité. L’utilisation d’une description fluide et d’une loi de frottement plus proche de la réalité physique semble donc diminuer l’importance de l’échantillon d’apprentissage, d’autant que l’utilisation d’un *prior* assez informatif est, comme on l’a vu, inévitable. Ce résultat montre bien que l’on est passé, en abandonnant le modèle de Coulomb bruité pour le modèle de Saint-Venant, d’un modèle purement empirique d’entrées/sorties qui fonctionne bien lorsque l’échantillon de calage est important à un modèle de type semi empirique où la connaissance physique injectée dans le modèle est capable de compenser la méconnaissance locale. Pour fixer les idées, on peut penser que le modèle de Saint-Venant permet l’obtention de résultats utilisables dès qu’une petite dizaine d’événements est disponible.

Figure 6.28 : Sensibilité de la distribution des distances d’arrêt à la quantité de données



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle de Saint Venant.

6.3.5. Modèle de propagation

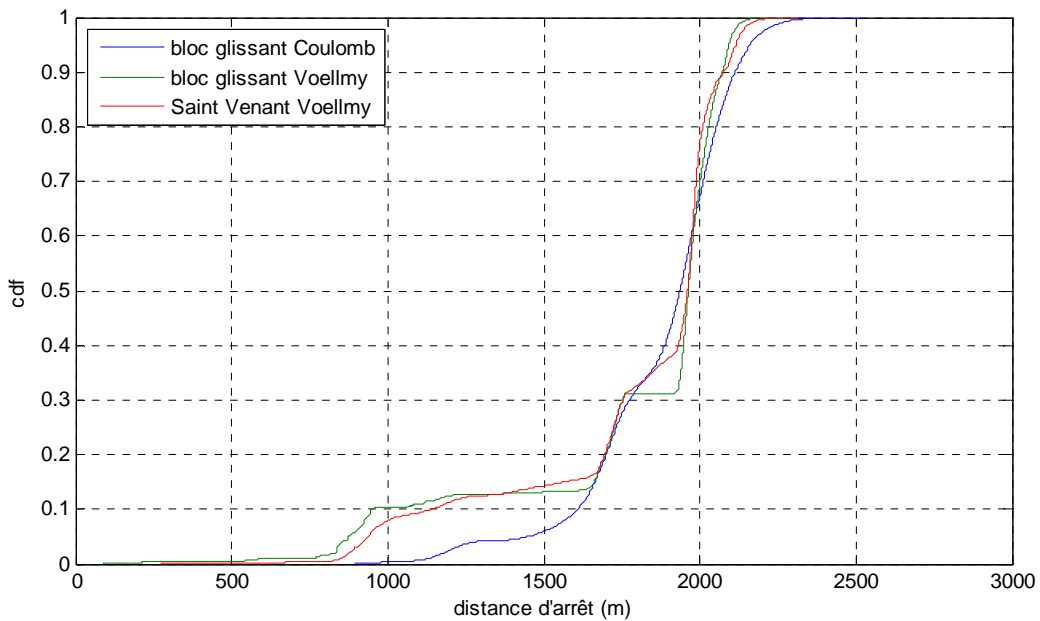
6.3.5.1. Résultats « bruts » des trois modèles sur un même couloir

Puisque le modèle de Coulomb bruité, le modèle de Voellmy bloc glissant et le modèle de Saint-Venant ont été appliqués sur un même site, c'est-à-dire le couloir EPA numéro 13 de Bessans, il est tentant de comparer les résultats obtenus. La Figure 6.29 présente les trois fonctions de répartition des distances d’arrêt présentées au chapitre 6.1. A première vue, les résultats sont globalement cohérents avec trois fonctions de répartition présentant des allures similaires. Les modèles de Voellmy bloc glissant et Saint-Venant sont très proches quelle que soit la distance d’arrêt, sauf au niveau des ruptures de pente où le modèle de Saint-Venant présente une forme plus régulière. On y reviendra plus tard. Quant au modèle de Coulomb bruité, l’écart avec les deux autres modèles est surtout marquant pour des distances d’arrêt très faibles et, dans une moindre mesure, pour des distances d’arrêt très fortes.

Le fait que le modèle de Coulomb bruité prédise beaucoup moins de distances d’arrêt faibles et très faibles que les autres modèles est lié à l’utilisation d’une vitesse initiale non nulle assez importante qui empêche les avalanches de s’arrêter très tôt même lorsqu’elles sont caractérisées par un frottement élevé. La différence entre les différents modèles pour les distances d’arrêt très faibles n’a donc pas de fondements physiques et elle n’est de toute façon pas très importante d’un point de vue opérationnel.

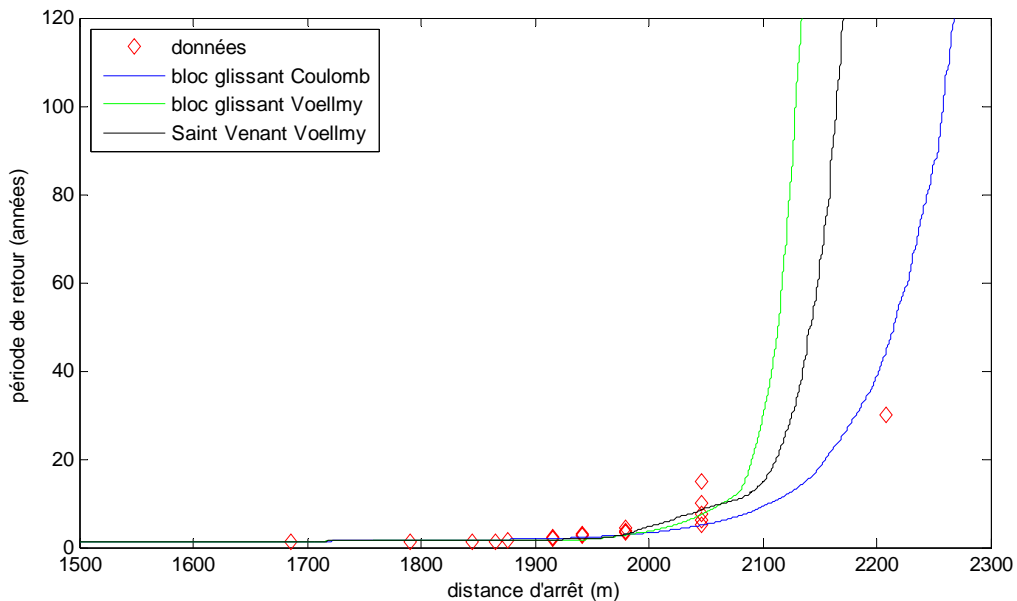
La différence entre les trois modèles pour des périodes de retour élevées est en revanche fondamentale. La Figure 6.30 confirme le caractère pessimiste du modèle de Coulomb bruité puisqu’il s’approche de la plus forte valeur observée alors qu’elle ne figure pas dans l’échantillon d’apprentissage utilisé pour l’inférence du modèle. Les deux autres modèles donnent des résultats assez similaires que l’on pourrait qualifier de façon relative comme « optimiste » pour le modèle de Voellmy bloc glissant et « légèrement optimiste » pour le modèle de Saint-Venant.

Figure 6.29 : Fonction de répartition des distances d'arrêt avec les trois modèles d'intensité



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Figure 6.30 : Période de retour avec les trois modèles d'intensité



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les données représentées sont celles utilisées pour le calage des modèles de Voellmy bloc glissant et Saint-Venant, i.e. 26 avalanches. Pour le calage du modèle de Coulomb, 40 avalanches sont utilisées.

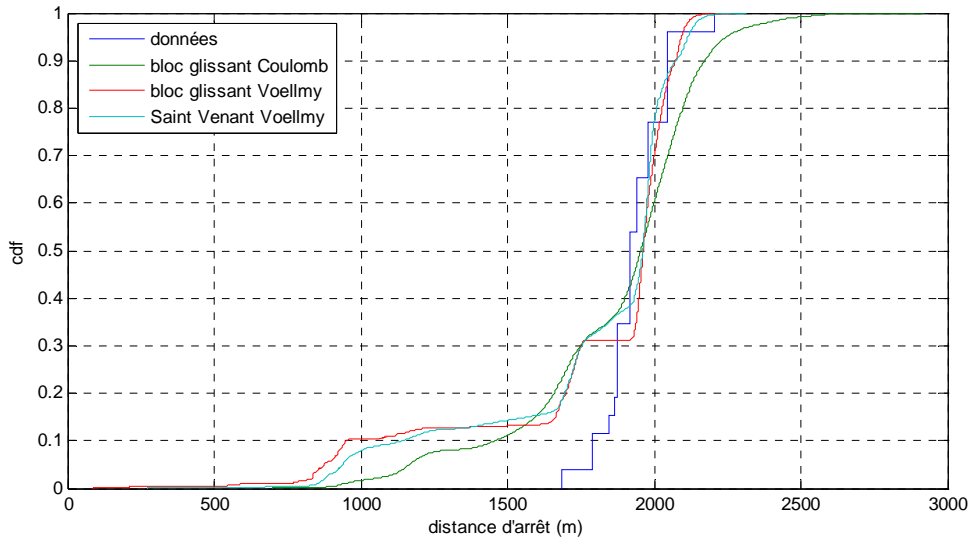
6.3.5.2. Fonctionnement à quantité d'information équivalente

La comparaison entre les différents modèles est encore plus pertinente si les données utilisées pour l'inférence sont identiques. On compare donc à présent les trois modèles sur les mêmes 26 événements utilisés au chapitre 6.1 avec les modèles de Voellmy bloc glissant et Saint-Venant et au chapitre 6.3.4.1 avec le modèle de Coulomb bruité.

La Figure 6.31 montre que les trois modèles prévoient bien davantage de distances d'arrêt faibles qu'il n'en existe dans les données. Cela s'explique par le fait que les très petites avalanches n'étaient souvent pas enregistrées dans l'EPA avant le toilettage et qu'elles ne le sont presque plus du tout depuis l'instauration d'un seuil d'observation. Du fait de ce décalage vers le bas de la distribution empirique, c'est le modèle de Coulomb bruité qui est le plus proche des données pour des distances

d'arrêt faibles. Mais c'est, comme on l'a vu, bien plus dû à l'introduction peu physique d'une vitesse initiale non nulle qu'à une représentation réaliste du fonctionnement du couloir. Il ne s'agit donc pas d'un « bon point » pour ce modèle. Pour des distances d'arrêt élevées, l'accord semble par contre très bon entre la distribution empirique des 26 données et les modèles de Voellmy et Saint-Venant, tandis que le modèle de Coulomb bruité, conformément aux conclusions du paragraphe 6.3.4, semble trop pessimiste.

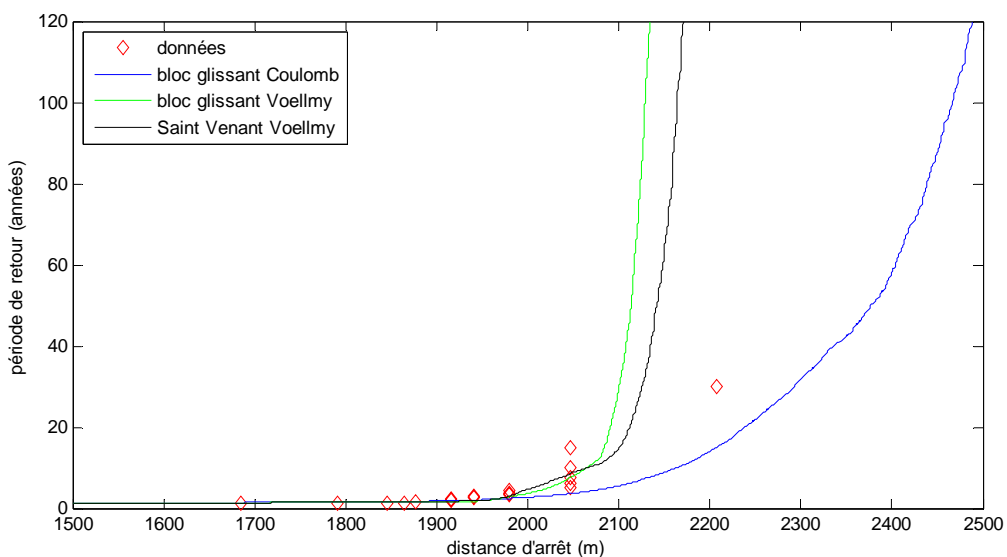
Figure 6.31 : Fonctions de répartition à quantité d'information équivalente pour l'inférence



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les données représentées sont les 26 avalanches utilisées pour le calage des trois modèles.

Si l'on regarde à nouveau spécifiquement le comportement des trois modèles pour des périodes de retour élevées (Figure 6.32), rien ne change par rapport à la Figure 6.30 pour les modèles de Saint-Venant et Voellmy bloc glissant. En revanche, on constate que le modèle de Coulomb bruité s'éloigne encore davantage des deux autres modèles. La raison en est, comme dit au chapitre 6.3.4, la diminution du nombre de données et surtout l'incorporation dans l'échantillon d'apprentissage de l'avalanche de distance d'arrêt supérieure à 2200 m qui tire fortement les valeurs de projet proposées vers le haut.

Figure 6.32 : Période de retour à quantité d'information équivalente pour l'inférence



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les données représentées sont les 26 avalanches utilisées pour le calage des trois modèles.

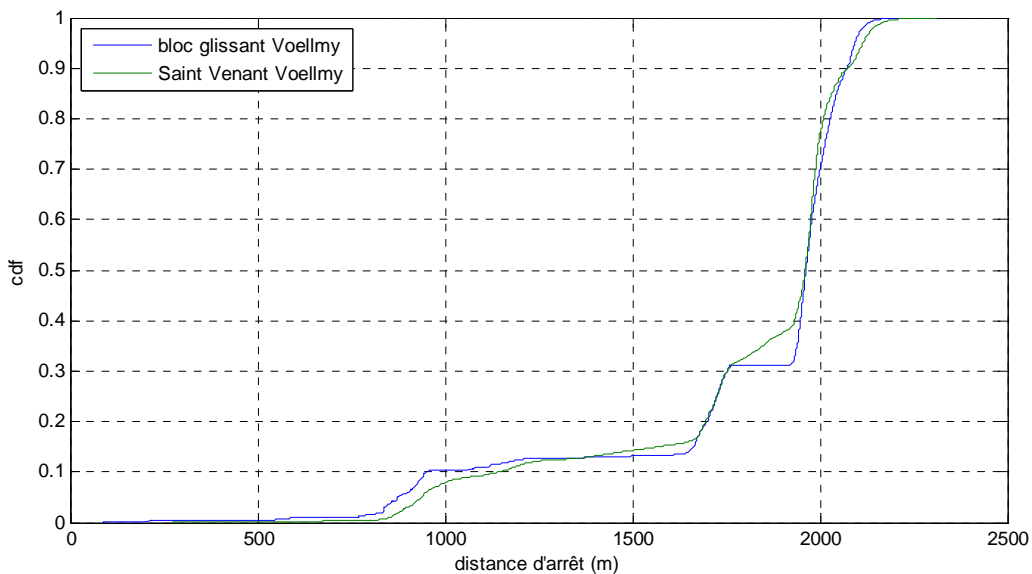
6.3.5.3. Comparaison fine des deux modèles avec la même loi de frottement

Les modèles de Voellmy bloc glissant et Saint-Venant sont particulièrement étroitement liés puisque le modèle stochastique est quasi équivalent, la seule différence étant la signification différente de la variable h , hauteur de bloc constante dans un cas contre hauteur de fracture dans le second. Les mêmes lois *a priori* ont, de plus, pu être employées, ce qui signifie qu'une information extérieure aux données très similaire a été injectée dans les deux modèles.

En termes de distribution des distances d'arrêt, les résultats obtenus avec les deux modèles sont, on l'a vu, assez concordants, même si le modèle de Saint-Venant est un peu plus pessimiste pour de très grandes périodes de retour (Figure 6.30 ou 6.32).

Dans le détail, la différence la plus marquante se situe au niveau des zones de ressaut, notamment celle située en début de zone d'arrêt pour le couloir étudié. Avec le modèle de bloc glissant, une franche cassure dans la fonction de répartition est observée, indiquant qu'avec ce modèle aucune avalanche ne peut s'arrêter sur les pentes fortes. Avec le modèle de Saint-Venant, la fonction de répartition obtenue est beaucoup plus régulière. Cela indique que toutes les abscisses d'arrêt peuvent être obtenues en fonction des valeurs de coefficient de frottement et de la capacité du modèle à adapter hauteur et répartition de la masse aux variations de la pente (Figure 6.33). Cette différence entre les deux modèles était déjà apparue au chapitre 4 où il a été montré que le modèle de Saint-Venant était bien mieux capable que le modèle de bloc glissant de reproduire sans biais les événements observés s'étant arrêtés au niveau de la zone de ressaut. Il semble donc bien que le modèle de Saint-Venant offre une vision plus réaliste que le modèle de bloc glissant de la répartition des distances d'arrêt, tout au moins dès lors que des irrégularités topographiques existent.

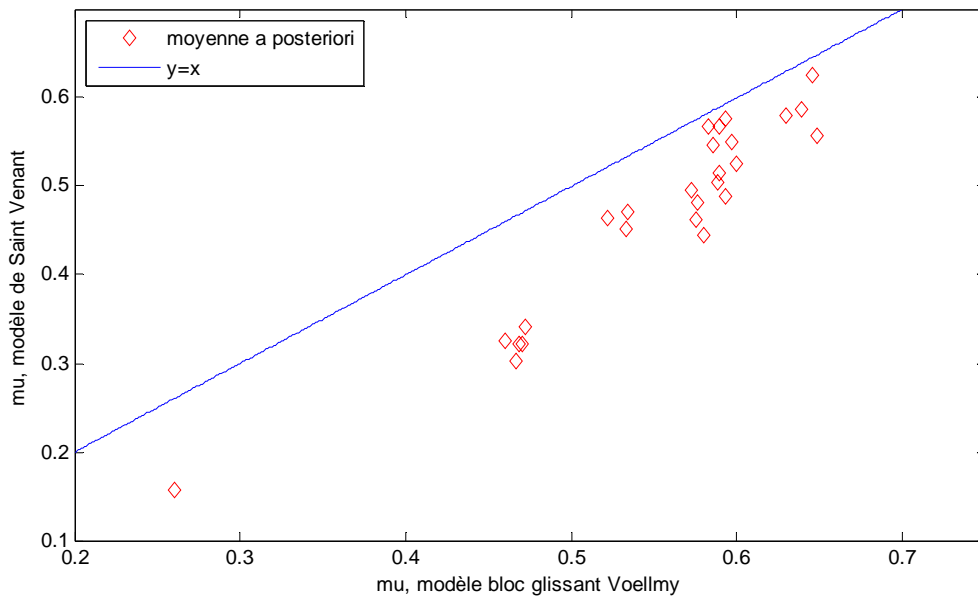
Figure 6.33 : Répartition des distances d'arrêt, modèles bloc glissant Voellmy et Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

Concernant les coefficients de frottement calibrés sur les événements observés, la Figure 6.34 rappelle qu'avec le modèle de bloc glissant, des valeurs significativement plus élevées du coefficient latent μ_{voel} sont obtenues par rapport à celles obtenues avec le modèle de Saint-Venant. Cet écart semble d'ailleurs d'autant plus fort que μ_{voel} est faible. Cette différence systématique est à relier à la différence également obtenue en termes d'estimation de ξ_{voel} dont la moyenne *a posteriori* est plus élevée avec le modèle de bloc glissant qu'avec le modèle fluide. Les valeurs calibrées de μ_{voel} s'ajustent donc pour conserver par compensation entre les deux coefficients de la loi de frottement les mêmes distances d'arrêt. Toutefois, les valeurs obtenues à l'aide du modèle de Saint-Venant sont plus proches des valeurs considérées comme « vraies », comme par exemple la valeur de 0.16 pour l'avalanche de distance d'arrêt la plus élevée. Cela plaide encore une fois pour un réalisme physique plus important du modèle de Saint-Venant.

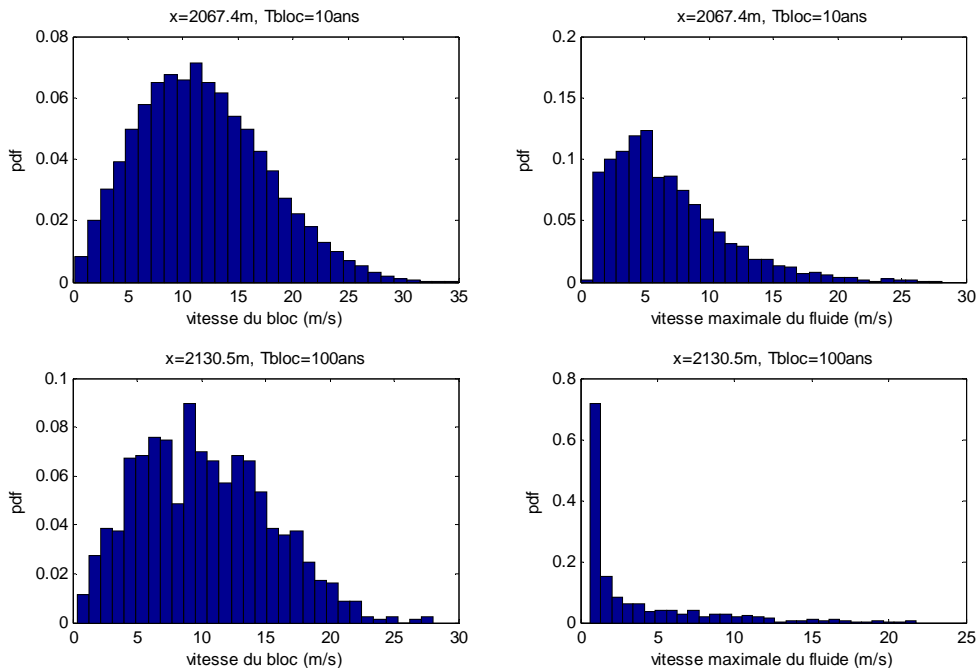
Figure 6.34 : Coefficients de frottement avec les modèles bloc glissant Voellmy et Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les moyennes a posteriori correspondant aux différents événements sont comparées.

Enfin les deux modèles peuvent être comparés en termes de distribution des vitesses aux mêmes abscisses. La Figure 6.35 présente ainsi les distributions des vitesses avec le modèle de bloc glissant et des vitesses maximales avec le modèle de Saint-Venant aux abscisses correspondant à des périodes de retour de 10 et 100 ans calculées avec le modèle de bloc glissant. Pour le modèle de bloc glissant, elles sont identiques à celles de la Figure 6.6. Pour le modèle de Saint-Venant, elles sont proches de celles présentées en Figure 6.11 sans toutefois être identiques car les abscisses auxquelles les distributions sont calculées sont légèrement différentes. En effet, le modèle de Saint-Venant étant un peu plus pessimiste que celui de bloc glissant pour des périodes de retour élevées, on a pour une période de retour donnée un quantile annuel un peu moins élevé avec le modèle de bloc glissant.

Figure 6.35 : Vitesses maximales avec les modèles de bloc glissant Voellmy et Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France.

A abscisse égale, on constate que les distributions conditionnelles à l'atteinte obtenues avec les deux modèles présentent des formes assez différentes : plutôt de « type Gamma » avec le modèle de bloc glissant et beaucoup plus asymétriques avec le modèle fluide. De plus, la vitesse moyenne est beaucoup plus forte avec le modèle de bloc glissant. C'est d'autant plus surprenant qu'il s'agit avec le modèle de Saint-Venant de vitesses maximales (Tableau 6.34). Quant à expliquer cette différence de vitesses, les valeurs des coefficients de frottement estimés constituent une piste intéressante. En effet, la valeur de ξ_{voel} bien plus élevée calibrée avec le modèle de bloc glissant favorise la simulation de vitesses élevées.

Tableau 6.34 : Comparaison des vitesses entre les modèles de Voellmy et Saint-Venant

	Moy.	SD	Skew	q_{2.5}	q_{97.5}
$v(m/s), x = 2067.4 m$, bloc glissant	11.8	5.7	0.4	2.2	24.2
$v(m/s), x = 2067.4 m$, fluide	6.8	4.5	0.6	1.4	18.4
$v(m/s), x = 2130.5 m$, bloc glissant	10.2	5.1	0.3	1.9	20.5
$v(m/s), x = 2130.5 m$, fluide	3.6	4.0	0.7	0.8	16.1

6.4. Calcul de pression

6.4.1. Pression en écoulement libre

Jusqu'ici, aucun calcul de pression n'a été effectué. Néanmoins, du fait du rôle fondamental de cette variable pour le zonage et le dimensionnement d'ouvrages de protection, la question se doit d'être abordée. Conditionnellement à une valeur fixée de la masse volumique ρ_N de la neige en écoulement que l'on prend égale à 300 kg/m^3 , l'obtention de la pression théorique en écoulement libre est facile à obtenir à partir de la distribution des vitesses. Le calcul est possible avec les modèles de Voellmy bloc glissant et de Saint-Venant. Les pressions obtenues correspondent à deux fois l'énergie cinétique par unité de volume, c'est-à-dire à un Cx constant égal à 2 dans l'eq. 3.11.

6.4.1.1. Avec le modèle de Voellmy bloc glissant

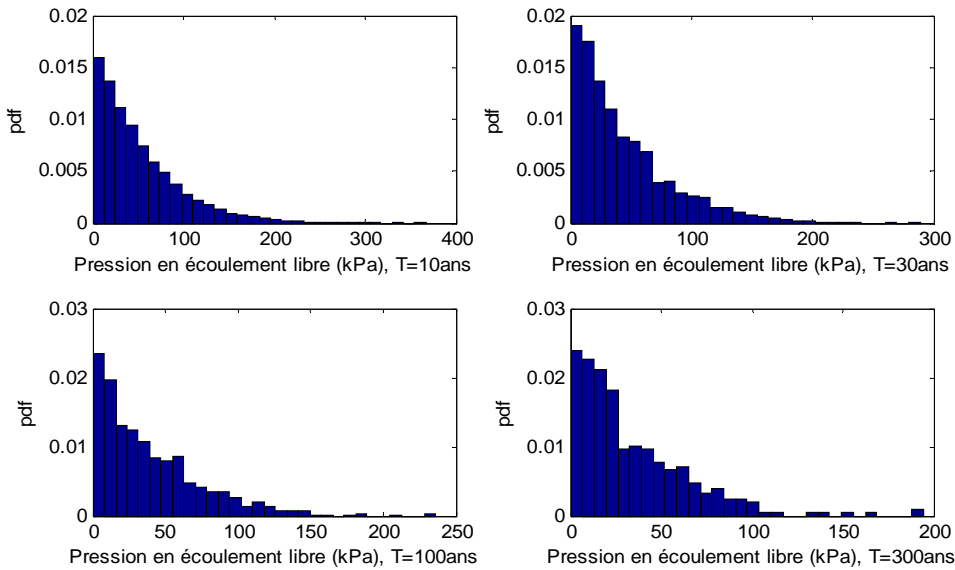
Du fait des fortes vitesses simulées avec le modèle de bloc glissant, les valeurs de pression obtenues sont assez fortes, avec par exemple une pression moyenne à l'abscisse centennale de 39.2 kPa et une borne supérieure de l'intervalle de crédibilité à 95% de 126.7 kPa.. Par rapport à la valeur de 30 kPa souvent retenue comme une valeur limite en termes d'urbanisme pour la construction dans les zones de montagne, de telles pressions sont considérables et mettent en doute la validité des distributions de vitesse obtenues avec ce modèle. De plus, la décroissance de la pression maximale et moyenne avec la période de retour est lente, ce qui ne fait que refléter la lente décroissance des vitesses des avalanches atteignant des abscisses de plus en plus importantes (Tableau 6.35).

Tableau 6.35 : Pression en écoulement libre et période de retour avec le modèle de bloc glissant

	Moy.	SD	q_{2.5}	q_{97.5}
$Pr_{T_{10}}$ (kPa)	51.4	46.4	1.5	175.0
$Pr_{T_{30}}$ (kPa)	43.8	40.4	1.4	149.3
$Pr_{T_{100}}$ (kPa)	39.2	35.7	1.1	126.7
$Pr_{T_{300}}$ (kPa)	33.7	31.2	0.8	103.8

Globalement, les distributions obtenues sont de type exponentiel quelle que soit la période de retour avec un mode très proche de zéro et un écart type d'un ordre de grandeur comparable à celui de la moyenne. La régularité des distributions décroît bien entendu avec la période de retour du fait de la diminution du nombre de simulations considérées (Figure 6.36).

Figure 6.36 : Pression théorique en écoulement libre avec le modèle de Voellmy bloc glissant

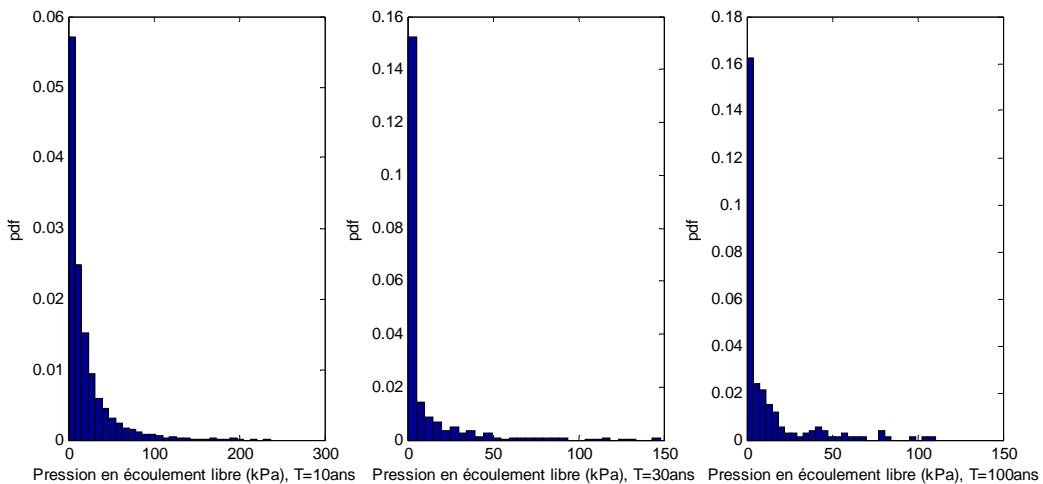


Couloir EPA 13 de la commune de Bessans, Savoie, France. Les pressions sont calculées avec $\rho_N = 300 \text{ kg/m}^3$ et $Cx = 2$.

6.4.1.2. Avec le modèle de Saint-Venant

Comme pour les variables du modèle, on limite les calculs à des périodes de retour inférieures ou égales à 100 ans de façon à limiter l'erreur de Monte Carlo. Les pressions moyennes sont nettement plus raisonnables qu'avec le modèle de bloc glissant puisqu'une valeur de 11.4 kPa correspond à l'abscisse centennale (Tableau 6.36), soit trois fois moins que la valeur précédemment obtenue à la même abscisse. Néanmoins, les valeurs moyennes n'ont pas beaucoup de sens avec ce modèle car les distributions sont très fortement asymétriques (Figure 6.37), avec des valeurs extrêmes presque aussi fortes qu'avec le modèle de bloc glissant. Ainsi le modèle de Saint-Venant semble prédire que les pressions exercées par la grande majorité des avalanches atteignant des abscisses élevées sont faibles, voire très faibles, mais que, de temps en temps, de très fortes valeurs peuvent être observées. Enfin, signalons que les valeurs prédites pour les avalanches trentennales sont plus faibles que pour les avalanches centennales. Cela est dû à l'irrégularité du profil topographique qui fait, comme cela avait déjà été relevé au moment de l'analyse des vitesses et des nombres de Froude, diminuer la dynamique de l'écoulement au niveau de l'abscisse trentennale sur ce cas d'étude.

Figure 6.37 : Pression maximale théorique en écoulement libre avec le modèle de Saint-Venant



Couloir EPA 13 de la commune de Bessans, Savoie, France. Les pressions sont calculées avec $\rho_N = 300 \text{ kg/m}^3$ et $Cx = 2$.

Tableau 6.36 : Pression en écoulement libre et période de retour avec le modèle de Saint-Venant

	Moy.	SD	q _{2.5}	q _{97.5}
Pr _{T₁₀} (kPa)	20.0	27.6	0.6	101.2
Pr _{T₃₀} (kPa)	8.7	19.7	0.2	74.2
Pr _{T₁₀₀} (kPa)	11.4	20.4	0.2	78.1

6.4.2. Distribution du C_x

La connaissance des distributions des hauteurs, vitesses et coefficients de frottement permet de calculer la distribution des C_x de façon à évaluer des pressions d'impact plus réalistes tenant compte de la rhéologie de l'écoulement. Du fait de la dépendance du C_x avec les caractéristiques de l'écoulement, les calculs ne sont effectués qu'avec le modèle de Saint-Venant. Cinq différentes formulations sont testées.

Les deux premières sont des lois expérimentales issues des travaux de Sovilla (2007) sur le site de la Sionne. La loi puissance de l'équation 3.12 doit être divisée par la masse volumique de la neige en écoulement pour se ramener au C_x (eq 6.5). On teste deux lois expérimentales obtenues pour deux événements du site de la Sionne. Elles correspondent à des valeurs assez différentes du paramètre ν_2 , c'est-à-dire à des degrés de dépendance différents du C_x avec la dynamique de l'écoulement. Pour la première, $\nu_1=5770$ et $\nu_2=1.9$, soit une dépendance de type « turbulente » du C_x avec le carré de la vitesse. Pour la seconde, $\nu_1=2508$ et $\nu_2=0.8$, soit une dépendance de type « visqueuse » du C_x avec la vitesse.

Les trois autres formulations testées sont trois « versions » de la formule proposée par Naaïm (2007). Dans l'équation qu'il a obtenue par analogie avec un fluide modèle ayant un comportement visqueux, la rhéologie du fluide intervient dans le calcul de pression au travers de la loi de frottement présente au dénominateur. En outre, la forme de l'obstacle considéré intervient explicitement au travers de son diamètre d_o (eq. 6.6). On teste cette formule avec un obstacle type de faible diamètre $d_o=0.25$ m, par exemple un pylône, avec un obstacle type plus large vérifiant $d_o=5$ m, par exemple une maison ou un tas freineur de protection, ainsi qu'avec un rapport h/d_o fixé et égal à 1. Dans ce dernier cas, la hauteur de l'écoulement est supposée être toujours du même ordre de grandeur que le diamètre de l'obstacle.

$$C_x = \frac{1}{2 \times \rho_N} \times \nu_1 \times Fr^{-\nu_2} \quad (\text{eq. 6.5})$$

$$C_x = 0.6 \times \left(1 + 9 \left(\frac{\frac{d_o \times Fr^2}{h}}{\left(\mu_{voel} + \frac{g}{\xi_{voel}} Fr^2 \right) \cos(\phi)} \right)^{-1/2} \right) \quad (\text{eq. 6.6})$$

Les distributions des C_x obtenues avec les 5 formulations à une abscisse décennale sont présentées en Figure 6.38. Elles présentent toutes les 5 des formes en cloche avec un mode bien marqué. Les deux distributions obtenues avec la formule de Naaïm en faisant varier le diamètre d'obstacle sont quasi symétriques mais divergent par les valeurs proposées qui sont près de trois fois supérieures pour l'obstacle de faible diamètre. Les trois autres distributions, c'est-à-dire les deux lois expérimentales et la formule de Naaïm avec h/d_o fixé, présentent quant à elles une légère asymétrie à droite indiquant que des valeurs nettement supérieures à la moyenne peuvent être obtenues.

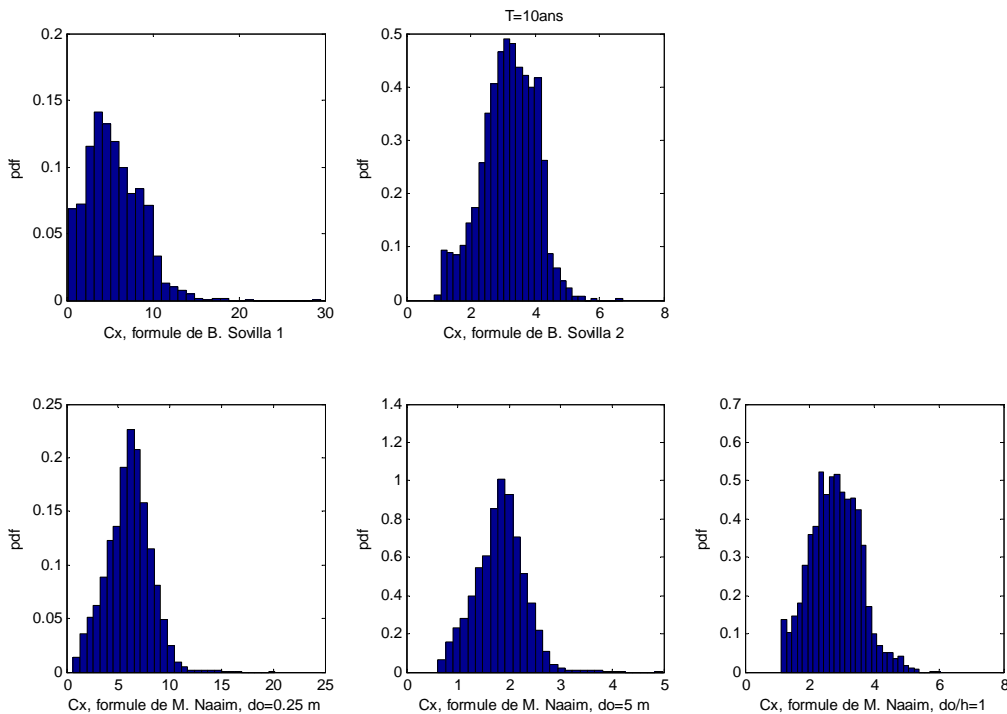
Seule la formule de Naaïm pour un diamètre de 5 m permet de retrouver un C_x de l'ordre de 2 correspondant à la pression en écoulement libre. Toutes les autres donnent des C_x en moyenne plus élevés et suggèrent donc que la pression en écoulement libre doit être majorée pour obtenir des pressions d'impact réalistes. C'est particulièrement vrai pour la formule de Naaïm avec un obstacle de faible diamètre qui propose un C_x à l'abscisse décennal de l'ordre de 6 et avec la formule de Sovilla avec la dépendance « turbulente » qui propose un C_x à l'abscisse décennale de l'ordre de 5.5. La

formule de Sovilla avec une dépendance visqueuse semble donner des résultats proches de ceux obtenus avec la formule de Naaim avec h/d_o fixé, puisque, dans les deux cas, les Cx à l'abscisse décennale sont de l'ordre de 3 (Tableau 6.37).

Tableau 6.37 : Cx à une abscisse décennale

	Moy.	SD	q _{2.5}	q _{97.5}
$Cx_{T_{10}}$, formule de Sovilla avec $\nu_2 = -1.9$	5.5	3.1	0.6	12.0
$Cx_{T_{10}}$, formule de Sovilla avec $\nu_2 = -0.8$	3.2	0.8	1.3	4.6
$Cx_{T_{10}}$, formule de Naaim avec $d_o = 0.25$ m	6.0	2.1	1.8	9.9
$Cx_{T_{10}}$, formule de Naaim avec $d_o = 5$ m	1.8	0.5	0.9	2.7
$Cx_{T_{10}}$, formule de Naaim avec $h/d_o = 1$	2.8	0.8	1.3	4.5

Figure 6.38 : Cx à une abscisse décennale



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle de Saint Venant.

Pour des périodes de retour plus élevées, des Cx encore plus importants sont obtenus. C'est particulièrement vrai à l'abscisse trentennale (Tableau 6.38) où les valeurs calculées sont supérieures à celles calculées à l'abscisse centennale (Tableau 6.39). Cela s'explique par les nombres de Froude plus faibles observés à l'abscisse trentennale du fait de la diminution locale de la dynamique de l'écoulement.

Pour les deux abscisses, on retrouve que la formule de Naaim permet d'obtenir des Cx de l'ordre de 2 pour des diamètres d'obstacle importants. De même, on retrouve que pour des diamètres d'obstacle plus faibles, des valeurs de Cx nettement plus importantes sont obtenues. De plus, les valeurs de Cx les plus fortes sont obtenues avec la formule de Sovilla avec la dépendance turbulente. Ces dernières, de l'ordre de 20, semblent trop fortes pour être réalistes. Cela suggère que le coefficient $\nu_2=1.9$, s'il pouvait refléter les propriétés particulières d'une avalanche, est trop élevé pour représenter un comportement moyen. Au contraire, la loi expérimentale de type visqueuse fournit des Cx d'ordre de grandeur assez comparables à ceux de la loi de Naaim, en particulier lorsque h/d_o est fixé. C'est naturel étant donné que, dans les deux cas, le Cx varie quasiment en fonction inverse de la vitesse.

Tableau 6.38 : Cx à une abscisse trentennale

	Moy.	SD	$q_{2.5}$	$q_{97.5}$
$Cx_{T_{30}}$, formule de Sovilla avec $v_2 = -1.9$	24.5	22.0	1.3	84.8
$Cx_{T_{30}}$, formule de Sovilla avec $v_2 = -0.8$	5.6	2.4	1.8	10.5
$Cx_{T_{30}}$, formule de Naaim avec $d_o = 0.25$ m	8.5	4.0	1.2	17.3
$Cx_{T_{30}}$, formule de Naaim avec $d_o = 5$ m	2.4	0.9	0.7	4.3
$Cx_{T_{30}}$, formule de Naaim avec $h/d_o = 1$	4.6	2.0	1.6	8.3

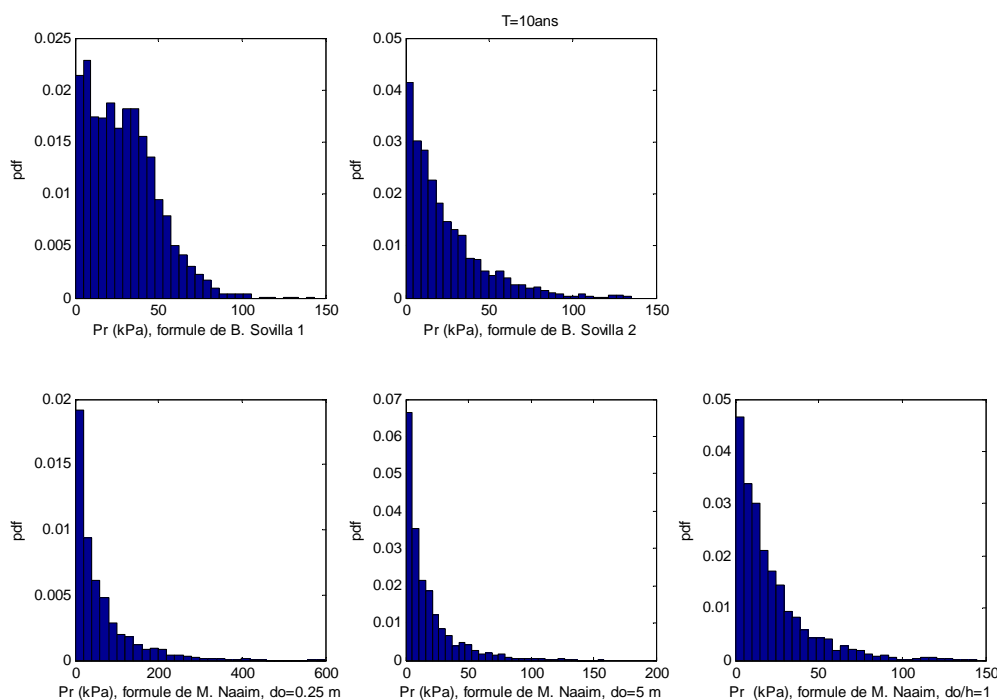
Tableau 6.39 : Cx à une abscisse centennale

	Moy.	SD	$q_{2.5}$	$q_{97.5}$
$Cx_{T_{100}}$, formule de Sovilla avec $v_2 = -1.9$	17.9	20.2	1.1	76.9
$Cx_{T_{100}}$, formule de Sovilla avec $v_2 = -0.8$	4.8	2.2	1.7	10.0
$Cx_{T_{100}}$, formule de Naaim avec $d_o = 0.25$ m	6.1	2.6	1.8	11.4
$Cx_{T_{100}}$, formule de Naaim avec $d_o = 5$ m	1.8	0.6	0.9	3.0
$Cx_{T_{100}}$, formule de Naaim avec $h/d_o = 1$	3.7	1.6	1.4	7.2

6.4.3. Calcul des pressions d'impact

Les distributions des vitesses et des Cx peuvent être facilement combinées pour obtenir celles des pressions d'impact. A une abscisse décennale, toutes les distributions données par les différentes formules à l'exception de celle obtenue à l'aide de la formule de Sovilla avec une dépendance turbulente présentent des formes de type exponentiel comme pour les pressions en écoulement libre : mode proche de zéro et écart type sensiblement égal à la moyenne (Figure 6.39).

Figure 6.39 : Pression décennale maximale tenant compte de la rhéologie



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle de Saint Venant.

Les valeurs de pression obtenues sont, à cette abscisse, moins différentes des valeurs en écoulement libre que l'on pouvait s'y attendre. C'est dû aux valeurs encore assez élevées des vitesses des

avalanches atteignant l'abscisse décennale qui empêchent la rhéologie d'influencer fortement la pression. Ainsi, seule la formule de Naaim pour des obstacles de faible diamètre donne des valeurs de pression d'impact très nettement supérieures à celles obtenues lorsque l'on fixe Cx à 2 (Tableau 6.40). Cette pression très élevée traduit bien que pour un très petit diamètre, la force s'exerce sur une toute petite surface, de sorte que l'obstacle est soumis à un très rude traitement.

Tableau 6.40 : Pression d'impact à une abscisse décennale

	Moy.	SD	q_{2.5}	q_{97.5}
$Pr_{T_{10}}$ (kPa), formule de Sovilla avec $\nu_2 = -1.9$	29.2	20.4	2.6	75.8
$Pr_{T_{10}}$ (kPa), formule de Sovilla avec $\nu_2 = -0.8$	23.8	23.1	1.3	85.0
$Pr_{T_{10}}$ (kPa), formule de Naaim avec $d_o = 0.25$ m	58.2	70.1	1.0	253.4
$Pr_{T_{10}}$ (kPa), formule de Naaim avec $d_o = 5$ m	17.7	21.5	0.4	78.1
$Pr_{T_{10}}$ (kPa), formule de Naaim avec $h/d_o = 1$	21.4	22.2	1.1	81.0

Pour des périodes de retour plus importantes, la forme très asymétrique des distributions de pression est conservée. Elle est même accentuée avec, quelle que soit la formule de calcul retenue, des pressions maximales très fortes susceptibles d'être générées (Figure 6.40). En moyenne, les pressions d'impact diminuent avec la période de retour, l'effet de l'augmentation du Cx étant compensé par une diminution de la vitesse. C'est particulièrement vrai pour l'abscisse trentennale où des Cx très importants avaient été calculés mais où les pressions d'impact moyennes données par les différents formules ne sont pas très élevées, en tout cas inférieures à la « norme » de 30 kPa (Tableau 6.41). Néanmoins, que ce soit à une abscisse trentennale ou centennale, les pressions obtenues soit avec les formules empiriques de Sovilla soit avec la formule de Naaim et $d_o = 0.25$ sont assez significativement supérieures à celles qui avaient été obtenues avec un Cx constant (Tableau 6.42). Ainsi, prendre en compte la rhéologie pour le dimensionnement d'enjeux pouvant être atteints par des avalanches exceptionnelles uniquement semble plus nécessaire qu'à l'abscisse décennale. Cela s'explique simplement par le fait que ces enjeux sont atteints par des avalanches globalement « en bout de course », de sorte que l'écoulement n'est plus très dynamique et que l'augmentation de la pression due à l'augmentation du Cx se perçoit alors sensiblement.

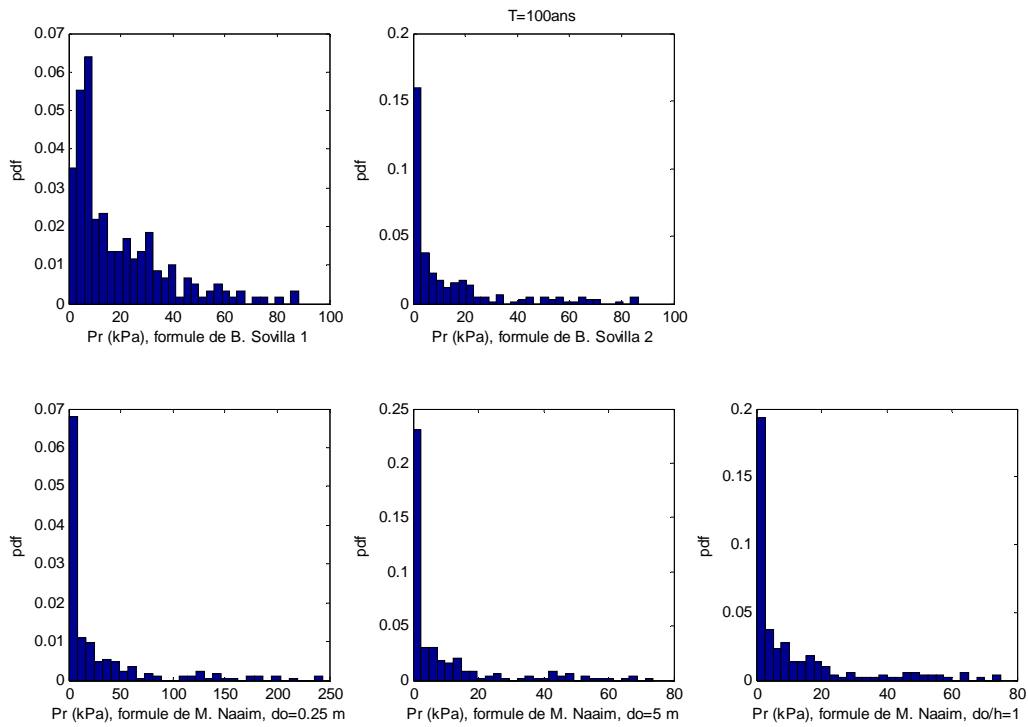
Tableau 6.41 : Pression d'impact à une abscisse trentennale

	Moy.	SD	q_{2.5}	q_{97.5}
$Pr_{T_{30}}$ (kPa), formule de Sovilla avec $\nu_2 = -1.9$	20.3	19.2	0.6	71.1
$Pr_{T_{30}}$ (kPa), formule de Sovilla avec $\nu_2 = -0.8$	12.1	19.5	0.5	76.6
$Pr_{T_{30}}$ (kPa), formule de Naaim avec $d_o = 0.25$ m	28.7	54.4	0.2	196.0
$Pr_{T_{30}}$ (kPa), formule de Naaim avec $d_o = 5$ m	8.5	16.5	0.1	64.8
$Pr_{T_{30}}$ (kPa), formule de Naaim avec $h/d_o = 1$	10.3	17.2	0.4	69.2

Tableau 6.42 : Pression d'impact à une abscisse centennale

	Moy.	SD	q_{2.5}	q_{97.5}
$Pr_{T_{100}}$ (kPa), formule de Sovilla avec $\nu_2 = -1.9$	19.2	18.3	1.2	68.2
$Pr_{T_{100}}$ (kPa), formule de Sovilla avec $\nu_2 = -0.8$	14.1	19.9	0.5	71.1
$Pr_{T_{100}}$ (kPa), formule de Naaim avec $d_o = 0.25$ m	30.1	50.6	0.2	188.4
$Pr_{T_{100}}$ (kPa), formule de Naaim avec $d_o = 5$ m	9.4	15.8	0.1	58.2
$Pr_{T_{100}}$ (kPa), formule de Naaim avec $h/d_o = 1$	11.6	17.0	0.4	62.8

Figure 6.40 : Pression centennale maximale tenant compte de la rhéologie

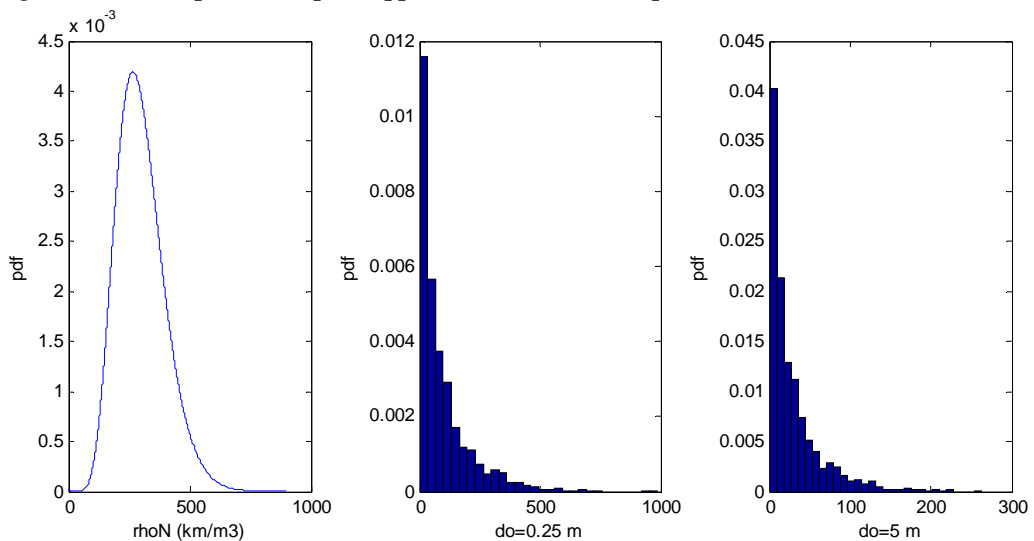


Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle de Saint Venant.

6.4.4. Effet sur la pression de l'incertitude sur la masse volumique de la neige

Pour finir, on a cherché à évaluer la sensibilité des valeurs proposées à la masse volumique de la neige. Pour ce faire, on a considéré que l'on ne connaissait plus sa valeur mais que sa distribution pouvait être évaluée de façon experte. Et l'on a moyenné les distributions des pressions obtenues précédemment sur cette distribution, ce qui revient à calculer la loi prédictive de la pression d'impact par rapport à la masse volumique considérée comme un paramètre. Rappelons que conformément aux équations proposées au chapitre 3, la masse volumique n'intervient pas dans la propagation dès lors que la variation de la densité au cours de l'écoulement est négligée.

Figure 6.41 : Loi prédictive par rapport à la densité de la pression maximale décennale



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle de Saint Venant.

Pour $p(\rho_N)$, une loi Gamma de moyenne 300 et d'écart type 100 a été choisie (Figure 6.41, gauche). Les lois prédictives des pressions décennales par rapport à ρ_N données par la formule de Naaim conservent une allure analogue à celles précédemment obtenues sans prendre en compte l'incertitude sur ρ_N , c'est-à-dire quasi exponentielle avec une moyenne dépendant fortement de la valeur de diamètre d'obstacle choisi. Néanmoins, on observe un décalage vers les fortes valeurs puisque, par exemple pour un diamètre d'obstacle de 5 m, la pression moyenne attendue à l'abscisse décennale était de 17.7 kPa et que l'on passe maintenant à 29.2 kPa, tandis que les valeurs maximales attendues à la même abscisse étaient de l'ordre de 80 kPa et que l'on passe maintenant à 130 kPa (Figure 6.41, droite). Il en va de même pour un diamètre de 0.25 m où l'on passe en moyenne de 58.2 à 96 kPa et en valeurs maximales d'un peu plus de 250 kPa à plus de 420 kPa (Figure 6.41, centre). Tenir compte de l'incertitude sur la masse volumique de l'écoulement conduit donc à prédire de plus fortes pressions d'impact en offrant la possibilité à de très fortes pressions d'être simulées, ces dernières correspondant bien entendu aux valeurs les plus élevées de ρ_N .

Pour conclure, un petit bilan des différentes pressions calculées à une abscisse décennale a été effectué (Tableau 6.43). Il apparaît que la pression en écoulement libre calculée avec un C_x fixé à 2 sous-estime la pression d'impact pour des diamètres d'obstacles faibles. C'est moins vrai pour des obstacles plus grands car la pression s'exerce alors sur des surfaces plus importantes. D'autant plus qu'à une abscisse décennale les écoulements sont encore assez dynamiques, de sorte que l'effet de la rhéologie sur la pression d'impact n'est pas encore très marqué. Enfin, les valeurs calculées présentent une forte sensibilité à la masse volumique de la neige et sa méconnaissance doit conduire à un dimensionnement prudent.

Tableau 6.43 : Bilan sur la pression d'impact à une abscisse décennale

	Moy.	SD	q _{2.5}	q _{97.5}
$Pr_{T_{10}}$ (kPa), C_x fixé à 2	20.0	27.6	0.6	101.2
$Pr_{T_{10}}$ (kPa), formule de Naaim avec $d_o = 0.25$ m, $\rho_N = 300$ kg/m ³	58.2	70.1	1.0	253.4
$Pr_{T_{10}}$ (kPa), formule de Naaim avec $d_o = 0.25$ m, ρ_N tiré selon $p(\rho_N)$	96.0	115.6	1.7	418.3
$Pr_{T_{10}}$ (kPa), formule de Naaim avec $d_o = 5$ m, $\rho_N = 300$ kg/m ³	17.7	21.5	0.4	78.1
$Pr_{T_{10}}$ (kPa), formule de Naaim avec $d_o = 5$ m, ρ_N tiré selon $p(\rho_N)$	29.2	35.5	0.6	128.9

6.5. Conclusion du chapitre 6

- *Trois modèles de prédétermination*

Dans ce chapitre, les trois modèles d'intensité du chapitre 4 et le modèle fréquentiel du chapitre 5 ont été assemblés pour obtenir trois modèles de prédétermination permettant le calcul des aléas de référence sur les cas d'étude choisis. Le lien quasi univoque entre période de retour et quantile de distance d'arrêt y a été mis à profit pour affecter une période de retour à chaque distance d'arrêt. L'ensemble des incertitudes a été exprimée sous la forme d'intervalles de crédibilité et il a été montré que la variance et même l'asymétrie de la distribution des distances d'arrêt de référence augmentait avec la période de retour considérée.

Les distributions des autres variables de sortie des différents modèles aux abscisses de grande période de retour ont également été analysées. En outre, une analyse de sensibilité à la plupart des hypothèses de modélisation utilisées a été effectuée et le calcul des pressions d'impact a été mené lorsqu'il était possible. Afin de guider le choix du lecteur que l'un des trois modèles proposés tenterait, leurs forces et faiblesses sont récapitulées dans le Tableau 6.44.

Le modèle dit de Coulomb se base sur un modèle de propagation inversible. Ses avantages sont la facilité de l'inférence, le temps calcul extrêmement réduit qu'il nécessite et sa parcimonie. Un avatar bruité permet aux calculs d'être effectués même sur des topographies de couloir raisonnablement irrégulières. En contrepartie, la connaissance physique disponible n'est que très peu valorisée et ce modèle se comporte comme un opérateur qui fonctionne bien uniquement lorsque la quantité

d'information disponible pour l'inférence est importante. De plus, seule la distribution des distances d'arrêt est utilisable.

Les modèles de propagation Voellmy bloc glissant et Saint-Venant permettent une représentation de plus en plus réaliste de l'avalanche, d'abord sous la forme d'une loi de frottement à deux coefficients puis d'une description fluide de l'écoulement. Ils donnent accès aux distributions des vitesses, hauteurs et coefficient de frottement, même aux abscisses uniquement atteintes par des avalanches de période de retour élevées.

Le calcul de la pression d'impact, grandeur la plus pertinente pour caractériser les dommages potentiels, est délicat du fait de l'absence de modélisation de la densité et de l'incertitude pesant sur le lien pression-vitesse. Néanmoins, les distributions obtenues ont confirmé l'importance de la prise en compte de la rhéologie à proximité immédiate de l'arrêt ainsi que la forte influence de la connaissance de la densité de la neige sur les valeurs de projet.

Tableau 6.44 : Bilan sur les modèles de prédétermination proposés

	Modèle de Coulomb (bruité)	Modèle de Voellmy bloc glissant	Modèle de Saint-Venant
Nature des données nécessaires	Dates de déclenchement, positions de départ et d'arrêt	Dates de déclenchement, positions de départ et d'arrêt, hauteur en écoulement	Dates de déclenchement, positions de départ et d'arrêt, hauteurs et longueurs de départ
Réalisme physique	Très faible	faible	Correct
Temps calcul - ordre de grandeur	Très faible – minutes	Correct - heures	Lourd à très lourd – jours à semaines
Sorties utilisables	Distribution des variables d'entrée, des distances d'arrêt et des aléas de référence	Distribution des variables d'entrée, des distances d'arrêt et des aléas de référence, distribution des vitesses du bloc	Distribution des variables d'entrée, des distances d'arrêt, des vitesses et des hauteurs au cours du temps
Comparaison relative des aléas de référence obtenus	Très « pessimiste »	Légèrement « optimiste »	Moyen
Corrélation entre variables du modèle	Très significative	Significative	Peu significative
Quantité minimale de données nécessaires	Une trentaine d'événements	Une dizaine d'événements	Une dizaine d'événements
Sensibilité aux coefficients de frottement	Forte (un seul coefficient latent)	Forte pour μ_{voel} , assez faible pour ξ_{voel}	Forte pour μ_{voel} , mais moins qu'avec le bloc glissant, assez faible pour ξ_{voel}
Sensibilité aux autres variables d'entrée	Assez faible	Assez faible pour x_{start} , Très faible pour h	Assez faible
Faisabilité du calcul de Pression	/	oui	oui
Fiabilité du calcul de Pression	/	douteuse	Aussi correcte que possible
Remarque	Modèle bruité moins physique nécessaire pour les couloirs à géométrie complexe	Possibilité de fixer ξ_{voel} pour réduire le temps calcul	/

- Des aléas de référence robustes, mais difficiles à valider

Les aléas de référence obtenus se sont révélés relativement robustes vis-à-vis des hypothèses de modélisation effectuées. Les trois modèles ont, en particulier, donné des résultats relativement cohérents en termes de distance d'arrêt lorsqu'ils ont été utilisés sur le même cas d'étude. Néanmoins, la forte influence du coefficient de frottement latent sur la variabilité de la distribution des distances

d'arrêt a été mise en évidence. De même, l'influence de l'information *a priori* sur les valeurs de projet a été établie, même si, pour le coefficient de frottement turbulent, elle ne s'est manifestée sur le cas d'étude traité qu'au niveau de la distribution des vitesses. Enfin, le modèle de Coulomb s'est montré nettement plus pessimiste que les deux autres modèles, en particulier à quantité d'information équivalente. En outre, les vitesses obtenues avec le modèle de Voellmy bloc glissant se sont révélées plus fortes à abscisse égale qu'avec le modèle de Saint-Venant.

La comparaison des résultats fournis par les différents modèles est toutefois difficile du fait de la nécessité de transformer des observations très bruitées pour les faire correspondre aux variables d'entrée et d'utiliser des lois *a priori* au moins partiellement informatives pour les modèles Voellmy bloc glissant et Saint-Venant. Plus généralement, la validation des aléas de référence obtenus reste un problème quasi insoluble du fait de la rareté des événements considérés. Acquérir davantage de confiance dans les modèles proposés nécessitera donc encore par la suite beaucoup de travail, notamment leur application à un grand nombre de cas d'études et la confrontation avec d'autres approches (modélisation déterministe, dendrochronologie, etc.).

7. Dimensionnement d'un ouvrage de protection

Préambule

Ce chapitre aborde le problème du dimensionnement d'un ouvrage de protection. Un exemple réel complexe illustre tout d'abord la difficulté à obtenir des aléas de référence multivariés. Le problème du dimensionnement d'une digue verticale est donc traité ensuite dans le cadre décisionnel présenté au chapitre 3.6. Les calculs des risques classique et bayésien sont effectués analytiquement avec un modèle d'aléa monodimensionnel explicite, puis numériquement à l'aide du modèle d'aléa Voellmy bloc glissant décrit au chapitre 4.2. Une étude de sensibilité paramétrique est menée et l'accent est mis sur l'impact de l'erreur d'estimation sur la décision.

7.1. Définir un aléa de référence bivarié pour le dimensionnement

7.1.1. L'exemple du paravalanche de Taconnaz

Sur le couloir EPA numéro 9 de la commune des Houches, site dit de Taconnaz, existe depuis la fin des années 80 un ouvrage paravalanche de très grande dimension protégeant différents hameaux. Plusieurs grosses avalanches ont souligné récemment le sous-dimensionnement de ce paravalanche : en 1999, une avalanche très dynamique de neige froide et sèche a débordé et endommagé le dispositif, tandis qu'au printemps 2006, une autre avalanche cette fois-ci de neige humide l'a en bonne partie rempli et a brisé certains des dispositifs freineurs qui s'y trouvaient. Le Cemagref a été consulté pour une expertise visant à améliorer cet ouvrage avec comme objectif de protection, conformément au guide PPRA (MEDD et al., 2004), l'événement centennal. Se pose donc en premier lieu le problème de la caractérisation d'un tel événement.



Figure 7.1 : Dépôt de l'avalanche d'avril 2006 sur le site de Taconnaz

Commune des Houches, Haute-Savoie, France.

Photo prise depuis la digue latérale droite du paravalanche.

La neige, lourde et humide, forme des boules.

La taille du dépôt par rapport aux arbres et au paravalanche dont on distingue à peine la digue en rive gauche est remarquable.

Photo N. Eckert/Cemagref.

L'ampleur exceptionnelle du couloir rend les calculs de propagation très lourds. De plus, la non stationnarité de la série des distances d'arrêt due à l'existence d'un premier ouvrage paravalanche rend compliqué le calcul de périodes de retour à partir des distances d'arrêt. Les résultats obtenus au chapitre 6 ont en outre montré que, même si la distance d'arrêt centennale peut être obtenue, choisir un seul scénario de référence à partir de distributions de hauteur, vitesse et/ou pression est extrêmement délicat.

L'approche développée jusqu'ici aurait donc difficilement pu être appliquée dans le cadre de l'expertise. C'est de ce fait une approche statistique plus simple et dérivée de la pratique en hydrologie qui a été mise en œuvre à partir des événements enregistrés dans l'EPA. Afin que le dispositif puisse remplir son rôle à la fois vis-à-vis des avalanches de neige lourde et/ou humide de grande ampleur (type 2006) et des avalanches de neige froide et sèche très énergétiques (type 1999), deux variables

ont dû être considérées : le volume de dépôt et une variable représentant la dynamique de l'écoulement. Pour cette dernière, c'est le nombre de Froude à l'entrée du dispositif paravalanche qui a été retenu. Les volumes de dépôt ont été obtenus classiquement par conversion du triplet longueur-largeur-hauteur figurant dans les carnets forestiers. Quant aux nombres de Froude, ils ont été obtenus par rétro-analyse des événements à l'aide d'un modèle de type Saint-Venant. Davantage d'informations concernant ce cas d'étude peuvent être trouvées en Annexe A4.

Le traitement statistique propre de la distribution des valeurs extrêmes d'un processus bivarié est, comme l'a brièvement explicité le chapitre 2, un problème délicat, faisant encore actuellement l'objet de recherches. Le principal verrou est lié à l'existence d'une infinité de modèles limites. L'exercice est considérablement facilité si l'on considère les deux variables marginales v_{stop} et Fr indépendantes.

Cela dispense de considérer la dépendance asymptotique entre les deux séries de valeurs extrêmes. La probabilité au non dépassement d'un doublet de valeurs définissant une période de retour est alors simplement le produit des probabilités marginales au non-dépassement correspondantes (eq. 7.1) et l'on peut se contenter de travailler sur les lois marginales. Cette hypothèse peut être justifiée d'un point de vue physique : la dynamique de l'écoulement est en effet sensée être déterminée par la qualité de la neige uniquement, de sorte que le volume de neige ne devrait pas intervenir. D'un point de vue statistique, la corrélation empirique entre Fr et v_{stop} est également loin d'être apparente sur le cas d'étude, le coefficient de corrélation empirique entre les nombres de Froude rétro-analysés et les volumes de dépôt valant seulement 0.28.

$$F(v_{stop}, Fr) = F(v_{stop}) \times F(Fr) = 1 - \frac{1}{T} \quad (\text{eq. 7.1})$$

7.1.2. Un modèle POT bivarié indépendant

7.1.2.1. Modèles POT marginaux

Plutôt qu'une approche directe visant à caler une loi sur l'ensemble des deux séries, la modélisation POT, courante dans le domaine des crues et présentée brièvement au chapitre 2, a été choisie. Elle consiste à ne considérer que les observations dépassant un certain seuil. Ce faisant, l'information utilisée pour l'inférence est appauvrie mais, en contrepartie, la détermination de quantiles de projet est relativement aisée. En effet, l'ajustement de la loi se fait uniquement sur les observations les plus élevées, de sorte que le problème du non enregistrement dans l'EPA des avalanches de faible ampleur ne se pose pas. Ensuite, le modèle prend en compte à la fois la fréquence et l'intensité des dépassements, de sorte que le passage aux périodes de retour est direct. Enfin, le problème du choix de modèle est considérablement simplifié grâce à la convergence des dépassements vers des lois limites de formes paramétriques bien connues (Pickands, 1975).

Pour les deux séries de volumes de dépôt et de nombres de Froude, on suppose donc les observations dépassant les seuils $s_{v_{stop}}$ et s_{Fr} distribuées selon des lois de Pareto généralisées (Coles, 2001) de paramètres de forme et d'échelle respectifs $(\beta_{v_{stop}}, \rho_{v_{stop}})$ et (β_{Fr}, ρ_{Fr}) (eq. 7.2). De plus, on suppose pour les deux séries le nombre total de dépassements sur la période d'observation T_{obs} tiré selon une loi de Poisson (eq. 7.3), les paramètres $\lambda_{v_{stop}}$ et λ_{Fr} représentant alors les fréquences moyennes au dépassement. Notons que, comme les deux variables sont supposées indépendantes, les deux seuils n'ont pas besoin d'être dépassés pour qu'un événement soit utilisé pour l'estimation, ce qui implique que des nombres de dépassements N_1 et N_2 différents peuvent être considérés.

$$\begin{cases} v_{stop} | v_{stop} \geq s_{v_{stop}} \sim GPD(\beta_{v_{stop}}, \rho_{v_{stop}}) \\ Fr | Fr \geq s_{Fr} \sim GPD(\beta_{Fr}, \rho_{Fr}) \end{cases} \quad (\text{eq. 7.2})$$

$$\begin{cases} \sum_{i=1}^{N_1} I_{\{v_{stop_i} \geq s_{v_{stop}}\}} \sim P(\lambda_{v_{stop}} \times T_{obs}) \\ \sum_{i=1}^{N_2} I_{\{Fr_i \geq s_{Fr}\}} \sim P(\lambda_{Fr} \times T_{obs}) \end{cases} \quad (\text{eq. 7.3})$$

Pour les deux séries, un rôle crucial est joué par le paramètre de forme, soit $\beta_{v_{stop}}$ et β_{Fr} respectivement. Celui-ci traduit l'incrément du quantile avec la période de retour. Son signe définit trois familles de fonctions de répartition correspondant aux domaines d'attraction du maximum d'un échantillon sur une période unitaire. D'un point de vue plus pratique, la possible variation de signe du paramètre de forme rend l'inférence délicate. La solution retenue a été de tester systématiquement les 3 familles de lois et de retenir le meilleur modèle sur la base du facteur de Bayes calculé grâce à la méthode de Kaas et Raftery présentée au chapitre 2.

Avec un modèle POT, le passage aux quantiles de projet est remarquablement facile. N'importe quel quantile annuel correspondant à une probabilité annuelle de non dépassement $1 - \frac{1}{T}$ peut en effet être

obtenu, en distinguant toutefois les cas des domaines de Fréchet et Weibull (eq 7.4) du domaine de Gumbel (eq 7.5). On peut donc, à partir de la loi *a posteriori* des paramètres, directement simuler les lois prédictives de n'importe quel quantile annuel marginal. Comme pour les lois prédictives des quantiles annuels de distance d'arrêt proposées au chapitre 6, celles ci résument l'information contenue dans les données ainsi que l'incertitude liée à l'estimation des paramètres. L'inversibilité de la loi des dépassements permettant d'obtenir l'expression analytique du quantile pour un jeu de paramètres donné rend cependant le calcul bien moins lourd.

$$\left\{ \begin{array}{l} v_{stop_T} = s_{v_{stop}} + \frac{1}{\beta_{v_{stop}}} \times \left(1 - \left(\frac{-1}{\lambda_{v_{stop}}} \times \ln \left(1 - \frac{1}{T} \right) \right)^{\frac{\beta_{v_{stop}}}{\rho_{v_{stop}}}} \right) \quad si \quad \beta_{v_{stop}} \neq 0 \\ Fr_T = s_{Fr} + \frac{1}{\beta_{Fr}} \times \left(1 - \left(\frac{-1}{\lambda_{Fr}} \times \ln \left(1 - \frac{1}{T} \right) \right)^{\frac{\beta_{Fr}}{\rho_{Fr}}} \right) \quad si \quad \beta_{Fr} \neq 0 \end{array} \right. \quad (eq. 7.4)$$

$$\left\{ \begin{array}{l} v_{stop_T} = s_{v_{stop}} - \frac{1}{\rho_{v_{stop}}} \ln \left(\frac{-1}{\lambda_{v_{stop}}} \ln \left(1 - \frac{1}{T} \right) \right) \quad si \quad \beta_{v_{stop}} = 0 \\ Fr_T = s_{Fr} - \frac{1}{\rho_{Fr}} \ln \left(\frac{-1}{\lambda_{Fr}} \ln \left(1 - \frac{1}{T} \right) \right) \quad si \quad \beta_{Fr} = 0 \end{array} \right. \quad (eq. 7.5)$$

7.1.2.2. Modélisation bivariée et courbes enveloppes

Reste, pour obtenir l'événement centennal recherché, à associer les résultats obtenus marginalement pour les volumes et les nombres de Froude. Pour ce faire, les équations 7.1 et 7.4 peuvent être combinées pour obtenir, dans le cas où les deux paramètres de forme $\beta_{v_{stop}}$ et β_{Fr} sont non nuls, l'eq. 7.6 donnant, en fonction des deux triplets de paramètres et des deux seuils des modèles POT marginaux, le nombre de Froude correspondant à une période de retour T donnée pour un volume de dépôt fixé. L'équation 7.6 définit une courbe enveloppe sur laquelle se trouve l'ensemble des couples (Fr, v_{stop}) correspondant à la période de retour T . Cela montre bien que, dans un cas bivarié, il n'existe plus de lien univoque entre probabilité au non dépassement et période de retour et que l'on se retrouve en particulier avec une infinité de scénarii centennaux. A noter que, ce faisant, on fait une entorse à la règle que l'on s'était fixée jusqu'à présent et consistant à ne calculer des périodes de retour que pour la distance d'arrêt.

$$Fr(T, v_{stop}) = s_{Fr} + \frac{1}{\beta_{Fr}} \times \left(1 - \left(\frac{-1}{\lambda_{Fr}} \times \ln \left(1 - \frac{1}{T} \right) - \frac{\lambda_{v_{stop}}}{\lambda_{Fr}} \times \left(1 - \beta_{v_{stop}} \times (v_{stop} - s_{v_{stop}}) \right)^{\frac{\rho_{v_{stop}}}{\beta_{v_{stop}}}} \right)^{\frac{\beta_{Fr}}{\rho_{Fr}}} \right) \quad (eq. 7.6)$$

7.1.3. Application

7.1.3.1. Etude de la série de volumes

L'étude porte sur une série de 46 événements en 102 hivers d'observation. La modélisation POT permet théoriquement de retenir plus d'une valeur par hiver. Cependant, les volumes de dépôt correspondant à différents événements successifs se produisant au cours d'un même hiver peuvent éventuellement se superposer dans le dispositif. Cela a conduit à n'accepter qu'un seul événement par hiver afin de garantir l'indépendance des dépassements. Afin de dimensionner la capacité de stockage du paravalanche, c'est le maximum hivernal qui a été retenu, de sorte que les données sont interprétées comme des volumes de neige dans le dispositif plutôt que comme des volumes correspondant à des avalanches bien individualisées. Notons toutefois que le cas de plusieurs avalanches au cours du même hiver ne se présente que quelques fois au cours de la période d'observation.

Le principal problème pratique lié à la mise en œuvre du modèle est le choix du seuil : celui-ci doit être suffisamment haut pour garantir la validité du modèle grâce à la convergence vers le modèle asymptotique mais suffisamment bas pour conserver une quantité de données satisfaisante. On a utilisé la solution empirique proposée par Coles (2001), c'est-à-dire la sélection du seuil qui correspond au début de la partie linéaire du graphique représentant la moyenne arithmétique des dépassements en fonction du seuil. Son application à la série de volumes de dépôt conduit à un seuil de 250 000 m³ et à une série de 20 dépassements (Tableau 7.1). La distribution obtenue est très asymétrique à cause de l'événement du 14 février 1950 qui dépasse très largement toutes les autres valeurs avec un volume supérieur à 4 millions de m³.

Tableau 7.1 : Statistiques descriptives des séries de dépassement

	v_{stop} (m ³)	Fr
Seuil	250 000	1.82
Moy.	806 375	2.57
SD	918 460	0.64
Skew	2.92	0.86
q₁₀	275 000	1.95
q₂₅	330 750	2.07
q₅₀	468 750	2.39
q₇₅	905 000	3.03
q₉₀	1 395 000	3.58

Les paramètres du modèle POT ont été inférés dans le cadre bayésien (Parent et Bernier, 2003a), avec utilisation systématique de lois *a priori* non informatives, et ce pour $\beta_{v_{stop}} > 0$, $\beta_{v_{stop}} = 0$ et $\beta_{v_{stop}} < 0$.

Le Tableau 7.2 montre que le modèle $\beta_{v_{stop}} < 0$ l'emporte assez nettement (Tableau 7.2). La sélection de ce modèle traduit une queue de distribution « lourde », c'est-à-dire avec des probabilités plus élevées associées aux quantiles très élevés qu'avec un modèle exponentiel simple où $\beta_{v_{stop}} = 0$.

Tableau 7.2 : Facteurs de Bayes pour les volumes de dépôt

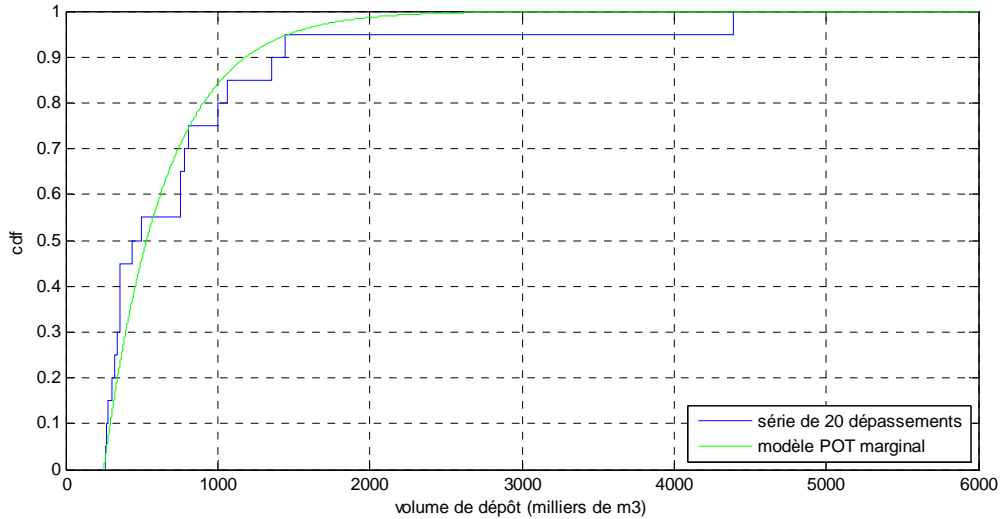
	$\beta_{v_{stop}} = 0$	$\beta_{v_{stop}} > 0$
$\beta_{v_{stop}} < 0$	29.4	7.70
$\beta_{v_{stop}} = 0$	/	0.26

Le Tableau 7.3 présente les principales caractéristiques des lois *a posteriori* des trois paramètres pour $\beta_{v_{stop}} < 0$. Le paramètre β est très faible, ce qui traduit tout de même un comportement assez proche de celui d'une loi exponentielle simple. L'adéquation entre données et moyenne *a posteriori* du modèle, sans être parfait, est satisfaisant (Figure 7.2). On voit en particulier que l'ajustement se fait bien pour les volumes entre 500 000 et 1 500 000 m³, mais que le maximum empirique de 1950 est réellement exceptionnel et de ce fait difficile à prendre en compte.

Tableau 7.3 : Lois a posteriori des paramètres du modèle POT pour les volumes de dépôt

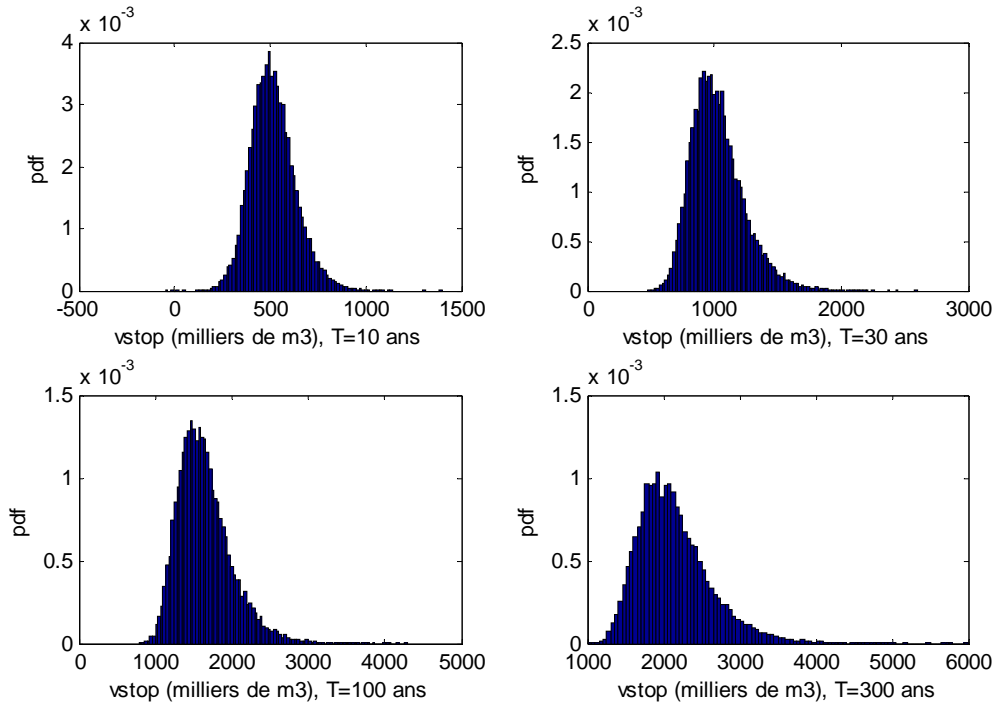
	Moy.	SD	q _{2.5}	q _{97.5}
$\beta_{v_{stop}}$	-0.00015	0.00006	-0.00024	-0.00002
$\lambda_{v_{stop}}$	0.199	0.045	0.121	0.294
$\rho_{v_{stop}}$	0.0025	0.0005	0.0016	0.0035

Figure 7.2 : Adéquation entre données et modèle pour les volumes de dépôt



Site de Taconnaz, commune des Houches, Haute-Savoie, France.

Figure 7.3 : Distribution prédictive des quantiles annuels pour les volumes de dépôt



Site de Taconnaz, commune des Houches, Haute-Savoie, France.

Les lois prédictives des quantiles annuels classiques correspondant à des périodes de retour de 10, 30, 100 et 300 ans ont été obtenues pendant l'inférence en calculant la valeur des quantiles à chaque itération. Elles montrent que, logiquement, la moyenne, la dispersion et l'asymétrie de la distribution des quantiles annuels augmentent avec la période de retour (Figure 7.3). Il apparaît en particulier que

la valeur centennale recherchée se situe autour de 1 600 000 m³. On voit également que l'incertitude autour de cette valeur reste forte puisque l'intervalle de crédibilité à 95% va jusqu'à plus de 2 600 000 m³ (Tableau 7.4). Enfin, l'influence de la donnée de 1950 est incontestable puisque, si les calculs sont refaits sans elle, une valeur de projet centennale moyenne inférieure de 500 000 m³ est obtenue.

Tableau 7.4 : Quantiles annuels pour les volumes de dépôt

	Moy.	SD	q_{2.5}	q_{97.5}
$v_{stop_{10}}$ (milliers de m ³)	515	118.8	304.4	775.7
$v_{stop_{30}}$ (milliers de m ³)	1028	209.2	700.8	1516
$v_{stop_{100}}$ (milliers de m ³)	1644	353	1112	2485
$v_{stop_{300}}$ (milliers de m ³)	2138	483.3	1423	3303

7.1.3.2. Etude de la série des nombres de Froude

L'analyse de la série des nombres de Froude a été menée de manière strictement analogue à celle des volumes. Le seuil retenu d'après Coles (2001) est 1.82. Il permet d'obtenir une sous-série de 20 dépassements indépendants (Tableau 7.1). Contrairement aux volumes, chaque valeur correspond bien évidemment à un événement bien individualisé. Le facteur de Bayes sélectionne nettement le domaine d'attraction de Weibull correspondant à β_{Fr} positif (Tableau 7.5). La distribution du nombre de Froude est ainsi bornée avec un nombre de Froude limite de 6.92. Le quantile centennal moyen s'établit à 3.58 (Tableau 7.6). A noter qu'il s'agit bien maintenant d'une valeur possédant une probabilité de dépassement annuelle de 1/100 et plus, comme au chapitre 6, du nombre de Froude associé à une avalanche d'abscisse centennale.

Tableau 7.5 : Lois a posteriori des paramètres du modèle POT pour les nombres de Froude

	Moy.	SD	q_{2.5}	q_{97.5}
β_{Fr}	0.196	0.116	0.010	0.411
λ_{Fr}	0.198	0.044	0.120	0.293
ρ_{Fr}	1.41	0.33	0.82	2.12

Tableau 7.6 : Quantiles annuels pour les nombres de Froude

	Moy.	SD	q_{2.5}	q_{97.5}
Fr_{10}	2.25	0.18	1.91	2.64
Fr_{30}	2.94	0.24	2.52	3.46
Fr_{100}	3.58	0.31	3.05	4.29
Fr_{300}	3.98	0.39	3.39	4.91

7.1.3.3. Analyse croisée

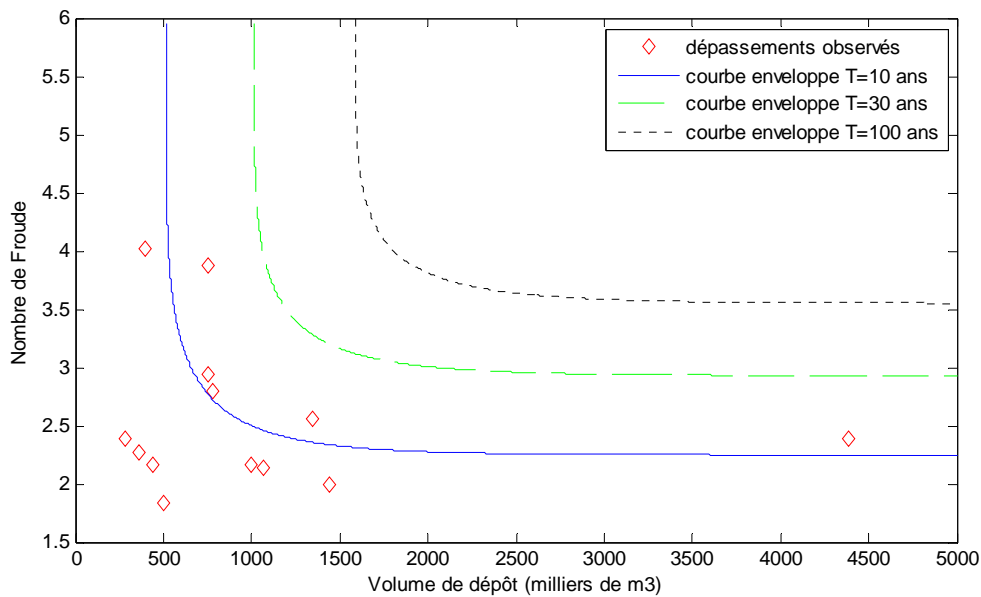
Les estimateurs des paramètres des deux modèles marginaux ont permis le calcul des courbes enveloppes moyennes correspondant à diverses périodes de retour selon l'éq. 7.6. Celles-ci sont comparées aux données dépassant les deux seuils marginaux en Figure 7.4. Du fait des domaines d'attraction respectifs des lois marginales, il y a un nombre de Froude limite, mais pas de volume limite théorique.

Les conditions de positivité dans l'équation 7.6 impliquent que, pour une période de retour donnée, les valeurs marginales constituent également des valeurs limites. Par exemple, pour une période de retour de 100 ans, le volume centennal marginal est le plus « petit » possible sans être jamais tout à fait atteint. Des volumes à peine plus grands correspondent à de très forts nombres de Froude, tandis que des volumes plus élevés correspondent à des nombres de Froude plus faibles. A noter enfin que les doubles dépassements observés correspondent, du point de vue bivarié, à des périodes de retour assez faibles : l'événement de 1950, très exceptionnel du point de vue du volume est plus que banal du point

de vue du nombre de Froude, de sorte que sa période de retour est de l'ordre de 20 ans à peine. Idem pour celui, très dynamique, de 1999 avec un nombre de Froude de 3.9 mais un volume relativement faible.

Puisqu'il n'est pas question de proposer une infinité de scénarii de référence, il a fallu extraire des courbes enveloppes quelques scénarii pertinents. La solution pratique retenue lors de l'expertise a été de considérer les deux couples correspondant aux bornes supérieures des intervalles de crédibilité à 95% sur les valeurs de projets marginales. Ainsi l'étude des volumes a-t-elle montré que l'on est sûr à 95% que le volume centennal est inférieur à 2.48 millions de m³. Le nombre de Froude correspondant définit un premier scénario de référence de type « gros volume ». De même, le couple nombre de Froude/volume associé à la borne supérieure de l'intervalle de crédibilité à 95% du nombre de Froude centennal définit un second scénario de référence de type « Froude fort » (Tableau 7.7).

Figure 7.4 : Courbes enveloppes pour différentes périodes de retour



Site de Taconnaz, commune des Houches, Haute-Savoie, France.

Tableau 7.7 : Scénarii centennaux retenus

	Scénario « Gros volume »	Scénario « Froude fort »
Volume de dépôt (millions de m³)	2.48	1.68
Nombre de Froude	3.64	4.29

7.1.4. Vers un modèle décisionnel

Ainsi une réponse à la question « quel est l'événement centennal à considérer sur ce site ? » a pu être apportée pour le dimensionnement fonctionnel de l'ouvrage. Mais la méthode proposée, outre le fait que l'indépendance entre nombre de Froude et volume de dépôt peut être contestée, est un peu bancal puisque l'on a été amené à choisir deux scénarii centennaux parmi une infinité.

De plus, le problème du calcul des pressions d'impact n'a pas été abordé. Celles-ci doivent pourtant être évaluées pour le dimensionnement structural sous peine de voir le dispositif endommagé, voire détruit, au premier événement. On peut évidemment calculer les pressions correspondant aux deux scénarii centennaux retenus selon l'une des approches proposées au chapitre 6.5. Néanmoins, il ne s'agira pas là de pressions centennales. En outre, étant donné que des pressions d'impact très fortes sont générées par des écoulements peu dynamiques, on est même sûr que des pressions bien plus fortes sont générées par des avalanches moins intenses du point de vue du volume et/ou du nombre de Froude. Ces dernières, comme celle d'avril 2006, ne menacent pas les enjeux mais sont susceptibles d'endommager le dispositif. Une modélisation statistique de la variable Pression doit donc, en toute rigueur, également être effectuée. Mais la modélisation d'un couple de valeurs extrêmes étant déjà difficile, le passage à un triplet rend l'approche proposée quasi ingérable, avec pour une période de retour non plus des courbes mais des surfaces enveloppes...

L'approche par aléa de référence montre donc ses limites pour le dimensionnement d'un ouvrage paravalanche du fait du caractère multivarié de l'aléa. On se propose par conséquent dans le restant du chapitre de dimensionner un ouvrage paravalanche à l'aide d'une approche décisionnelle permettant de minimiser la perte moyenne attendue plutôt que par un calcul de période de retour. L'ouvrage de protection envisagé restera cependant bien plus simple que le paravalanche de Taconnaz puisqu'on se limitera, conformément au cadre général proposé au chapitre 3.6, au cas d'une digue verticale. Les deux modèles décisionnels qui vont à présent être proposés restent cependant des explorations de pistes de recherche. Du point de vue de l'ingénierie, ils ne sont pas encore opérationnels et n'auraient donc pas permis d'apporter des réponses plus pertinentes que celles fournies par l'approche par aléa de référence au problème spécifique posé par le site de Taconnaz.

7.2. Un modèle d'optimisation simplifié analytique

7.2.1. Des hypothèses supplémentaires

7.2.1.1. Un modèle POT conjugué simplifié

Afin de mettre en place l'implémentation d'une approche décisionnelle de dimensionnement d'une digue paravalanche conforme au cadre proposé au chapitre 3.6, on se limite dans un premier temps à la modélisation des distances d'arrêt excédant l'abscisse de la digue. De plus, on suppose les dépassements distribués selon une loi exponentielle avec un paramètre d'échelle ρ dont l'inverse correspond à la moyenne des dépassements (eq.7.7), de sorte que $y = y_o = x_{stop_o}$ et $\theta_M = \rho$. De façon plus classique, on modélise la fréquence des dépassements par un modèle poissonien de paramètre $\theta_F = \lambda$ (eq. 7.8). Le modèle fréquentiel est donc identique à celui du chapitre 5 mais a_t représente à présent les nombres de dépassements annuels plutôt que les nombres totaux d'occurrences annuelles. Par rapport au modèle utilisé pour les volumes et les nombres de Froude sur le site de Taconnaz, ce modèle d'aléa est un modèle POT dont le paramètre de forme a été fixé à zéro et le seuil fixé arbitrairement à l'abscisse de la digue. L'abscisse d'arrêt x_{stop_r} correspondant à n'importe quelle période de retour T est extrêmement simple à obtenir (eq. 7.9).

$$l(x_{stop_o} | \rho, x_{stop_o} > x_d) = \rho \times \exp(-\rho \times (x_{stop_o} - x_d)) \quad (\text{eq. 7.7})$$

$$l(a_t | \lambda) = \frac{\lambda^{a_t}}{a_t!} \times \exp(-\lambda) \quad (\text{eq. 7.8})$$

$$x_{stop_r} = \frac{1}{\rho} \times \ln(\lambda T) + x_d \quad (\text{eq. 7.9})$$

L'intérêt d'un tel modèle d'aléa simplifié est de permettre le calcul analytique des lois *a posteriori* moyennant le choix de lois *a priori* conjuguées. Le calcul a déjà été détaillé au chapitre 5.1 pour le modèle fréquentiel. Pour le modèle d'intensité, c'est également un *prior* de type Gamma indexé par les paramètres (a_ρ, b_ρ) qui convient (eq. 7.10). La loi *a posteriori* de ρ est alors toujours Gamma avec des paramètres (a'_ρ, b'_ρ) vérifiant $a'_\rho = a_\rho + S(N)$ et $b'_\rho = b_\rho + N$ où $S(N)$ est la somme des dépassements. Bien entendu, on retrouve, comme pour le modèle fréquentiel, que pour un nombre d'événements assez important, la moyenne *a posteriori* de ρ converge vers son estimateur du maximum de vraisemblance, c'est-à-dire l'inverse de la moyenne des dépassements observés.

$$\pi(\rho | a_\rho, b_\rho) = \frac{a_\rho^{b_\rho}}{\Gamma(b_\rho)} \times \rho^{b_\rho-1} \times \exp(-a_\rho \times \rho) \quad (\text{eq.7.10})$$

7.2.1.2. Utilité et coût à partir des distances d'arrêt de référence

Le calcul du coût et de l'utilité associés à chaque événement s'effectue par combinaison du modèle d'aléa et de la fonction de coût retenue. Afin de poursuivre les calculs analytiques, on retient la formulation des coûts correspondant à un dommage maximal dès lors que l'enjeu est atteint, à savoir

l'éq. 3.38. Par rapport aux notations du chapitre 3.6, on a $E_d [a | \theta_F, \theta_M] = \lambda$, la dépendance vis-à-vis du paramètre d'intensité traduisant le fait que seuls les dépassements de l'abscisse de la digue sont considérés.

De plus, si l'on utilise la réduction de distance d'arrêt de l'équation 3.30 exprimant $x_{stop}(h_d)$ en fonction de la distance d'arrêt de référence x_{stop_o} , de la hauteur de la digue h_d et de la hauteur de l'écoulement de référence h_o , le coût correspondant à chaque événement et à chaque hauteur de digue peut être exprimé en fonction de l'événement de référence et de la hauteur de la digue uniquement (eq. 7.11). L'utilité associée traduit alors simplement le fait que la construction de la digue n'apporte un bénéfice que pour les avalanches qui atteignent l'enjeu sans digue mais ne l'atteignent plus avec la digue (eq. 7.12).

$$C(h_d, x_{stop_o}, a) = C_o \times h_d + A_{ct} \times \lambda \times C_1 \times I_{\left\{ (x_{stop_o} - x_d) \times \left(1 - \alpha \times \frac{h_d}{h_o} \right) \geq x_b - x_d \right\}} \quad (\text{eq. 7.11})$$

$$u(h_d, x_{stop_o}, a) = C_o \times h_d + A_{ct} \times \lambda \times C_1 \times I_{\left\{ \frac{x_b - x_d}{\left(1 - \alpha \times \frac{h_d}{h_o} \right)} + x_d \leq x_{stop_o} < x_b \right\}} \quad (\text{eq. 7.12})$$

7.2.1.3. Fonctions de risque analytiques

Les hypothèses de modélisation de l'aléa et des coûts permettent l'intégration analytique de la fonction d'utilité par rapport à la distribution des aléas selon l'éq. 3.40. En effet, il s'agit simplement d'intégrer une loi exponentielle entre des bornes dépendant de l'abscisse de l'enjeu, de l'abscisse de la digue, de la hauteur de l'écoulement de référence et de la hauteur de la digue ainsi que de son pouvoir dissipatif (eq. 7.13). La fonction de risque s'obtient par changement de variable. Elle dépend, outre les grandeurs précédemment mentionnées, des deux paramètres du modèle d'aléa et des coûts de construction et de destruction (eq. 7.14).

$$R_C(h_d, \rho, \lambda) = C_o \times h_d + A_{ct} \times \lambda \times C_1 \times \rho \times \int_{\frac{x_b - x_d}{1 - \alpha \times \frac{h_d}{h_o}} + x_d}^{x_b} \exp(-\rho \times (x_{stop_o} - x_d)) \times dx_{stop_o} \quad (\text{eq. 7.13})$$

$$R_C(h_d, \rho, \lambda) = C_o \times h_d + A_{ct} \times C_1 \times \lambda \times \left(\exp\left(\frac{-\rho \times (x_b - x_d)}{1 - \alpha \times \frac{h_d}{h_o}} \right) - \exp(-\rho \times (x_b - x_d)) \right) \quad (\text{eq. 7.14})$$

$$R_B(h_d) = C_o \times h_d + A_{ct} \times C_1 \times \int_{\lambda=0}^{\infty} \int_{\rho=0}^{\infty} \lambda \times \left(\exp\left(\frac{-\rho \times (x_b - x_d)}{1 - \alpha \times \frac{h_d}{h_o}} \right) - \exp(-\rho \times (x_b - x_d)) \right) \times p(\lambda, \rho | data) d\rho \times d\lambda \quad (\text{eq. 7.15})$$

$$R_B(h_d) = C_o \times h_d + A_{ct} \times C_1 \times \frac{b'_\lambda}{a'_\lambda} \times \left(\left(\frac{a'_\rho}{a'_\rho + \frac{x_b - x_d}{1 - \alpha \times \frac{h_d}{h_o}}} \right)^{b'_\rho} - \left(\frac{a'_\rho}{a'_\rho + x_b - x_d} \right)^{b'_\rho} \right) \quad (\text{eq. 7.16})$$

Le risque bayésien est, conformément à l'éq. 3.42, l'espérance mathématique du risque classique par rapport à la loi jointe *a posteriori* du couple ρ, λ (eq. 7.15). Grâce aux propriétés du modèle, cette dernière se décompose en produit des lois marginales *a posteriori*. L'intégrale sur λ correspond à l'espérance d'une loi Gamma. En outre, la propriété de sommation à 1 d'une loi de probabilité permet le calcul de l'intégrale sur ρ , de sorte qu'une expression simple de la fonction de risque dépendant

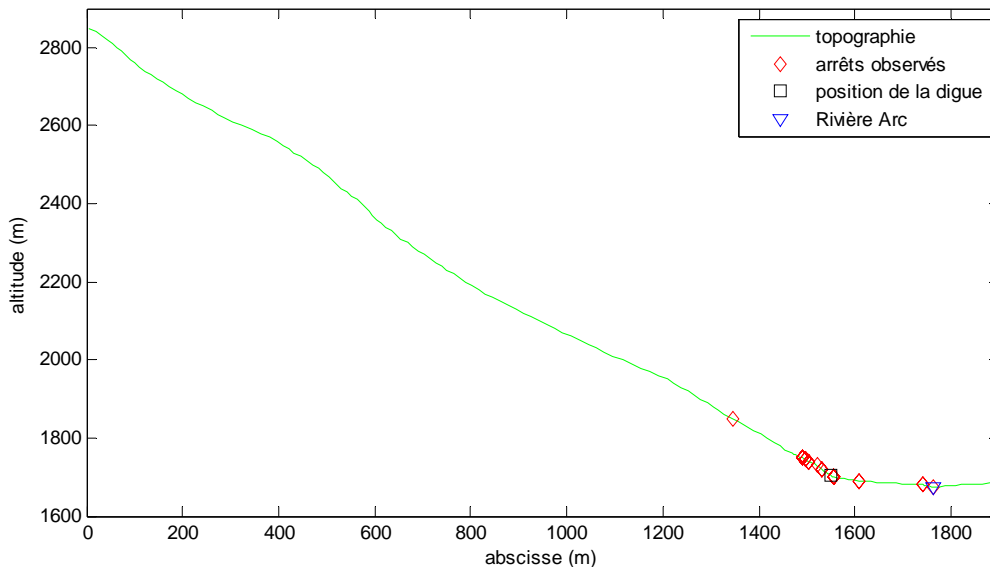
pour l'aléa uniquement des paramètres (a'_ρ, b'_ρ) et (a'_λ, b'_λ) peut être obtenue (eq 7.16). Malheureusement, la minimisation analytique des eqs. 7.14 et 7.16 n'est pas possible. Mais le calcul numérique des deux hauteurs optimales de digue est en revanche très simple.

7.2.2. Application

7.2.2.1. Cas d'étude choisi

Pour pouvoir utiliser le modèle d'aléa très simplifié proposé, la zone d'arrêt se doit d'être particulièrement régulière. On introduit donc un nouveau couloir correspondant au site EPA numéro 30 de la commune de Bessans. D'ampleur intermédiaire entre celle des couloirs 5 et 13 de la même commune avec une longueur de 1763 m et un dénivelé total de 1175 m, il se termine tout comme le couloir 13 voisin dans la vallée de l'Arc. Il est peu chenalisé et la pente moyenne est forte (Tableau 7.8). Le profil de pente est assez tourmenté avec plusieurs ressauts qui rendent difficile la modélisation numérique. Par contre, la zone d'arrêt est formée d'une zone de pente très régulière précédant la rivière et rendant possible l'utilisation d'un modèle stochastique très simple de distribution des distances d'arrêt. Davantage d'informations concernant ce cas d'étude peuvent être trouvées en Annexe A3.

Figure 7.5 : Cas d'étude pour le modèle décisionnel analytique



Couloir EPA 30 de la commune de Bessans, Savoie, France.

On suppose, ce qui n'est pas le cas, qu'un enjeu potentiellement exposé existe en zone d'arrêt à une abscisse correspondant à une période de retour allant de 10 à 1000 ans. Comme précédemment, on travaillera plus précisément aux abscisses correspondant aux périodes de retour classiques de 30, 100 et 300 ans. La position de la digue à dimensionner est fixée au début de la zone d'arrêt puisque les travaux sur les écoulements avalanches indiquent que le pouvoir dissipatif d'une digue est maximum lorsqu'elle est située là où l'énergie de l'écoulement est la plus forte. La pratique courante en ingénierie paravalanche étant de considérer le point où la pente locale passe sous les 10° comme le début de la zone d'arrêt, on fixe $x_d=1550$ m pour le site étudié (Figure 7.5).

Sur la période 1973–2003, 28 avalanches ayant dépassé x_d ont été enregistrées. La plus longue distance d'arrêt enregistrée correspond à l'abscisse de la rivière Arc, mais au-delà, le terrain reste quasiment plat, de sorte que des avalanches extrêmes sont susceptibles d'aller encore plus loin.

Pour le calcul des coûts, la hauteur de l'écoulement de référence, c'est-à-dire de l'avalanche sans digue, est nécessaire. Comme elle n'est pas modélisée avec le modèle d'aléa utilisé, on est contraint de la fixer à une valeur constante d'un événement à l'autre. Une hauteur de référence moyenne de 1 m a été choisie, ce qui implique $h_d < \frac{1}{\alpha} = 7.15$ m pour que l'équation 3.16 décrivant la réduction de la

distance d'arrêt soit valide. On fixe également dans un premier temps $C_o=5530$ €/m, $C_1=300\,000$ € et $A_{ct}=25$. Les coûts de construction et de destruction correspondent à ceux d'une petite digue sans géotextile et d'une maison individuelle, tandis que A_{ct} est obtenu selon l'éq. 3.35 avec un taux d'intérêt annuel constant de 4%.

Tableau 7.8 : Principales caractéristiques du cas d'étude pour le modèle décisionnel analytique

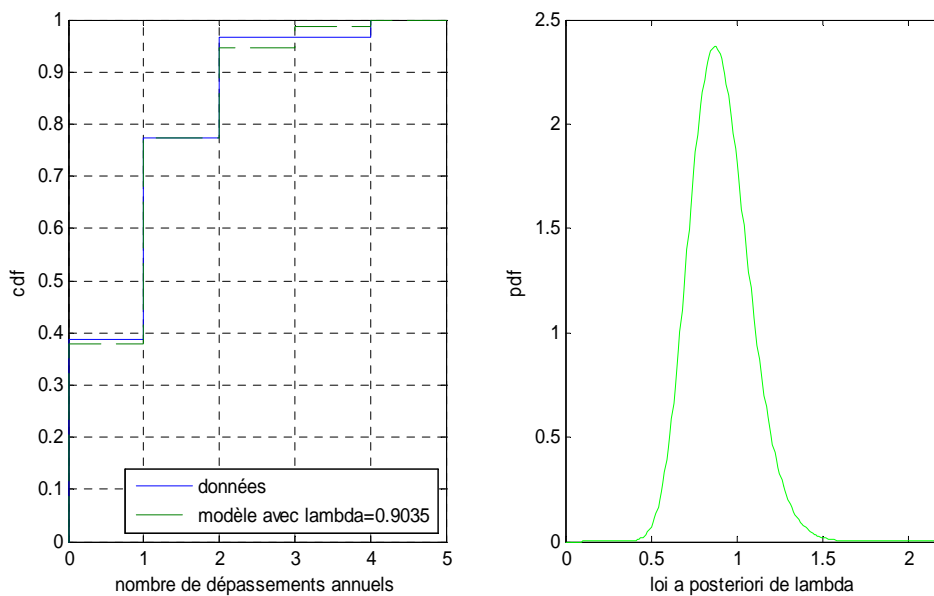
Altitude maximale (m)	2850
Altitude minimale (m)	1675
Dénivelé total (m)	1175
Longueur totale (m)	1763
Pente moyenne (deg)	38.2
Abscisse de la digue (m)	1550
Altitude de la digue (m)	1704
Nombre d'années d'observation	31
Nombre de dépassements	28

7.2.2.2. Inférence bayésienne du modèle POT simplifié

Des priors peu informatifs sont utilisés : on fixe $a_\lambda = a_\rho = 0.001$ et $b_\lambda = b_\rho = 0.01$ respectivement, de sorte que $a'_\lambda = 31.001$, $b'_\lambda = 28.01$, $a'_\rho = 1208.7$ et $b'_\rho = 28.01$. Les Figures 7.6 et 7.7 présentent les lois *a posteriori* Gamma des deux paramètres ainsi qu'une comparaison entre fonctions de répartition des données et de la moyenne *a posteriori* du modèle. Les deux lois *a posteriori* des paramètres présentent des variances bien inférieures à celles des lois *a priori* reflétant l'information apportée par les données. Les moyennes *a posteriori* des paramètres du modèle s'obtiennent aisément à partir des lois *a posteriori* puisque $E_{a_\lambda, b_\lambda}[\lambda] = \hat{\lambda} = \frac{b'_\lambda}{a'_\lambda} = 0.9035$ et $E_{a_\rho, b_\rho}[\rho] = \hat{\rho} = \frac{b'_\rho}{a'_\rho} = 0.0232$.

Le paramètre de fréquence correspond à la fréquence moyenne de dépassement de l'abscisse de la digue tandis que le paramètre d'intensité est l'inverse de la moyenne des dépassements. L'abscisse de la rivière Arc correspond ainsi à une période de retour légèrement supérieure à 100 ans tandis que l'abscisse millénale est située 80 m au-delà de la rivière.

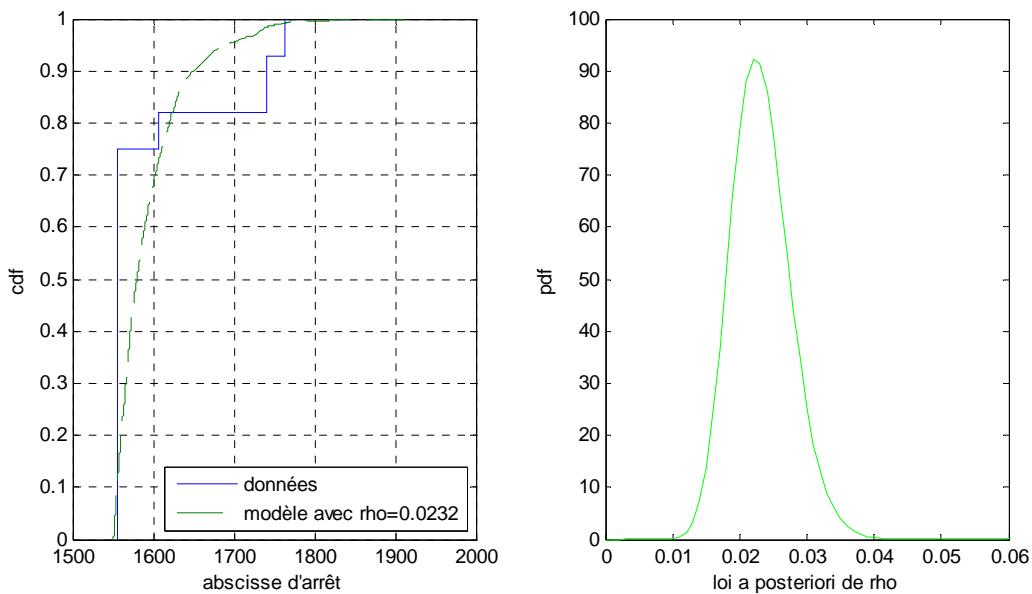
Figure 7.6 : Inférence du modèle fréquentiel



Couloir EPA 30 de la commune de Bessans, Savoie, France. Données et modèle sachant $\lambda = \hat{\lambda}$ (gauche) et loi *a posteriori* du paramètre (droite)

Le modèle fréquentiel sachant $\hat{\lambda} = \hat{\lambda}$ est, comme pour les couloirs étudiés au chapitre 5, très proche des observations. Il représente donc bien la distribution des nombres des dépassements annuels (Figure 7.6). L'adéquation entre le modèle d'intensité sachant $\hat{\rho} = \hat{\rho}$ et les dépassements observés est moins bonne (Figure 7.7). Une partie de cette inadéquation est due à la dépendance de la distance d'arrêt vis-à-vis du profil de pente qui fait s'arrêter de nombreuses avalanches aux abscisses correspondant à une diminution nette de la pente, ce que ne reproduit évidemment pas un modèle statistique simple. Néanmoins cette inadéquation est également due aux erreurs d'observations, ou plutôt aux erreurs de conversion entre abscisse et altitude déjà mentionnées au chapitre 6 pour le couloir 5 de Bessans et qui rendent la distribution empirique faussement discrète. En effet, pour les avalanches s'arrêtant en zone plane, la relation $z=f(x)$ est mal définie de sorte que la véritable abscisse d'arrêt est mal connue et que la distribution réelle des distances d'arrêt est nettement plus régulière que la distribution des observations ne le suggère.

Figure 7.7 : Inférence du modèle d'intensité analytique



Couloir EPA 30 de la commune de Bessans, Savoie, France. Données et modèle sachant $\hat{\rho} = \hat{\rho}$ (gauche) et loi a posteriori du paramètre (droite)

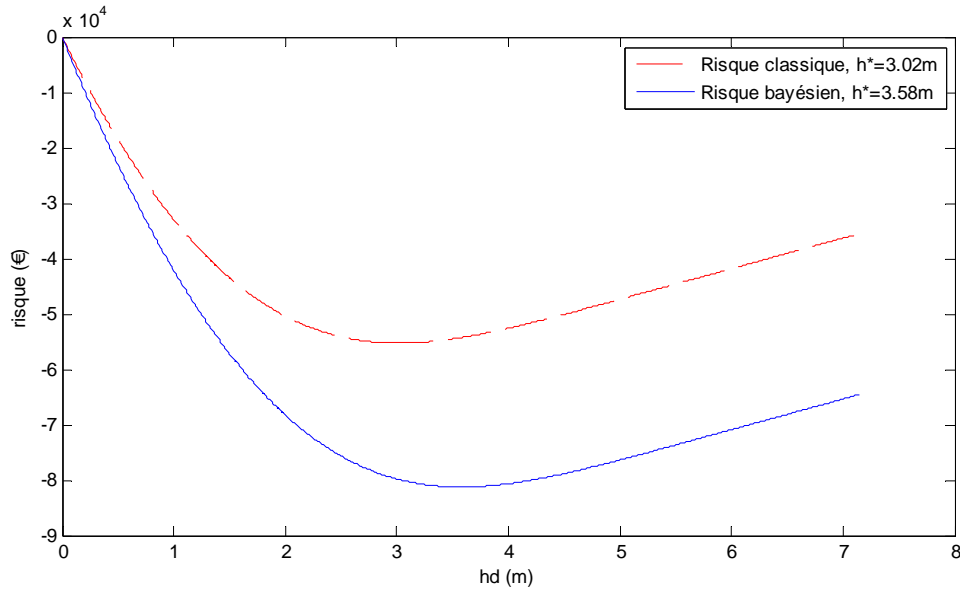
7.2.2.3. Optimisation de la hauteur de la digue pour un enjeu centennal

- *Risque classique*

La Figure 7.8 présente la fonction de risque obtenue pour un enjeu situé à une abscisse centennale. Elle est toujours négative, indiquant que, quelle que soit la hauteur de la digue, la perte moyenne attendue décroît, ce qui justifie le projet de construction. De plus, la fonction de risque présente un minimum net correspondant à une hauteur de digue de 3.02 m.

Pour une hauteur de digue nulle, le bénéfice de la construction est bien évidemment nul puisqu'il s'agit de l'état de référence. Le bénéfice attendu augmente avec la hauteur de la digue jusqu'à une valeur de 55 000 € correspondant à l'optimum économique. Pour des hauteurs de digue plus importantes, le bénéfice moyen attendu décroît car l'effort de protection additionnel ne compense plus l'augmentation du coût de construction. Pour des hauteurs de digue très élevées, la fonction de risque tend d'ailleurs à croître linéairement avec la hauteur de digue, indiquant que le terme de construction devient prédominant par rapport au terme de dommages dans l'équation 3.38.

Figure 7.8 : Risques analytiques pour un bâtiment situé à une abscisse centennale



Couloir EPA 30 de la commune de Bessans, Savoie, France. Modèle décisionnel analytique. $\lambda = 0.9035$ et $\rho = 0.0232$ pour le risque classique. $a'_\lambda = 31.001$, $b'_\lambda = 28.01$, $a'_\rho = 1208.7$ et $b'_\rho = 28.01$ pour le risque bayésien. $C_o = 5530 \text{ €/m}$, $C_1 = 300\,000 \text{ €}$, $\alpha = 0.1376$, $h_o = 1 \text{ m}$ et $A_{ct} = 25$.

- *Risque bayésien*

La Figure 7.8 présente également le risque bayésien à la même abscisse. Comme des *priors* très peu informatifs ont été utilisés pour l'inférence, les deux fonctions correspondent à la même information et peuvent être comparées.

Le risque bayésien est assez proche du risque classique en termes de forme et de valeurs. Néanmoins, l'optimum bayésien est plus élevé que l'optimum classique avec une différence absolue de 0.56 m, soit une différence relative de 18.5% (Tableau 7.9). De plus, le bénéfice attendu de la construction est plus élevé, puisqu'il est de 81 000 € à la hauteur optimale, soit une différence de 40% (Tableau 7.10). Ces résultats sont dus à l'incorporation de l'incertitude liée à l'aléa dans le calcul de risque. En effet, utiliser le *posterior* plutôt qu'un estimateur ponctuel ajoute l'erreur d'estimation à la variabilité naturelle, ce qui est particulièrement critique pour les événements les plus rares justifiant la construction d'une digue. Il est de ce fait compréhensible qu'un effort de protection accru apparaisse économiquement justifié quand l'erreur d'estimation est prise en compte pour la décision.

Enfin, il doit être noté que la fonction de risque bayésienne présente un optimum moins marqué que la fonction de risque classique. Une gamme assez large de hauteurs de digue correspond donc à peu près au même bénéfice moyen attendu. Cela reflète honnêtement la quantité limitée d'information locale disponible pour l'inférence à travers l'effet lissant de la moyenne sur le *posterior*. De ce fait, un optimum peu marqué ne doit pas être vu comme un désavantage de l'approche bayésienne par rapport à l'approche classique. Au contraire, d'un point de vue pratique, différents choix sont alors possibles, donnant ainsi au décideur la possibilité de privilégier différents compromis entre effort de protection et risque résiduel sur la base d'une efficacité économique identique.

Tableau 7.9 : Hauteurs optimales avec le modèle décisionnel analytique

	T ₁₀	T ₃₀	T ₁₀₀	T ₃₀₀	T ₁₀₀₀
x_{stop_T} (m)	1645.7	1693.1	1745.1	1792.5	1844.4
h_C^* (m)	5.42	4.32	3.02	1.77	0.34
h_B^* (m)	5.69	4.73	3.58	2.47	1.19
$\delta_{h^*} = h_B^* - h_C^*$ (m)	0.27	0.41	0.56	0.7	0.85
δ_{h^*} (%)	5	9.5	18.5	39.6	250

Tableau 7.10 : Bénéfices associés aux hauteurs optimales avec le modèle décisionnel analytique

	T ₁₀	T ₃₀	T ₁₀₀	T ₃₀₀	T ₁₀₀₀
b_C^* (€)	718 851	224 126	55 249	11 095	250
b_B^* (€)	781 613	271 277	81 219	23 275	3382
$\delta_{b^*} = b_C^* - b_B^*$ (€)	62 762	47 151	25 970	12 180	3132
δ_{b^*} (%)	8.7	21.0	47.0	109.8	1254.4

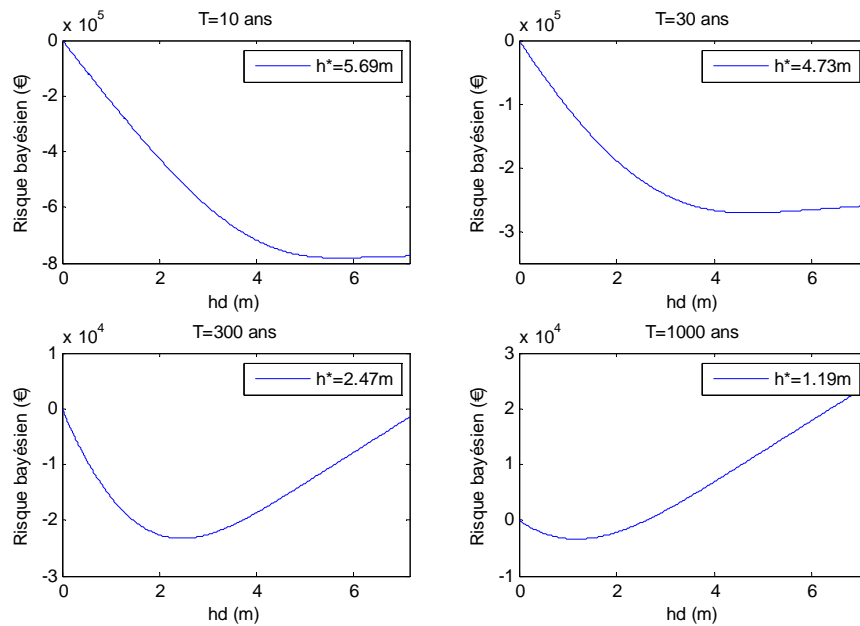
7.2.3. Analyse de sensibilité

Comme l'expression analytique des risques classiques et bayésiens est disponible, une analyse de sensibilité aux différents facteurs influençant la décision optimale est facile à mettre en œuvre sous la forme d'une étude paramétrique. En plus de l'effet déjà discuté de l'erreur d'estimation, les effets de l'intensité de l'aléa, de la fréquence de l'aléa, de la position de l'enjeu, des coûts de construction et de destruction ainsi que du pouvoir dissipatif de la digue peuvent être étudiés. Les résultats sont illustrés principalement avec les fonctions de risque bayésien, mais des résultats très similaires peuvent être obtenus avec les fonctions de risque classique.

7.2.3.1. Intensité des avalanches et position de l'enjeu

La fonction de risque bayésien dépend d'une combinaison entre l'intensité de l'aléa et la position de l'enjeu à travers la somme $a'_p + x_b - x_d$. La dépendance en $x_b - x_d$ plutôt que vis-à-vis de l'abscisse elle-même confirme que la topographie de la zone d'arrêt n'est pas prise en compte dans le modèle puisque seule la distance à la digue compte. En outre, tout changement dans l'intensité de l'aléa peut être compensé par un changement dans la position de l'enjeu pour obtenir la même fonction de risque. Pour une distance $x_b - x_d$ donnée, la hauteur de digue optimale augmente avec l'intensité de l'aléa. Réciproquement, pour une intensité d'aléa donnée, la hauteur optimale décroît avec la distance de l'enjeu jusqu'à devenir nulle, indiquant que, logiquement, pour un enjeu situé extrêmement loin de la digue, celle-ci n'est d'aucune utilité (Figure 7.9). Ainsi le Tableau 7.9 montre que la hauteur optimale décroît de 5.69 m à 1.19 m quand la période de retour augmente de 10 à 1000 ans.

Figure 7.9 : Sensibilité du risque bayésien analytique à la position de l'enjeu



Couloir EPA 30 de la commune de Bessans, Savoie, France. Modèle décisionnel analytique. $a'_\lambda = 31.001$, $b'_\lambda = 28.01$, $a'_p = 1208.7$ et $b'_p = 28.01$, $C_o = 5530 \text{ €/m}$, $C_1 = 300\,000 \text{ €}$, $\alpha = 0.1376$, $h_o = 1 \text{ m}$ et $A_{ct} = 25$.

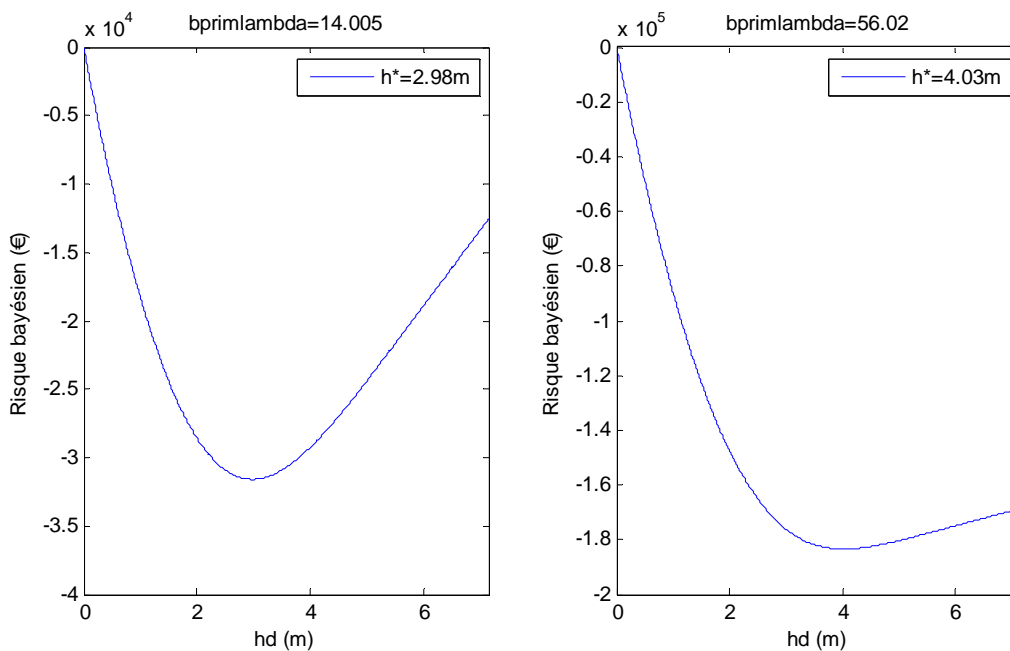
Le bénéfice moyen attendu à l'optimum économique est très grand pour des périodes de retour "faibles", par exemple 782 000 € pour un bâtiment situé à une abscisse décennale (Tableau 7.10). En effet, beaucoup d'avalanches qui auraient détruit le bâtiment sans la digue sont alors stoppées par la digue avant l'abscisse du bâtiment grâce à la réduction de distance d'arrêt. Mais la hauteur optimale est à cette abscisse très mal définie avec un bénéfice quasi constant pour $5m \leq h_d < 7.14m$. Au contraire, pour un bâtiment situé à l'abscisse millénaire, la fonction de risque est positive sur une large gamme de hauteurs, indiquant que la construction est alors économiquement inefficace. La fonction de risque est toutefois négative autour de l'optimum, mais avec un bénéfice moyen attendu très faible d'environ 3000 € qui questionne fortement le principe même d'une construction.

Notons également qu'une différence importante entre les optimums classiques et bayésiens de même qu'entre les bénéfices moyens attendus classiques et bayésiens existe quelle que soit la position de l'enjeu, avec toujours $h_B^* \geq h_C^*$ et $b_B^* \geq b_C^*$. De plus, ces différences augmentent avec la période de retour, par exemple de 5% à 250% pour la hauteur optimale (Tableau 7.9) et de 8.7% à plus de 1250% pour le bénéfice moyen attendu (Tableau 7.10) pour des positions d'enjeu correspondant à des périodes de retour allant de 10 à 1000 ans. Prendre en compte l'erreur d'estimation pour la décision favorise donc un dimensionnement de l'ouvrage plus prudent tout particulièrement pour des enjeux affectés par les événements les plus extrêmes seulement. Un tel résultat est naturel étant donné que l'erreur d'estimation affecte en particulier les quantiles les plus élevés de la distribution des aléas, rendant ainsi plus plausibles les événements les plus extrêmes.

7.2.3.2. Fréquence des avalanches

Le terme de dommages de la fonction de risque est proportionnel à la fréquence avalancheuse de dépassement. De ce fait, la hauteur optimale augmente avec la fréquence avalancheuse de dépassement. De plus, le bénéfice moyen attendu lui est presque proportionnel pour de faibles hauteurs de digue. Ainsi, la hauteur optimale diminue à 2.98 m tandis que le bénéfice correspondant est presque divisé par deux si la fréquence avalancheuse de dépassement passe de 0.9 à 0.45 avalanches par an. Au contraire, la hauteur optimale augmente jusqu'à 4.03 m et le bénéfice attendu est presque doublé si la fréquence avalancheuse de dépassement augmente de 0.9 à 1.8 avalanches par an. (Figure 7.10)

Figure 7.10 : Sensibilité du risque bayésien analytique à la fréquence moyenne de dépassement



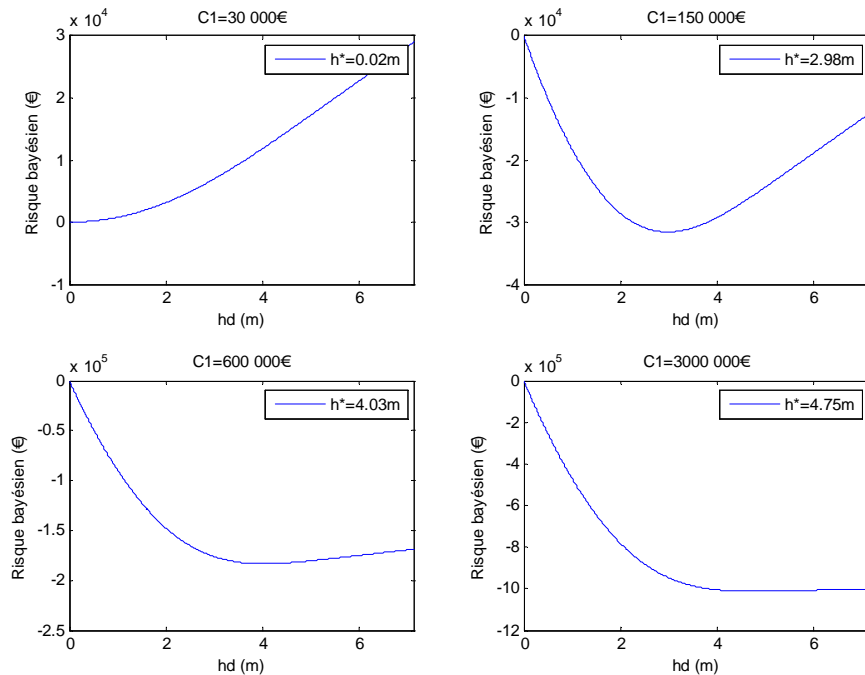
Couloir EPA 30 de la commune de Bessans, Savoie, France. Modèle décisionnel analytique. $a'_\lambda = 31.001$, $a'_\rho = 1208.7$ et $b'_\rho = 28.01$. $x_b = x_{stop100}$, $C_o = 5530$ €/m, $C_1 = 300\,000$ €, $\alpha = 0.1376$, $h_o = 1$ m et $A_{ct} = 25$.

7.2.3.3. Rapport entre coût de construction et coût de destruction

L'effet des coûts sur la fonction de risque peut être étudié au travers du rapport entre le prix de la construction et le coût des dommages actualisés, soit $\frac{A_{ct} \times C_1}{C_o}$. Il existe, comme on pouvait s'y attendre, un rapport minimal pour que l'optimum existe. La Figure 7.11 montre, par exemple, que lorsque le rapport est divisé par dix en fixant la valeur C_1 de l'enjeu à 30 000 € au lieu de 300 000 €, ce rapport minimal est presque atteint, de sorte que la construction n'est alors économiquement que peu justifiée.

Au contraire, pour des rapports très élevés, la hauteur optimale n'augmente que très doucement avec le rapport. Cela indique que la digue est déjà suffisamment haute pour stopper presque toutes les avalanches avant l'enjeu exposé. La Figure 7.11 montre ainsi qu'avec un rapport multiplié par 10 en fixant la valeur C_1 de l'enjeu à 3 000 000 €, la hauteur optimale n'est augmentée que de 0.72 m de plus que si le rapport est multiplié par deux en fixant C_1 à 600 000 €. En outre, pour des valeurs aussi élevées du rapport, la hauteur optimale est à nouveau très mal définie avec une fonction de risque présentant une allure identique à celle observée en Figure 7.9 pour un bâtiment situé à une abscisse décennale. Entre les deux cas extrêmes d'un rapport très faible ou très fort, la hauteur optimale est par contre bien définie et augmente avec le rapport.

Figure 7.11 : Sensibilité du risque bayésien analytique au coût de destruction



Couloir EPA 30 de la commune de Bessans, Savoie, France. Modèle décisionnel analytique. $a'_\lambda = 31.001$, $b'_\lambda = 28.01$, $a'_\rho = 1208.7$ et $b'_\rho = 28.01$. $x_b = x_{stop_{100}}$, $C_o = 5530 \text{ €/m}$, $\alpha = 0.1376$, $h_o = 1 \text{ m}$ et $A_{ct} = 25$.

7.2.3.4. Pouvoir dissipatif de la digue

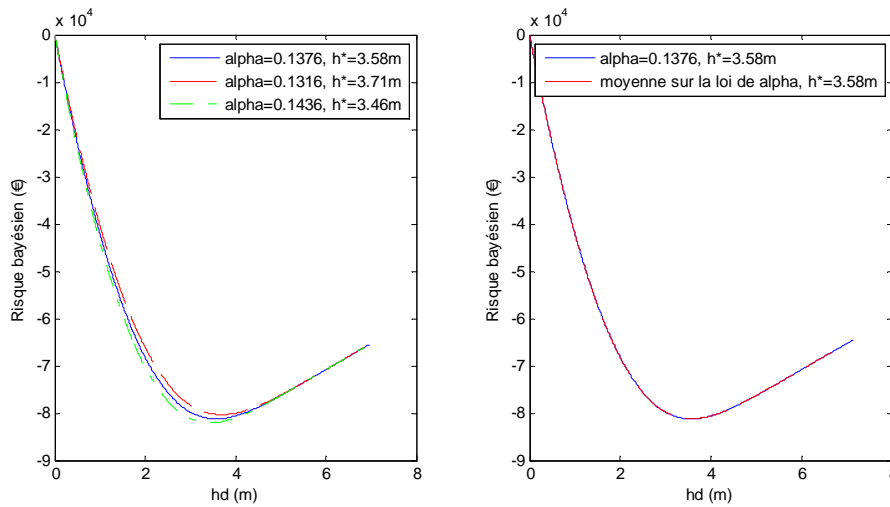
Jusqu'à présent, le paramètre α a été considéré comme parfaitement connu et de ce fait fixé à sa valeur expérimentale. Une première possibilité pour analyser la sensibilité du risque bayésien au choix de α est de remplacer l'estimateur ponctuel α précédemment utilisé par les valeurs correspondant aux bornes inférieures et supérieures de son intervalle de crédibilité à 95% obtenu à partir de l'analyse des données expérimentales présentée en Figure 3.4. La Figure 7.12 (gauche) montre que les fonctions de risque ainsi obtenues ne sont pas très différentes de celles calculées sachant α . Néanmoins, une

certaine sensibilité à la valeur de α existe, en particulier à proximité de la hauteur optimale, celle-ci étant de 3.71 m au lieu de 3.46 m si la valeur la plus pessimiste pour α est retenue.

Une seconde possibilité est de traiter α comme un paramètre du modèle statistique et de calculer un « vrai » risque bayésien en moyennant le risque précédemment obtenu sur sa loi *a posteriori*. Sous l'hypothèse d'indépendance de α par rapport aux paramètres ρ, λ du modèle d'aléa, la « vraie » fonction de risque bayésien $R_B'(h_d)$ est alors obtenue d'après l'éq. 7.17 où $p(\alpha|data)$ est la loi *a posteriori* de α obtenue à partir de l'analyse des données expérimentales de la Figure 3.4. Cette dernière est gaussienne avec un modèle conjugué normal-Gamma inverse pour la régression linéaire. L'intégration doit cependant être effectuée numériquement. La Figure 7.12 (droite) montre que $R_B'(h_d)$, et $R_B(h_d)$ sont quasiment identiques, ce qui indique une très faible sensibilité sur ce cas d'étude à l'erreur d'estimation concernant le pouvoir dissipatif de la digue.

$$R_B'(h_d) = E_\alpha [R_B(h_d)] = \int R_B(h_d) \times p(\alpha|data) \times d\alpha \quad (\text{eq. 7.17})$$

Figure 7.12 : Sensibilité du risque bayésien analytique au paramètre α .



Couloir EPA 30 de la commune de Bessans, Savoie, France. Modèle décisionnel analytique. $a'_\lambda = 31.001$, $b'_\lambda = 28.01$, $a'_\rho = 1208.7$ et $b'_\rho = 28.01$. $x_b = x_{stop100}$, $C_o = 5530 \text{ €/m}$, $C_1 = 300\,000 \text{ €}$, $h_o = 1 \text{ m}$ et $A_{ct} = 25$.

7.2.4. Vers un modèle multivarié

7.2.4.1. Un modèle d'aléa simpliste

- Influence de la topographie

Pour des variables physiques comme un débit de crue ou le volume de dépôt d'une avalanche étudié au chapitre 7.1, l'utilisation des modèles POT est justifiée par la théorie des valeurs extrêmes. Pour les distances d'arrêt des avalanches, la convergence asymptotique n'est plus valable du fait de la dépendance de la distribution des distances d'arrêt vis-à-vis de la topographie (Meunier et Ancey, 2004). Néanmoins, le modèle du rapport d'arrêt présenté au chapitre 2 est d'un usage assez courant dans le domaine des avalanches. Les dépassements du « point beta » correspondant au début de la zone d'arrêt y sont normalisés par une grandeur dépendant de la topographie du couloir et modélisés par une loi de Gumbel (McClung et Lied, 1987) ou une loi GEV (Keylock, 2005). Le choix du seuil n'est, dans ce cas, pas basé sur des considérations théoriques (Coles, 2001). Le modèle statistique est alors considéré comme un modèle empirique acceptable plutôt que comme un modèle théorique limite. De plus, il n'est appliqué que sur des couloirs très réguliers de type scandinave pour lesquels l'hypothèse que la probabilité de dépassement d'une abscisse donnée ne dépend que de la distance au point beta est tenable. Ce point de vue pragmatique est également adopté avec ce premier modèle décisionnel puisque le seuil choisi, c'est-à-dire l'abscisse de la digue, correspond aussi au début de la

zone d'arrêt et que le modèle n'est appliqué que sur des couloirs très réguliers. L'hypothèse de l'indépendance de la distribution des distances d'arrêt vis à vis du profil de pente se traduit du point de vue décisionnel par le fait que la fonction de risque ne dépend que de la distance entre l'enjeu et la digue.

- *Monovarié contre multivarié*

Au delà du problème de la dépendance topographique, le choix d'un modèle POT monovarié pour décrire l'intensité des avalanches peut bien évidemment être critiqué parce qu'il ne propose pas une vision multivariée de l'aléa. La hauteur en écoulement n'est, en particulier, pas modélisée, et ce alors qu'elle est nécessaire à la quantification de la réduction de la distance d'arrêt par la digue. Elle doit donc être gardée constante d'une avalanche à l'autre, ce qui sous estime nécessairement la variabilité des aléas sur le site étudié.

Un cadre de travail monovarié très similaire pourrait être construit en reprenant le modèle POT appliqué aux volumes dans l'exemple de Tacconnaz, en quantifiant l'effet de la digue en termes de volume stocké (Faug, 2004), et en exprimant les dommages causés aux enjeux en fonction du volume excédant la digue. Un tel modèle décisionnel, souvent utilisé en hydrologie (Parent et Bernier, 2007), s'intéresserait à une variable quantitative importante pour le dimensionnement d'un paravalanche mais ne tiendrait pas compte de la distance entre la digue et les enjeux, ce qui n'est évidemment pas approprié pour un fluide à la rhéologie complexe comme la neige. Il semble donc indispensable pour progresser de représenter l'aléa avalancheux par un modèle d'aléa multivarié du type de ceux présentés au chapitre 4.

7.2.4.2. Une évaluation des coûts rudimentaire

- *Une valeur économique globale*

La quantification de la vulnérabilité sur un cas d'étude réel implique, au minimum, le recensement de tous les bâtiments concernés, du mobilier qu'ils incluent et bien sûr du nombre d'habitants permanents ou temporaires qu'ils abritent. Plus généralement, l'ensemble des enjeux sociaux et environnementaux peuvent également être évalués. L'ensemble de ce travail complexe nécessite, comme dit au chapitre 2, une approche interdisciplinaire (Fuchs et al., 2004 ; Fuchs et Bründ, 2005 ; Keiler, 2004. Celui-ci n'a pas du tout été engagé dans la cadre de cette thèse pour laquelle les enjeux ont été limités à un seul bâtiment situé à une abscisse ponctuelle. En particulier, le problème critique de l'évaluation du coût de la vie humaine n'a pas été abordé. L'approche proposée se limite donc au cas où tous les enjeux situés sur un site sont proches les uns des autres de façon à pouvoir les ramener à une abscisse ponctuelle, et à une analyse de type assurantielle où l'on accepte de quantifier le prix d'une vie humaine.

Aller plus loin nécessiterait donc à la fois d'inclure dans le modèle décisionnel une analyse de la vulnérabilité plus approfondie et de trouver une manière plus satisfaisante de traiter le problème des coûts humains. La piste islandaise exprimant le risque comme une probabilité annuelle de décès constitue à ce titre une piste intéressante mais insuffisante pour traiter à la fois des enjeux humains et matériels.

- *Modélisation des coûts*

De même, la formulation des coûts proposée peut légitimement être critiquée puisque les hypothèses d'un coût de construction linéaire avec l'obstacle, de l'absence de terme traduisant l'endommagement de la digue et d'une destruction totale de l'enjeu peuvent toutes trois être remises en cause.

Pour le coût de construction, le modèle pourrait être facilement amélioré en étudiant différents projets d'ingénierie de façon à en extraire une formulation empirique de l'augmentation du coût avec la hauteur de l'ouvrage. Un terme de dommage traduisant le coût d'entretien annuel de la digue pourrait également être ajouté. Cependant, outre le fait qu'il est normalement faible par rapport aux autres dommages si le dimensionnement structural de l'ouvrage est bien effectué, il peut plus simplement être considéré comme inclus dans le terme C_o moyennant la définition d'une période d'amortissement et un calcul d'actualisation analogue à celui effectué pour les dommages causés aux enjeux. Enfin, en ce qui concerne le terme de dommages, il pourrait bien entendu être affiné pour prendre en compte la dépendance du dommage avec la vitesse et/ou la pression (Jonasson et al., 1999 ; Barbolini et al., 2004b). Malheureusement, une telle amélioration se heurte au caractère monovarié du modèle d'aléa

qui ne donne aucune information concernant la vitesse et/ou la pression au niveau des enjeux. De plus, un dommage en marche d'escalier permet de contourner la difficulté à effectuer le lien entre intensité de l'avalanche et niveau des dommages. Une telle formulation peut donc être considérée comme une réponse à l'état actuel de méconnaissance tout en constituant le pendant naturel aux hypothèses très fortes utilisées pour le modèle d'aléa. Aller plus loin en termes de quantification des dommages, notamment sous la forme d'une relation de vulnérabilité dépendant de l'intensité des avalanches, n'est donc intéressant qu'en association avec un modèle d'aléa multivarié. C'est justement l'objet du paragraphe suivant

7.3. Un modèle d'optimisation numérique

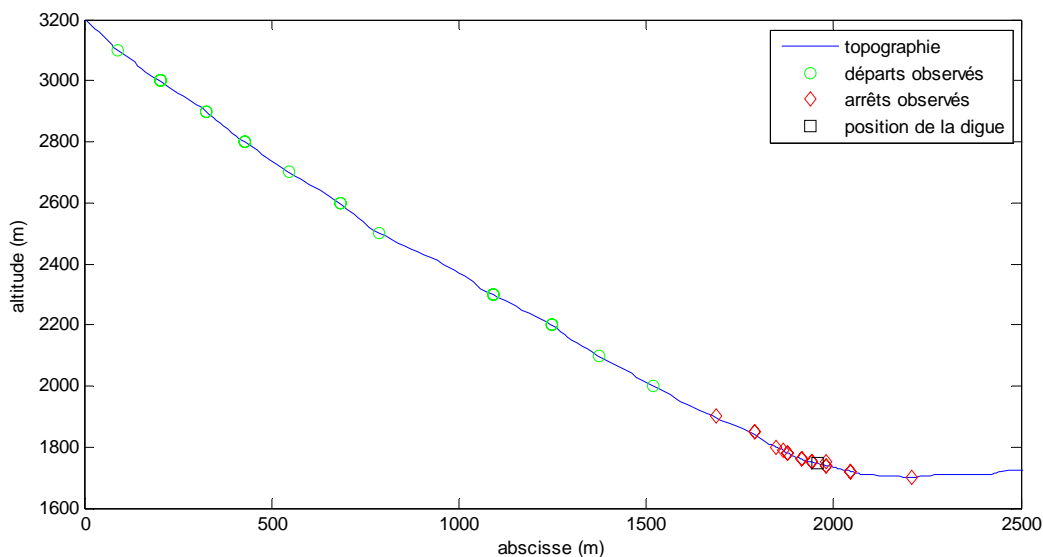
7.3.1. Modèle et cas d'étude choisi

On reprend à présent l'un des cas d'étude traité précédemment, à savoir le couloir EPA numéro 13 de Bessans, et on utilise comme modèle d'aléa multivarié le modèle Voellmy bloc glissant dont l'inférence a été effectuée au chapitre 4.2 et qui a permis la simulation des aléas de référence au chapitre 6 sur ce couloir.

On suppose, comme pour le modèle analytique, qu'un bâtiment potentiellement exposé existe en zone d'arrêt à une abscisse correspondant à une période de retour allant de 10 à 1000 ans. Pour sa protection, on envisage la construction d'une digue verticale en début de zone d'arrêt, soit à une abscisse $x_d=1956.5$ m (Figure 7.13). On conserve le même modèle d'effet de la digue sur l'écoulement que pour le modèle analytique, mais on fixe cette fois ci et une fois pour toute $\alpha = 0.14$, valeur la plus couramment retenue en ingénierie.

Le modèle de coût utilisé est celui de l'eq. 3.37, c'est-à-dire que l'on suppose entre deux seuils v_m et v_M un coût des dommages linéaire avec la vitesse de l'avalanche. Comme pour le modèle analytique, on fixe $C_o=5530$ €/m, $C_1=300\,000$ € et $A_{ct}=25$. On rappelle également que « conformément » à la directive suisse (Salm et al., 1990), on fixe $v_m = 3.3$ m/s et $v_M = 10$ m/s.

Figure 7.13 : Cas d'étude pour le modèle décisionnel complet



Couloir EPA 13 de la commune de Bessans, Savoie, France.

7.3.2. Distribution des dépassements

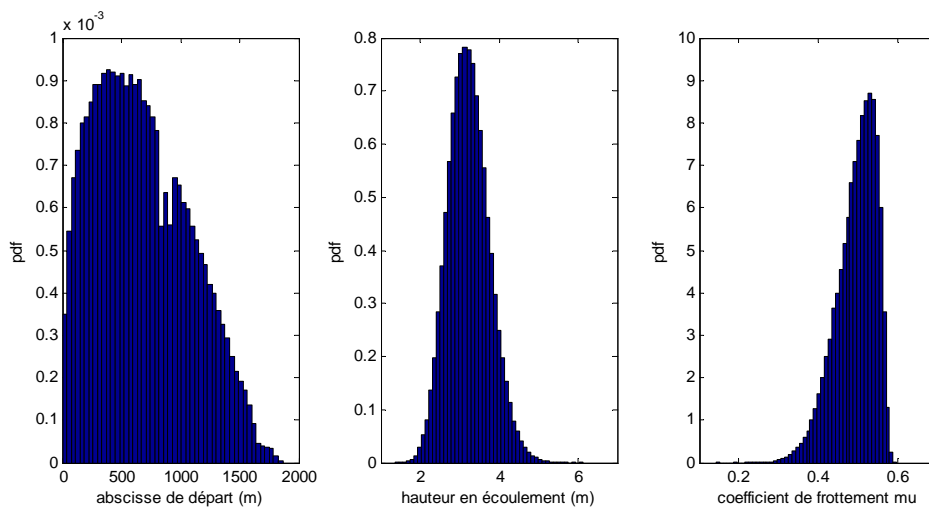
A partir de la distribution moyenne *a posteriori* des distances d'arrêt présentée au chapitre 6, la distribution moyenne des dépassements $(x_{start}, h, \mu_{voel}, x_{stop}, v_x | x_{stop} \geq x_d)$ été extraite en ne considérant que les événements dont la distance d'arrêt excède l'abscisse de la digue. On obtient ainsi la distribution de l'intensité des avalanches de référence sur lesquelles on va optimiser la hauteur de la

digue. La fréquence moyenne de dépassement s'obtient en pondérant la fréquence avalancheuse moyenne par la probabilité de dépassement moyenne (eq. 7.18). Elle s'élève à 0.506 avalanches par an, ce qui indique que l'abscisse de la digue correspond presque à une période de retour de 2 ans.

$$\hat{E}_d [a | \theta_F, \theta_M] = \hat{\lambda} \times \left(1 - \hat{F}_{xstop}(x_d) \right) \quad (\text{eq. 7.18})$$

La Figure 7.14 présente la distribution des variables d'entrée conditionnellement au dépassement. La plupart des avalanches dépassant l'abscisse de la digue se déclenchent donc dans les pentes supérieures du couloir. Les hauteurs des écoulements de référence correspondantes s'échelonnent de 2 à 5 m, et leur distribution présente une légère asymétrie à droite. Les coefficients de frottement μ_{voel} des avalanches dépassant l'abscisse de la digue s'échelonnent entre 0.2 et 0.6 et leur distribution présente une très forte asymétrie à gauche. Par rapport aux distributions moyennes *a posteriori* présentées au chapitre 4, seule la distribution des coefficients de frottement est ainsi notablement modifiée, ce qui traduit les corrélations du Tableau 4.23. En effet, les corrélations étant très faibles entre d'une part l'abscisse de départ et l'abscisse d'arrêt et d'autre part la hauteur en écoulement et l'abscisse d'arrêt, la distribution des abscisses de départ et des hauteurs en écoulement conditionnellement au dépassement est logiquement très proche de la moyenne *a posteriori* du modèle. Au contraire, le coefficient μ_{voel} est fortement corrélé à la distance d'arrêt, de sorte que la distribution du coefficient conditionnellement au dépassement est très différente de la moyenne *a posteriori*. Plus précisément, la forte asymétrie de la distribution de μ_{voel} conditionnellement au dépassement traduit le fait que seules les avalanches possédant un coefficient de frottement faible sont susceptibles d'atteindre de grandes distances d'arrêt. Cela avait déjà été observé au chapitre 6 avec l'analyse des distributions de coefficient de frottement correspondant aux avalanches de grande période de retour.

Figure 7.14 : Distribution des variables d'entrée des avalanches de référence



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant. Les simulations sont effectuées conditionnellement à la moyenne *a posteriori* des paramètres du modèle d'aléa. Seuls les dépassements de l'abscisse de la digue sont considérés

La distribution tronquée des distances d'arrêt (Figure 7.18, haut) correspond à la queue de la distribution des distances d'arrêt étudiée au chapitre 6.1. Elle reflète la géométrie de la zone d'arrêt avec de nombreuses avalanches s'arrêtant lorsque la pente locale décroît, en particulier pour $x > 2080$ m, et au contraire moins d'avalanches s'arrêtant lorsque la pente augmente légèrement, par exemple autour de $x = 1975$ m. Plus précisément, cette distribution tronquée se présente comme un mélange de plusieurs lois exponentielles et seule son extrémité correspondant à $x > 2080$ m aurait pu être modélisée par une seule loi exponentielle. Le modèle analytique présenté au paragraphe précédent ne peut donc pas être utilisé avec l'abscisse de digue choisie mais pourrait éventuellement être employé pour un ouvrage de protection situé à une abscisse $x_d > 2080$ m, et de ce fait atteint seulement par des avalanches extrêmes.

Pour chaque dépassement, l'effet de n'importe quelle hauteur de digue sur la distance d'arrêt peut être calculé comme pour le modèle analytique à l'aide de l'équation 3.30. De même, son effet sur la vitesse au niveau de n'importe quel enjeu situé en zone d'arrêt peut être calculé analytiquement en calculant la dissipation ponctuelle à l'abscisse de la digue grâce à l'eq. 3.31 puis en propageant la vitesse ainsi obtenue sur la pente supposée constante entre l'abscisse de la digue et l'enjeu considéré grâce à l'eq. 3.32. Rappelons qu'au chapitre 6.2, les distributions des vitesses à l'abscisse bisannuelle, c'est à dire celle de la digue, ont été étudiées et qu'elles s'échelonnent entre 0 et 40 m/s, avec une moyenne de 17 m/s.

7.3.3. Calcul des fonctions de risque et dimensionnement optimal

7.3.3.1. Temps de calcul

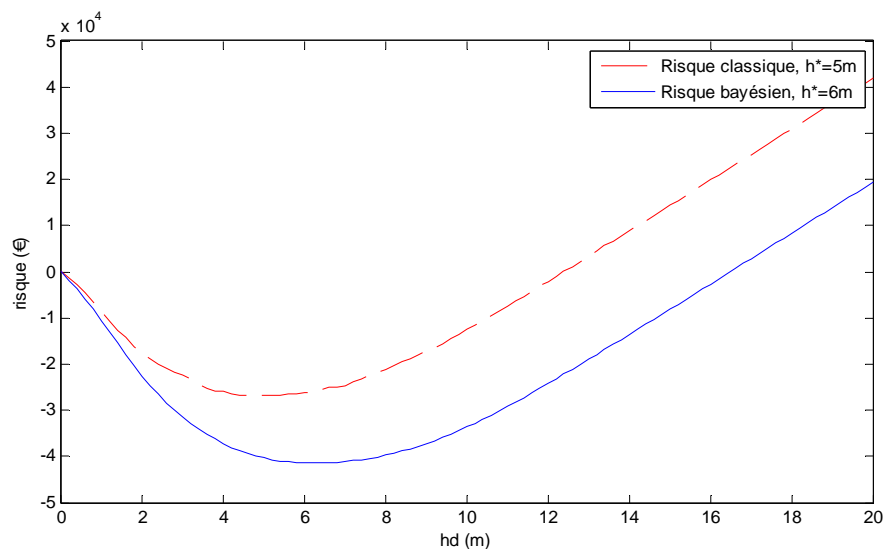
Conformément aux remarques faites au chapitre 3.6 concernant le compromis à trouver entre temps de calcul et précision de l'approximation numérique, la hauteur de digue a , pour le cas d'étude, été discrétisée à 0.2 m près, avec une hauteur maximale de 20 m qui correspond à la hauteur maximale pour laquelle l'eq. 3.20 est valide avec un écoulement de référence d'une hauteur de 2.8 m. Pour le risque classique, effectuer 300 000 itérations du modèle de propagation et calculer l'utilité associée à chaque hauteur de digue pour cinq positions d'enjeu conformément à la méthode décrite au chapitre 3.6 a nécessité environ 3 heures. Pour le risque bayésien, 5000 valeurs de la loi *a posteriori* de θ_M, θ_F ont été considérées, et pour chaque couple de paramètres, 5000 avalanches ont été simulées, permettant le calcul de la fonction de risque pour cinq positions d'enjeux conformément à la méthode décrite au chapitre 3.6. Le temps de calcul total s'est élevé à un jour environ.

7.3.3.2. Fonctions de risque pour une distance d'arrêt centennale

La Figure 7.15 présente les deux fonctions de risque classique et bayésien obtenues pour un enjeu situé à une abscisse centennale. Elles sont toutes deux relativement régulières, indiquant qu'un nombre suffisant de hauteurs de digue a été utilisé et, plus généralement, que l'erreur de Monte Carlo est faible. De plus, comme avec le modèle analytique, elles présentent chacune un optimum net, 5 m et 6 m respectivement, et un incrément linéaire avec la hauteur de la digue pour des hauteurs de digue élevées.

Le risque classique est négatif pour des hauteurs de digue inférieures à 13 m, ce qui, comme attendu, indique que la construction de la digue apporte un bénéfice économique sauf si une digue trop haute est construite. Plus précisément, le bénéfice moyen attendu croît de 0, valeur obtenue pour l'état de référence sans digue, jusqu'à 26 863 € à la hauteur optimale puis décroît à nouveau.

Figure 7.15 : Risques numériques pour un bâtiment situé à une abscisse centennale



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle décisionnel numérique. $C_o = 5530 \text{ €/m}$,

$$C_1 = 300\,000 \text{ € et } A_{ct} = 25.$$

Le risque bayésien est, comme pour le modèle analytique, relativement proche du risque classique en termes de forme et de valeur. L'optimum bayésien est cependant supérieur de 20% à l'optimum classique (Tableau 7.11). De plus, le bénéfice moyen attendu est également nettement plus élevé avec une différence relative de 54% (Tableau 7.12). Ainsi, si la procédure de dimensionnement bayésienne est choisie plutôt que la procédure classique, un dimensionnement plus prudent est préconisé et un bénéfice moyen plus important est attendu. Comme pour le modèle analytique, la prise en compte de la pénurie d'information locale semble donc peser sur le processus décisionnel. Les 26 avalanches utilisées pour l'inférence conduisent à des lois *a posteriori* encore suffisamment dispersées pour rendre délicate l'évaluation de l'intensité des avalanches les plus extrêmes et influencer ainsi la décision dans le sens de la sécurité. Une autre explication possible et vraisemblablement combinée de l'écart entre les résultats fournis par les procédures classique et bayésienne est l'information *a priori* utilisée pour l'inférence. Un *prior* assez informatif a été utilisé, en particulier pour le coefficient de frottement turbulent, et il plaide peut être aussi pour une décision plus prudente que si seules les données avaient été utilisées.

La principale différence avec le modèle analytique est que les deux hauteurs optimales sont bien plus élevées. C'est dû aux hauteurs bien plus importantes des avalanches de référence considérées, celles-ci allant de 2 à 5 m dans ce cas alors que la hauteur de référence était fixée à 1 m avec le modèle analytique. Par contre, la différence relative entre les deux hauteurs optimales est du même ordre de grandeur dans ce cas (20%) qu'avec le modèle analytique (18.5%).

7.3.4. Analyse de sensibilité

7.3.4.1. Position de l'enjeu

Une étude paramétrique des fonctions de risque très similaire à celle effectuée avec le modèle simplifié peut être effectuée avec ce modèle moyennant un effort de calcul substantiel. Comme la quasi totalité des résultats sont très similaires, seule l'effet de la position des enjeux, point critique du point de vue opérationnel, est ici investigué.

Pour des périodes de retour allant de 10 à 300 ans, la hauteur optimale (Tableau 7.11) et le bénéfice moyen attendu (Tableau 7.12) décroissent lorsque l'abscisse de l'enjeu augmente. L'optimum bayésien passe ainsi de 14.2 m à 2.8 m, tandis que le bénéfice moyen correspondant diminue de 522 965 € à 4247 €. Ces résultats sont très naturels : comme pour le modèle simplifié, une digue haute est très avantageuse pour de faibles périodes de retour car, sans digue, de nombreuses avalanches atteignent l'enjeu, de sorte qu'un bénéfice important est apporté par une digue haute. Au contraire, pour de très grandes périodes de retour, très peu d'avalanches atteignent l'enjeu même sans digue, de sorte que seule une petite digue est économiquement justifiable. Pour un enjeu situé à une abscisse millénale, la hauteur optimale n'existe même plus, indiquant que la construction d'un ouvrage n'a alors plus aucune efficacité économique.

Tableau 7.11 : Hauteurs optimales avec le modèle décisionnel numérique

	T ₁₀	T ₃₀	T ₁₀₀	T ₃₀₀	T ₁₀₀₀
h_C^* (m)	13.2	8.6	5.0	0.6	0.0
h_B^* (m)	14.2	9.8	6.0	2.8	0
$\delta_{h^*} = h_B^* - h_C^*$ (m)	1.0	1.2	1.0	2.2	0
δ_{h^*} (%)	8	14	20	367	/

Aussi longtemps que l'optimum existe, une différence importante entre les optimums classique et bayésien existe, de même qu'entre les bénéfices moyens attendus. De plus, ces différences augmentent avec la période de retour, par exemple de 8% à 366% pour la hauteur optimale (Tableau 7.11) et de 5% à plus de 900% pour le bénéfice moyen attendu (Tableau 7.12) pour des positions d'enjeu correspondant à des périodes de retour allant de 10 à 300 ans. On retrouve bien, comme avec le modèle analytique, que prendre en compte l'erreur d'estimation pour la décision favorise un dimensionnement de l'ouvrage plus prudent tout particulièrement pour des enjeux affectés par les événements les plus extrêmes seulement. Enfin, quelle que soit la période de retour, les hauteurs optimales obtenues sont toujours plus hautes qu'avec le modèle analytique du fait de la hauteur plus

élevée des événements de référence. Les bénéfices attendus sont cependant moins élevés du fait d'une formulation du coût des dommages moins pessimiste. On y reviendra au paragraphe 7.3.4.3.

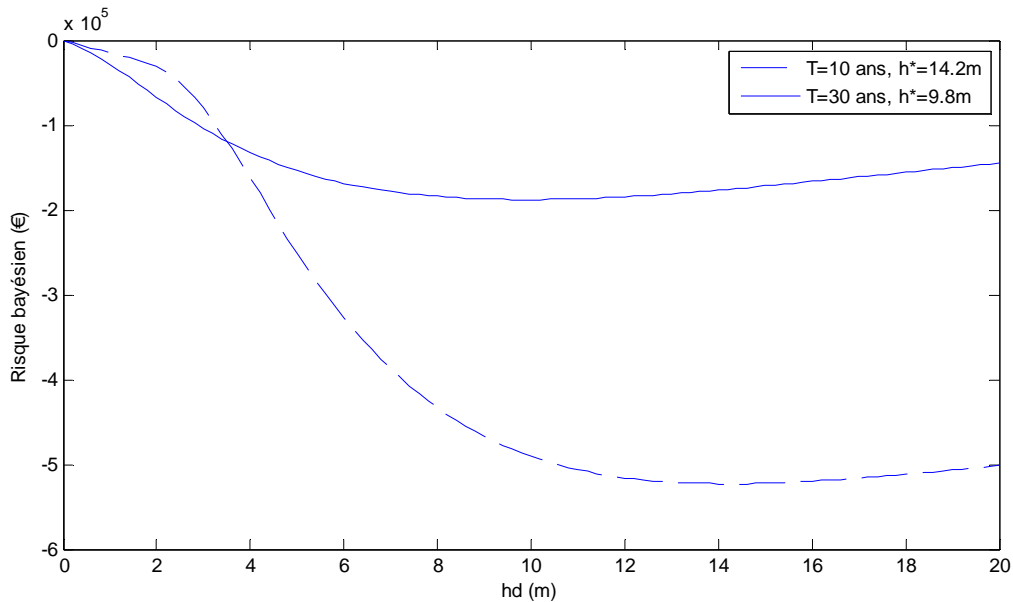
Tableau 7.12 : Bénéfices associés aux hauteurs optimales avec le modèle décisionnel numérique

	T ₁₀	T ₃₀	T ₁₀₀	T ₃₀₀	T ₁₀₀₀
b_C^* (€)	497 824	163 499	28 863	422	0
b_B^* (€)	522 965	187 868	41 465	4247	0
$\delta_{b^*} = b_C^* - b_B^*$ (€)	25 141	24 369	14 602	3824	0
δ_{b^*} (%)	5	15	54	905	/

7.3.4.2. Topographie de la zone d'arrêt

Dans le modèle analytique présenté au paragraphe précédent, les fonctions de risque dépendent uniquement de la distance entre la digue et l'enjeu et ont une forme analogue quelle que soit cette distance. Avec ce modèle, au contraire, la topographie de la zone d'arrêt influence directement la distribution des distances d'arrêt et des vitesses en chaque point du profil. La sensibilité des fonctions de risque à la topographie en zone d'arrêt peut donc être investiguée.

Figure 7.16 : Risques bayésiens numériques pour des abscisses décennale et trentennale



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle décisionnel numérique. $C_o = 5530$ €/m,

$$C_1 = 300\,000 \text{ € et } A_{cr} = 25.$$

La Figure 7.16 montre par exemple que, pour un enjeu situé à une abscisse décennale, le risque bayésien présente une forme surprenante. En effet, le bénéfice moyen attendu augmente très doucement avec la hauteur de la digue pour de faibles hauteurs de digue puis très fortement au voisinage de la hauteur optimale. Il peut même être noté que pour de faibles hauteurs de digue, le bénéfice moyen attendu est plus faible pour un enjeu situé à une abscisse décennale que pour un enjeu situé à une abscisse trentennale. L'explication est que la pente moyenne entre la digue et l'abscisse décennale est assez forte ($\phi_b = 16.3$ deg). La plupart des avalanches qui atteignent l'abscisse décennale sans digue ont de ce fait assez d'énergie pour l'atteindre également avec une petite digue, et ce sans perdre beaucoup de vitesse. Une petite digue protège donc très mal un enjeu situé à une abscisse décennale et apporte donc un bénéfice faible. Pour des enjeux situés à des abscisses correspondant à des périodes de retour plus élevées comme 30 ans, la pente moyenne de la zone d'arrêt est nettement plus faible, de sorte que même une petite digue apporte une protection significative, rendant ainsi la fonction de risque bien plus régulière. Au final, l'influence de la

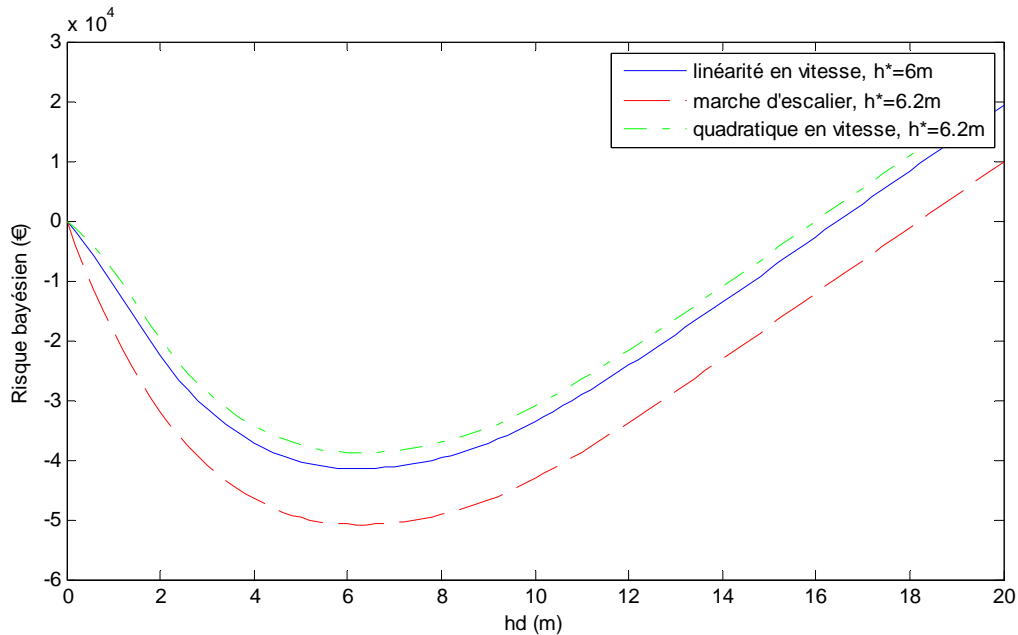
topographie sur la fonction de risque est toutefois surtout nette pour des hauteurs de digue bien inférieures à la hauteur optimale, elle ne semble donc pas beaucoup influencer le choix de la valeur de projet.

Au delà du problème de la forme de la fonction de risque, ce résultat questionne l'hypothèse de considérer la pente constante entre la digue et l'enjeu. En particulier, le ralentissement de la dynamique de l'écoulement au voisinage de l'abscisse trentennale mis en lumière au chapitre 6.5 et conduisant à une augmentation locale du C_x n'est ainsi pas pris en compte. Néanmoins, cette hypothèse est nécessaire à la validité de la formule utilisée pour la dissipation d'énergie au voisinage de la digue. Elle doit donc être effectuée sous peine de devoir reconsidérer la quantification des effets de l'obstacle. La méthode proposée ne peut donc être utilisée que sur des couloirs présentant une géométrie de zone d'arrêt assez simple.

7.3.4.3. Dépendance en vitesse du coût des dommages

Pour l'optimisation, une fonction de coût linéaire en vitesse a jusqu'ici été utilisée. Dans les analyses de risque proposées par les auteurs norvégiens, italiens et islandais citées au chapitre 2, des formulations de vulnérabilité ajustées sur des événements catastrophiques en Islande et Autriche et dépendant du carré de la vitesse ont toutefois été préférées. En outre, au paragraphe précédent présentant le modèle analytique simplifié, une fonction de coût encore plus rudimentaire de type marche d'escalier a été utilisée. La sensibilité du dimensionnement optimal au choix de la fonction de coût a de ce fait été étudiée en refaisant les calculs avec la fonction de coût en marche d'escalier et avec une fonction de coût dépendant quadratiquement de la vitesse entre les mêmes seuils $v_m = 3.3 \text{ m/s}$ et $v_M = 10 \text{ m/s}$.

Figure 7.17 : Sensibilité du risque bayésien à la dépendance en vitesse du coût de destruction



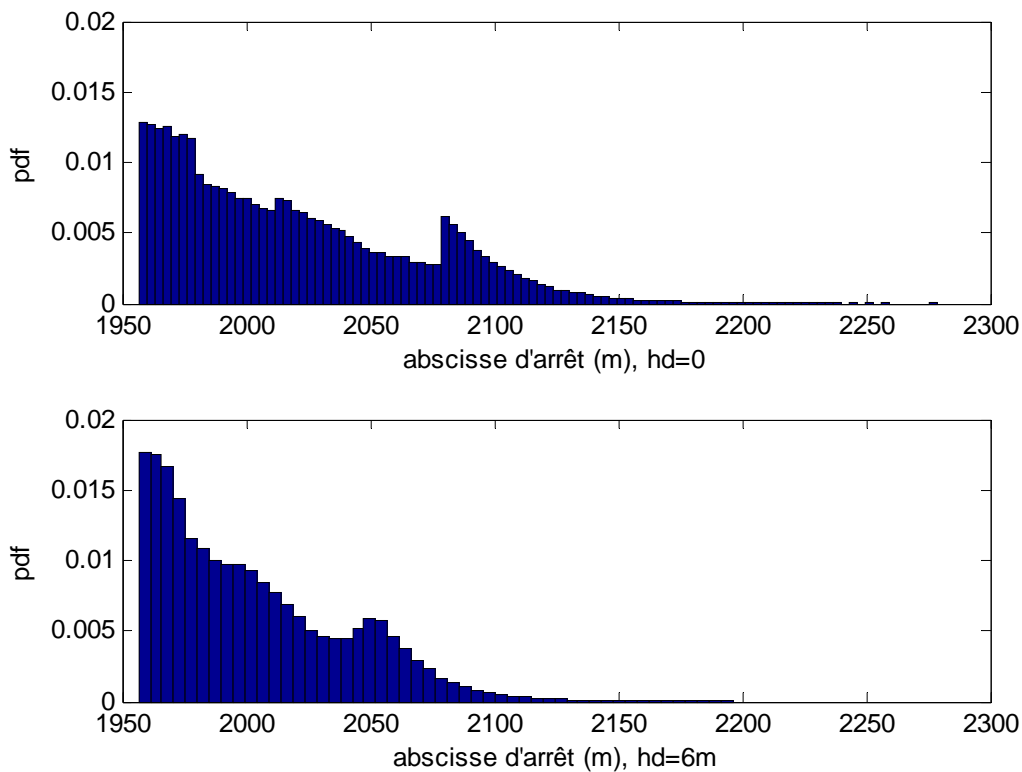
Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle décisionnel numérique. L'enjeu est situé à une abscisse centennale.

La Figure 7.17 montre une très faible différence entre les 3 fonctions de coût en termes de hauteur optimale puisque l'écart entre les valeurs obtenues avec les trois fonctions de coût correspond à peine à 0.2 m, soit le pas de discrétisation utilisé. La différence est par contre plus importante du point de vue du bénéfice moyen attendu. En effet, avec le coût en marche d'escalier, la digue apporte un bénéfice maximal égal à la valeur de l'enjeu dès lors qu'elle réduit la distance d'arrêt suffisamment pour que ce dernier ne soit pas atteint alors qu'il aurait dû l'être sans la digue. Cela explique qu'avec cette fonction de coût, une hauteur optimale non nulle avait pu être obtenue à une abscisse millénale sur le couloir 30 de Bessans.

Pour les deux autres fonctions de coût, la digue n'apporte un bénéfice égal à la valeur de l'enjeu que pour des avalanches ayant des vitesses au niveau de l'enjeu supérieures à v_M sans la digue et inférieures à v_m avec la digue. Pour toutes les autres avalanches, le bénéfice dépend de la réduction de vitesse au niveau de l'enjeu mais reste inférieur à la valeur de l'enjeu. Le bénéfice moyen attendu de la construction de la digue, en particulier de la digue optimale, est donc plus petit avec les fonctions de coût dépendant de la vitesse qu'avec la fonction de coût en marche d'escalier. Avec la fonction de coût quadratique en vitesse, le bénéfice moyen attendu est même un peu plus faible qu'avec la fonction de coût linéaire en vitesse parce que pour une vitesse v_b donnée, le coût est plus élevé avec le coût linéaire qu'avec le coût quadratique.

Au final, ces résultats indiquent que la détermination de la hauteur de digue optimale n'est pas très sensible au choix de l'une des trois fonctions de coût. Cela suggère que si seule la hauteur optimale est évaluée, une formulation de coût très fruste peut suffire. La sensibilité du bénéfice attendu au choix de la fonction de coût indique en revanche que si une estimation économique précise de l'efficacité de la digue est nécessaire, le choix de la fonction de coût doit être bien réfléchi. La fonction en marche d'escalier permet une quantification rapide du bénéfice maximal à attendre de l'ouvrage, tandis que l'une des deux autres fonctions de coût permet d'obtenir une valeur de bénéfice plus faible mais vraisemblablement plus réaliste.

Figure 7.18 : Distribution des distances d'arrêt sans et avec ouvrage



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant. Les paramètres du modèle sont fixés à leur moyenne a posteriori. Seules les avalanches dépassant l'abscisse de la digue sont considérées. La hauteur de la digue correspond à l'optimum bayésien.

7.3.5. Risque résiduel

7.3.5.1. Efficacité économique et risque résiduel

Quand la hauteur optimale a été déterminée, il est nécessaire de vérifier l'efficacité de l'ouvrage choisi afin de tester si l'optimum économique correspond à un niveau de protection raisonnable. En particulier, si la hauteur optimale est faible, son choix peut légitimement être contesté par les populations locales qui préféreront sans doute payer un peu plus et être mieux protégés même si ce

n'est pas tout à fait économiquement justifié. Pour quantifier le risque résiduel sur le cas d'étude traité, la moyenne *a posteriori* de la distribution des dépassements a été recalculée avec une hauteur de digue de 6 m correspondant à l'optimum bayésien pour un enjeu situé à une abscisse centennale.

7.3.5.2. Augmentation de la période de retour de la distance d'arrêt

La Figure 7.18 montre que la distribution des dépassements présente une asymétrie à droite bien moins marquée avec la digue que sans. Seules quelques avalanches dépassent en effet l'abscisse de 2100 m avec la digue. Cette impression visuelle est confirmée si la période de retour associée à chaque abscisse est recalculée avec la digue (Tableau 7.13). Pour une abscisse donnée, la période de retour est en effet très significativement augmentée avec la digue. L'abscisse correspondant à une période de retour de 100 ans sans la digue correspond par exemple à une période de retour de 1006 années avec la digue. Ce rapport de 10 entre la période de retour avec digue et la période de retour de référence suggère que l'ouvrage de protection possède une efficacité très satisfaisante. Le Tableau 7.13 montre en outre que ce rapport augmente avec la période de retour, puisqu'il passe de 3.4 à 64.4 pour des abscisses correspondant à des périodes de retour de référence allant de 10 à 1000 ans. Cela indique que la digue est particulièrement efficace pour protéger les enjeux situés très loin en zone d'arrêt et atteints par les événements les plus extrêmes seulement.

Tableau 7.13 : Augmentation de la période de retour avec la hauteur d'obstacle optimale

x_{stop} (m)	2067.4	2101.4	2130.5	2158.0	2188.3
T_1 (années), sans digue	10	30	100	300	1000
T_2 (années), $h_d = 6m$	33.8	203.4	1005.9	4440.0	64 379.4
T_2/T_1	3.4	6.8	10.1	14.8	64.4

7.3.5.3. Réduction de la vitesse au niveau de l'enjeu

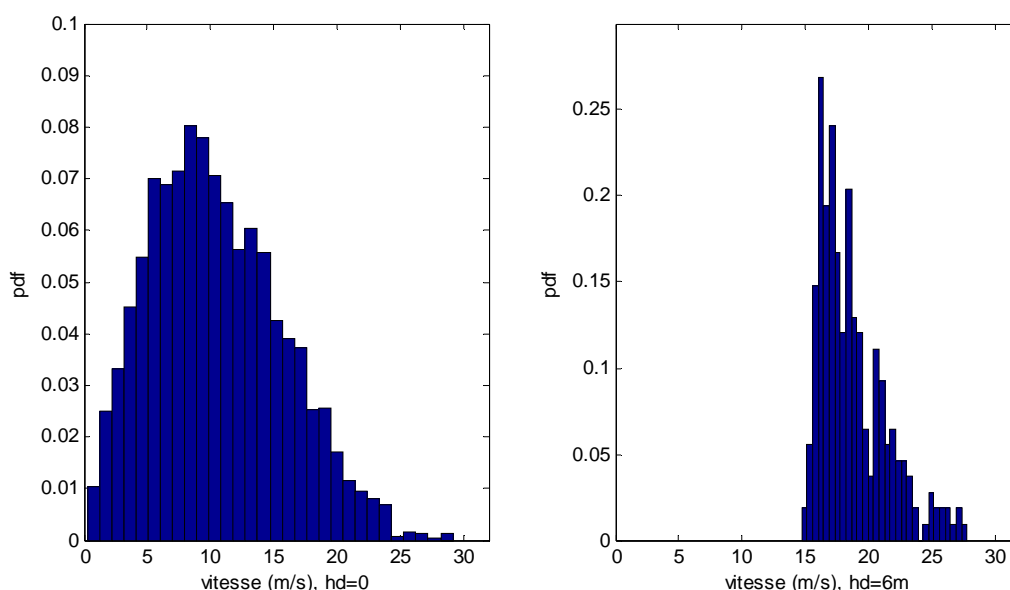
L'impression de grande efficacité de la digue déduite de l'analyse de la distribution des distances d'arrêt avec et sans digue doit toutefois être nuancée par l'analyse des distributions des vitesses avec et sans digue. La Figure 7.19 présente ainsi la distribution des vitesses des avalanches atteignant l'abscisse centennale avec et sans la digue. La distribution sans digue, déjà analysée au chapitre 6, est rappelée pour faciliter la comparaison. Avec la digue, la distribution est plus irrégulière parce que le nombre d'événements considérés a été divisé par 10, ce qui augmente significativement l'erreur de Monte Carlo. Néanmoins, ce qui apparaît clairement, c'est qu'avec la digue, toutes les vitesses à l'abscisse centennale sont comprises entre 15 et 28 m/s, de sorte que la distribution avec digue correspond quasiment à la queue de la distribution sans digue. Cela indique que la digue est très efficace pour arrêter avant l'abscisse centennale les avalanches qui atteignent cette abscisse avec une vitesse faible sans digue mais qu'elle est incapable de beaucoup freiner les avalanches atteignant l'abscisse centennale avec de fortes vitesses.

L'explication est que la pente moyenne entre la digue et l'abscisse centennale est encore non négligeable ($\phi_b = 12.7$ deg). Les avalanches possédant un coefficient de frottement μ_{voel} inférieur à $\tan^{-1}(\phi_b) = 0.225$ ne décélèrent donc pas entre la digue et l'enjeu. En outre, bien qu'elles perdent de l'énergie au niveau de la digue, le tronçon de pente est suffisamment long pour qu'au contraire elles réaccélèrent, de sorte que la vitesse limite prévue par le modèle de Voellmy est atteinte au niveau de l'enjeu avec ou sans la digue. Quant aux avalanches ayant un coefficient de frottement μ_{voel} légèrement supérieur à 0.225, elles perdent un peu d'énergie au niveau de la digue mais ne sont ensuite que très peu freinées, de sorte qu'elles atteignent l'enjeu avec une vitesse encore forte. La Figure 6.7 montre d'ailleurs que des valeurs de μ_{voel} inférieures à 0.3 sont effectivement assez fréquentes pour les avalanches atteignant l'abscisse centennale sans digue sur ce couloir.

Ce résultat questionne à nouveau fortement l'hypothèse d'une zone d'arrêt de pente constante. Mais il suggère également que la digue optimale proposée est sans doute insuffisante pour empêcher la destruction totale par les avalanches les plus extrêmes des enjeux situés très loin en zone d'arrêt. En d'autres termes, la digue de projet semble très efficace en «fonctionnement courant» en protégeant les enjeux contre la plupart des avalanches susceptibles de les affecter mais insuffisante pour prévenir les

événements les plus catastrophiques pour lesquels un autre ouvrage de protection devrait sans doute être envisagé, par exemple une seconde petite digue située très près des enjeux.

Figure 7.19 : Distribution des vitesses à une abscisse centennale sans et avec ouvrage



Couloir EPA 13 de la commune de Bessans, Savoie, France. Modèle Voellmy bloc glissant. La distance d'arrêt est centennale sans obstacle mais millénaire avec l'obstacle

7.3.6. Hypothèses résiduelles et développements futurs

7.3.6.1. Modèle d'aléa

- D'un modèle décisionnel analytique à un modèle numérique

Avec le modèle décisionnel du chapitre 7.2, tous les calculs peuvent être effectués analytiquement, à l'exception de l'optimisation des fonctions de risque. De plus, la fonction de coût en marche d'escalier permet une quantification aisée du bénéfice maximal attendu et une étude paramétrique peut être conduite moyennant un effort de calcul très réduit. Ce modèle décisionnel est de ce fait bien adapté à une première investigation d'un couloir avalancheux. Malheureusement, ce modèle décisionnel ne peut être utilisé que pour un couloir très régulier du fait de la dépendance des distances d'arrêt vis-à-vis du profil de pente. De plus, la vision monovariée de l'aléa rend difficile une amélioration du modèle par une quantification plus réaliste des effets de l'obstacle et des dommages.

Passer à un modèle multivarié d'aléa incluant un modèle de propagation numérique comme détaillé au chapitre 7.3 évite d'avoir à proposer un modèle statistique pour les distances d'arrêt, de sorte que la vraie topographie peut être prise en compte dans le calcul. De plus, les effets de l'obstacle peuvent alors être calculés à l'aide de la hauteur de l'écoulement correspondant à l'événement considéré et une quantification des dommages faisant intervenir la vitesse au niveau de l'enjeu peut être utilisée. Evidemment, les temps de calcul sont alors nettement augmentés, tout particulièrement pour l'obtention de la fonction de risque bayésienne et si la sensibilité aux différents facteurs est investiguée.

Bien sûr, comme le modèle de Voellmy bloc glissant reste une approximation assez grossière de l'écoulement avalancheux, le choix de ce modèle d'aléa peut malgré tout être légitimement critiqué. D'autant qu'il a été montré au chapitre 6 que par rapport à un modèle fluide il avait tendance à surestimer les distributions des vitesses aux abscisses correspondant à de grandes périodes de retour. Adopter une approche totalement similaire avec le modèle de Saint-Venant proposant notamment une hauteur d'écoulement réaliste au niveau de l'abscisse de la digue plutôt qu'une hauteur de bloc semble donc une piste prometteuse. Mais les temps de calcul s'en trouveront alors encore fortement augmentés. En outre, il conviendra de voir comment la relation proposée pour la réduction de la distance d'arrêt peut être adaptée à un écoulement dont la hauteur au niveau de la digue varie au cours du temps.

- *Influence de l'ouvrage*

Avec un modèle décisionnel, la modélisation de l'aléa ne se limite pas à l'écoulement. Une représentation réaliste des effets de l'ouvrage de protection sur l'écoulement doit également être proposée. C'est d'autant plus difficile pour la neige que la perturbation locale au niveau de l'ouvrage est difficile à propager jusqu'aux enjeux du fait de la faible connaissance de la rhéologie du matériau. Dans les deux modèles décisionnels proposés, une formulation analytique complète des effets de l'obstacle a été utilisée : réduction des distances d'arrêt uniquement pour le modèle analytique et réduction des distance d'arrêt et effet sur le profil de vitesse pour le modèle numérique multivarié, le tout avec à chaque fois une hypothèse de pente constante entre la digue et les enjeux considérés. Ce modèle d'influence de l'ouvrage sur la distance d'arrêt et le profil de vitesse a été obtenu théoriquement par une approximation au premier ordre de la dissipation d'énergie et vérifié en laboratoire à l'aide d'expériences sur modèles réduits avec des écoulements rapides de matériaux granulaires non cohésifs impactant une digue verticale. Néanmoins, au delà du seul problème de la sensibilité au paramètre α , il peut également être remis en cause.

Du point de vue physique, son caractère un peu fruste est discutable et sa validité est limitée au cas d'une digue verticale impactée par des écoulements rapides non cohésifs, c'est-à-dire par des avalanches denses de neige sèche. Afin d'aller plus loin, les effets de l'obstacle sur un écoulement quelconque pourraient être calculés numériquement pour chaque hauteur d'obstacle (Naaim et al., 2004) de façon à quantifier de manière plus exacte la perturbation locale et sa propagation. L'hypothèse de pente constante en zone d'arrêt pourrait ainsi être abandonnée. Outre son caractère moins élégant, le modèle numérique ainsi obtenu deviendrait beaucoup plus lourd du point de vue du temps de calcul. Mais il ne fait pas de doute qu'avec l'augmentation de la puissance de calcul des ordinateurs, cela sera bientôt une voie de progrès toute tracée.

D'un point de vue plus pratique, le génie paravalanche ne se limite pas, comme l'a montré l'exemple de Tacconnaz, au cas relativement simple des digues verticales impactées par des avalanches denses de neige sèche. Cette première approche mériterait de ce fait d'être approfondie pour prendre en compte une gamme plus large d'aléas (avalanches denses de neige humide, avalanche aérosol, etc.) et différents types d'ouvrages de protection (dignes déviateurs, tas freineurs, etc.), de façon à proposer différents outils pour répondre aux différentes situations rencontrées en contexte opérationnel.

7.3.6.2. *Evaluation des coûts*

Le passage à un modèle multivarié a permis d'inclure dans le modèle décisionnel une formulation des dommages analogue à celles proposées dans la littérature. Malheureusement, elle reste très fruste, notamment parce que, comme on l'a vu au chapitre 6, la pression d'impact d'une avalanche est très difficile à calculer. Améliorer le modèle décisionnel du point de vue de l'évaluation des coûts de destruction nécessite donc d'attendre des progrès dans le domaine de la quantification des dommages subis par une structure impactée par une avalanche. Néanmoins, étant donné la faible sensibilité au choix de la fonction de coût du dimensionnement optimal obtenue au chapitre 7.3.4.3, ce n'est pas la priorité première pour le développement des approches décisionnelles dans le domaine des avalanches.

7.3.6.3. *Variable de décision*

L'intérêt de combiner différentes digues pour protéger différents enjeux a été rapidement évoqué au chapitre 7.3.5 du fait de l'inefficacité de la digue à protéger les enjeux situés très loin en zone d'arrêt contre les avalanches les plus extrêmes. De même, pour traiter un cas de figure aussi complexe que celui de Tacconnaz présenté en introduction de ce chapitre, différents ouvrages de protection doivent être combinés afin d'offrir une protection contre les différents types d'aléa susceptibles de se produire sur ce site. Ces deux exemples soulignent les limites d'une variable de décision monodimensionnelle et plaident pour aborder le problème du dimensionnement d'ouvrages paravalanches dans un cadre décisionnel multivarié. En particulier, une première amélioration des modèles décisionnels proposés consisterait à optimiser simultanément la hauteur et la position de la digue projetée plutôt que de fixer d'emblée son abscisse au début de la zone d'arrêt.

7.4. Conclusion du chapitre 7

Ce chapitre a traité du problème du dimensionnement d'un ouvrage paravalanche. Après une introduction illustrant la difficulté à travailler avec un aléa de référence, les calculs ont été menés dans

un cadre décisionnel, tout d'abord de façon analytique à l'aide d'un modèle simplifié puis numériquement avec le modèle d'aléa Voellmy bloc glissant et une fonction de coût dépendant de la vitesse au point d'impact. Le modèle décisionnel analytique qui fournit une estimation rapide de la perte maximale attendue a permis la mise en œuvre d'une étude paramétrique portant sur l'intensité et la fréquence de l'aléa, le rapport entre coûts de construction et de destruction ainsi que le pouvoir dissipatif de la digue. Le modèle numérique qui offre quant à lui une estimation plus précise du bénéfice apporté par l'ouvrage a permis de mettre en évidence l'importance de la topographie de la zone d'arrêt sur l'allure des fonctions de risque et de quantifier le risque résiduel après construction de l'ouvrage. Pour les deux modèles, le caractère systématiquement pessimiste de l'optimum bayésien prenant en compte l'erreur d'estimation due à l'insuffisance de données locales a été mis en évidence. Etant donné l'absence par ailleurs d'autres approches décisionnelles en génie paravalanche, ce chapitre reste néanmoins exploratoire. L'étude a de ce fait été restreinte au problème de l'optimisation de la hauteur d'une digue paravalanche impactée par des avalanches denses. Des hypothèses simplificatrices très fortes ont en outre été effectuées, notamment pour la modélisation des coûts. Il reste donc beaucoup de place pour des développements futurs de cette thématique.

8. Synthèse, conclusion et perspectives

8.1. Bilan des résultats obtenus

8.1.1. Construction d'un cadre de travail intégrateur rigoureux

8.1.1.1. Bibliographie, clarification et transposition

Cette thèse a été effectuée à l'interface entre plusieurs disciplines : nivologie, modélisation déterministe des aléas naturels, hydrologie stochastique ainsi que statistique bayésienne. Les références intéressantes étaient donc très nombreuses et différents axes ont été explorés : connaissances de base concernant la neige et les avalanches, modèles de propagation des avalanches, modélisation bayésienne hiérarchique... Bien que la matière n'ait évidemment pas été épuisée, un large tour d'horizon synthétisé au chapitre 2 a été effectué. Il a permis de clarifier les concepts essentiels à la modélisation stochastique des avalanches, en référence notamment au domaine de l'hydrologie.

Ponctuellement, des domaines plus inattendus ont également été source d'inspiration, telle l'épidémiologie statistique mise à profit au chapitre 5 pour le développement d'une analyse régionale des fréquences avalancheuses. Cette thèse a ainsi contribué à montrer que la discipline très jeune qu'est la modélisation stochastique des avalanches pouvait se nourrir de la connaissance développée dans d'autres domaines, même apparemment très éloignés au premier abord.

8.1.1.2. Assemblage des connaissances grâce à la modélisation stochastique

Comme cela a été rappelé au chapitre 2, les méthodes traditionnelles de prédétermination des avalanches sont plutôt de nature déterministe. Elles définissent des aléas de référence en termes probabilisés de manière un peu artificielle, voire abusive. Une partie importante du travail de thèse a donc consisté en la redéfinition plus rigoureuse d'un cadre de travail stochastique général intégrant les différents types de connaissances et permettant le zonage et le dimensionnement d'ouvrages de protection à l'échelle d'un site avalancheux de l'Enquête Permanente sur les Avalanches.

Dans ce cadre, proposé au chapitre 3 et utilisé concrètement dans les chapitres 4 et 7, la connaissance naturaliste est intervenue classiquement sous la forme de la description topographique du site mais aussi sous la forme de lois *a priori* informatives. La connaissance de la physique des avalanches a été intégrée grâce à un modèle numérique de propagation. La donnée historique de terrain a permis l'apprentissage local de façon à compenser l'incapacité du modèle numérique à traduire sans calibration la réalité de l'ensemble des processus physiques. Le liant a été apporté par le modèle stochastique décrivant la variabilité du phénomène à l'échelle du site étudié. Une fois inféré, il a permis la prédiction des aléas de référence recherchés. Il y a donc une forte complémentarité entre l'expertise naturaliste, la modélisation numérique de la dynamique des écoulements et la collecte des données avalancheuses. Cette thèse a contribué à montrer que la prédétermination est plus rigoureuse si l'ensemble des connaissances est assemblé grâce à la modélisation stochastique.

8.1.2. Construction d'outils de prédétermination locale

8.1.2.1. Des hypothèses de modélisation

A partir du cadre général de travail, différentes hypothèses simplificatrices ont été nécessaires pour la mise en œuvre du calcul des aléas de référence. Elles ont concerné la nature des variables modélisées, le choix d'un modèle de propagation et d'opérateurs stochastiques décrivant la variabilité du phénomène ainsi que les *priors*. La fréquence avalancheuse a toujours été modélisée par une approche poissonnienne simple proposée au chapitre 5.1. Par contre, l'intensité de l'avalanche a été modélisée au chapitre 4 par un nombre de plus en plus important de variables correspondant à trois modèles locaux de complexité croissante et permettant une description de plus en plus réaliste de la propagation. Pour les trois modèles proposés, l'éventuelle corrélation existant entre les différentes variables a été explicitement prise en compte. Les lois *a priori* nécessaires à la mise en œuvre de l'inférence

bayésienne ont été obtenues soit par élicitation, soit grâce à des sites analogues permettant l'apprentissage, soit en choisissant des *priors* très peu informatifs.

8.1.2.2. *Un modèle très simplifié*

Un modèle très simplifié dit « de Coulomb » a été proposé au chapitre 4.1. Il possède la particularité d'être inversible, c'est-à-dire que les coefficients de frottement correspondant aux avalanches observées peuvent être obtenus analytiquement de façon déterministe. Pour ce modèle, le problème de l'inférence statistique est donc déconnecté de la propagation, ce qui réduit considérablement les difficultés techniques liées à son utilisation. Les distributions de probabilité de l'ensemble des paramètres du modèle ainsi que les distributions prédictives de la distance d'arrêt et de n'importe quel quantile de cette variable utilisée pour le zonage ont de ce fait pu être obtenues moyennant un effort de calcul réduit, de même que n'importe quelle période de retour. En outre, ce modèle s'est révélé bien adapté du fait de sa simplicité à la comparaison de différentes structures de modèle et à des études de sensibilité. D'un point de vue méthodologique, le modèle de Coulomb a donc été utilisé comme un modèle de base dont l'inférence et l'utilisation prédictive étaient faciles. Il illustre bien la capacité à obtenir des résultats utilisables à partir du cadre général proposé au chapitre 3.

En revanche, il est apparu que pour des géométries de sites irrégulières, le modèle de Coulomb ne pouvait être utilisé car il était incapable de reproduire correctement les observations. Une variante de ce modèle avec un terme d'erreur et une vitesse initiale non nulle a donc été proposée au chapitre 3.1.5. Ce modèle dit « de Coulomb bruité » s'est avéré capable de fonctionner correctement sur un site avalancheux à la géométrie raisonnablement complexe, mais uniquement si un nombre important de données était disponible pour la calibration. Une autre limite importante de ce modèle est que ses sorties opérationnelles sont restreintes à la distance d'arrêt car le modèle prend en compte très peu de variables et modélise les champs de vitesse de façon incorrecte. En contrepartie, il est bien adapté à la structure des données de l'EPA puisque seules les positions de départ et d'arrivée sont nécessaires à son fonctionnement.

8.1.2.3. *Deux modèles non inversibles plus réalistes*

Les parties 2 et 3 du chapitre 4 ont proposé deux modèles de prédétermination relativement similaires. Le premier d'entre eux s'articule, comme le modèle de Coulomb, autour d'un modèle de propagation de type bloc glissant, tandis que le second s'inscrit dans le cadre de la mécanique des milieux continus à l'aide du formalisme de Saint-Venant. Tous deux utilisent une loi de frottement de type Voellmy et sont par conséquent non inversibles. Ils permettent une représentation multivariée plus réaliste qu'avec le modèle de Coulomb ou son avatar bruité de l'aléa avalancheux en donnant accès à plusieurs grandeurs supplémentaires telles la hauteur et la vitesse en n'importe quelle abscisse du couloir.

Pour les deux modèles de propagation, le modèle stochastique est quasiment identique, la hauteur moyenne en zone de départ « remplaçant » simplement dans le modèle de Saint-Venant la hauteur moyenne du bloc. Le raffinement de la description est toutefois bien plus grand avec le modèle de Saint-Venant puisque ce sont les champs de vitesse et de hauteur complets qui sont modélisés plutôt que la hauteur constante et la vitesse moyenne d'un bloc.

Des algorithmes de calibration probabilistes permettant de surmonter la difficulté technique de l'inférence ont été développés. Ils sont très similaires pour les deux modèles, l'une de leurs caractéristiques communes étant la nécessité d'utiliser un *prior* relativement informatif pour le coefficient de frottement turbulent. Avec le modèle de Voellmy bloc glissant, une étude sur données simulées a permis d'en vérifier le bon fonctionnement.

Les lois *a posteriori* des différents paramètres ont pu être simulées, donnant accès à des estimateurs ponctuels, aux incertitudes associées et aux coefficients de corrélation inter paramètres. Les phénomènes de compensation entre certains couples de paramètres et entre coefficients de frottement ont pu être établis. Avec les deux modèles, de faibles corrélations entre le coefficient de frottement latent et les autres variables d'entrée ont été obtenues sur le cas d'étude traité, mais il est apparu que les estimateurs des paramètres de corrélation pouvaient être biaisés par les erreurs de discrétisation numérique et d'observation, sans toutefois affecter gravement la prédiction des distances d'arrêt de référence. L'adéquation entre données et modèles a, quant à elle, été satisfaisante, même si la distribution des variables latentes inférée s'est révélée être bien plus resserrée que la distribution moyenne *a posteriori*, celle-ci correspondant à un comportement moyen du fait de la faiblesse des corrélations inter variables.

Le temps calcul, déjà bien accru avec le modèle « Voellmy bloc glissant » par rapport au modèle de Coulomb, est devenu problématique avec le modèle de Saint-Venant. De ce fait, le calcul des lois prédictives des aléas de référence n'a été mené qu'avec le modèle de Voellmy bloc glissant. Il a néanmoins été montré que le nombre d'itérations nécessaires à l'inférence pouvait être considérablement réduit en fixant le coefficient de frottement turbulent sans, là encore, affecter gravement la prédiction des distances d'arrêt de référence.

8.1.2.4. Prédiction, zonage, dimensionnement et effet d'un obstacle

Le chapitre 6 a proposé la mise en œuvre complète du calcul des aléas de référence nécessaires au zonage et au dimensionnement structural d'un ouvrage de protection avec les trois modèles sur différents cas d'étude. Le lien quasi univoque entre période de retour et quantile de distance d'arrêt y a été mis à profit pour affecter une période de retour à chaque distance d'arrêt. L'ensemble des incertitudes a été exprimé sous la forme d'un intervalle de crédibilité quantifiant bien mieux l'erreur de prédiction qu'un intervalle de confiance non paramétrique de type erreur de Monte Carlo et permettant au décideur d'apprécier la précision des valeurs de projet qui lui sont proposées. Réciproquement, il a été montré que la distribution des périodes de retour correspondant à une abscisse donnée était également accessible.

Le modèle de Coulomb et son avatar bruité ont permis l'obtention de résultats en bon accord avec les chroniques historiques et la topographie dès lors que cette dernière restait « raisonnable ». Mais seule la distance d'arrêt de référence était alors disponible, ce qui ne répond que partiellement aux besoins du zonage et pas à ceux du dimensionnement d'ouvrages paravalanches. Les deux autres modèles ont donné accès à la distribution des autres variables d'intérêt pour une distance d'arrêt de référence donnée, autorisant ainsi, en plus du zonage, le dimensionnement structural d'un ouvrage. Sont accessibles en particulier les vitesses aux abscisses uniquement atteintes par des avalanches de période de retour élevées, ce qui permet de prévoir, au moins partiellement, les sollicitations auxquelles seraient soumis un éventuel obstacle ou un enjeu. Le calcul de la pression d'impact, grandeur la plus pertinente pour caractériser les dommages potentiels s'est toutefois révélé délicat du fait de l'absence de modélisation de la densité et de l'incertitude pesant sur le lien pression-vitesse. Néanmoins les distributions de pression obtenues ont confirmé l'importance de la prise en compte de la rhéologie à proximité immédiate de l'arrêt ainsi que la forte influence de la connaissance de la masse volumique de la neige sur les valeurs de projet.

Une analyse de sensibilité aux différentes hypothèses de modélisation a été menée. Le rôle prépondérant du coefficient de frottement latent a été mis en évidence, de même que l'importance de l'information *a priori* pour certains paramètres. En termes de distance d'arrêt, les trois modèles de prédétermination ont donné des résultats relativement cohérents sur le site où ils ont été comparés. Mais à information équivalente, le modèle de Coulomb s'est montré nettement plus pessimiste que les deux autres, donnant même des résultats irréalistes avec une faible quantité de données. Notamment, les vitesses obtenues avec le modèle de Voellmy bloc glissant se sont également révélées plus fortes à abscisse égale qu'avec le modèle de Saint-Venant. Plus généralement, un comportement physiquement plus réaliste a été observé avec le modèle de Saint-Venant, notamment en termes de valeurs de coefficients de frottement associées aux avalanches les plus extrêmes. La comparaison des résultats fournis par les différents modèles a toutefois été difficile du fait de la nécessité de transformer les observations pour les faire correspondre aux variables d'entrée modélisées et d'utiliser des lois *a priori* au moins partiellement informatives pour les deux modèles utilisant une loi de frottement « de Voellmy ». Ce dernier point rejoint le problème bien connu de la validation des événements de référence correspondant à des probabilités très faibles.

Les aléas de référence obtenus se sont montrés fortement influencés par les corrélations intervariables avec le modèle de Coulomb et son avatar bruité. Avec les deux autres modèles, au contraire, les corrélations intervariables obtenues ont été de plus en plus faibles au fur et à mesure que le modèle de propagation se complexifiait, de sorte que la distribution des aléas de référence qui a été obtenue sur le cas d'étude avec le modèle de Saint-Venant s'est révélée proche de celle correspondant à des variables d'entrée indépendantes. Il est donc apparu que ces corrélations, bien captées par le modèle complètement empirique de Coulomb, étaient moins nécessaires au bon fonctionnement d'un modèle plus physique.

Enfin, les effets d'une digue verticale ont été inclus dans le modèle d'aléa grâce à une loi semi empirique issue d'expériences sur les milieux granulaires. Les effets de la digue sur la loi des aléas ont

de ce fait pu être quantifiés au chapitre 7.3 de façon à appréhender précisément le risque résiduel après construction d'une digue de hauteur donnée. Le dimensionnement fonctionnel de l'ouvrage, c'est-à-dire son choix en fonction de l'effet protecteur recherché, peut ainsi s'effectuer non pas vis-à-vis d'un aléa de référence mais vis-à-vis de la variabilité du phénomène tout entier.

8.1.3. Du local au spatial et du stationnaire au non stationnaire

Pour le calcul de périodes de retour, la fréquence avalancheuse moyenne a été nécessaire. Etant donné que l'hypothèse d'indépendance entre intensité et fréquence des avalanches a été effectuée au cours de la thèse, la façon dont est obtenue la fréquence avalancheuse moyenne n'influence pas du tout le calcul des aléas de référence.

Une modélisation poissonnienne locale a tout d'abord été proposée pour l'ensemble des modèles de prédétermination au chapitre 5.1. Au chapitre 5.2, un élargissement spatial en a été proposé à l'aide d'un modèle inspiré de l'épidémiologie statistique et permettant d'estimer et de prédire le nombre moyen d'avalanches se produisant par couloir et par an à l'échelle communale. Le choix de l'échelle spatiale utilisée reste discutable mais il a été guidé en premier lieu par l'imprécision de la localisation des sites avalancheux pour les événements anciens. Les résultats obtenus suggèrent que prendre en compte la spatialisation de l'information est une voie à creuser dans le domaine des avalanches. La structure spatiale a en effet pu être isolée du bruit et 73% de la variabilité des occurrences avalancheuses ont pu être expliqués par le voisinage une fois les données prétraitées pour prendre en compte les observations manquantes. Il a également pu être établi que la structure spatiale des occurrences avalancheuses présente une portée relativement courte et correspond grossièrement à celle des massifs montagneux, avec toutefois des excès ou des déficits locaux. Des covariables nivométéorologiques ont pu être intégrées par krigeage dans le modèle pour en améliorer le pouvoir prédictif. Différents modèles spatiaux dérivés du modèle de base n'ont par contre pas permis d'obtenir des résultats bien meilleurs, en particulier en prédiction.

Aux chapitres 4.3 et 4.4, une extension spatio-temporelle du modèle a été développée. Les résultats obtenus ont montré une nette amélioration de l'adéquation du modèle aux données dès lors qu'une composante non stationnaire traduisant la variabilité interannuelle du climat a été introduite. Celle-ci n'a pas semblé présenter de tendance ou de rupture forte. Par contre, elle s'est révélée suffisamment marquée pour permettre la classification, avec néanmoins des niveaux de confiance faible, des hivers passés en trois régimes correspondant à des années d'activité faible moyenne ou forte. Enfin, l'ajout d'une structure spatio-temporelle du second ordre a permis de faire quasiment disparaître le bruit local et d'identifier des effets locaux dans l'espace et le temps interprétables en termes d'effort d'observation. Le modèle final obtenu a été appliqué en guise d'illustration sur trois régions des Alpes françaises, permettant de montrer que leur fonctionnement était partiellement couplé et que la variabilité du climat était un peu plus forte dans les Alpes du Sud que dans les Alpes du Nord.

Cette partie de la thèse a ainsi marqué une première inflexion majeure par rapport au sujet initial qui était celui de la prédétermination locale en contexte stationnaire. Le manque d'information locale a été compensé par la recherche de cohérence spatiale. De plus, l'hypothèse d'invariance de la distribution des nombres d'avalanches annuels a été reconsidérée de façon à prendre en compte la variabilité interannuelle du forçage climatique. Deux problématiques actuelles de la recherche sur les avalanches ont ainsi été abordées : celle, très fondamentale et pour l'instant très peu traitée, de la quantification des effets du changement climatique sur l'activité avalancheuse. Et celle, plus appliquée mais incontournable pour la gestion du risque, de la prédétermination sur des couloirs non documentés.

Cette dernière est généralement traitée par l'extrapolation de connaissances physiques à partir d'un couloir modèle (Sigurdsson et al., 1998) ou l'utilisation d'une norme régionalisée (Salm et al., 1990). La méthode d'interpolation spatiale proposée défend plutôt l'idée d'une juste pondération de l'information locale disponible par l'information apportée par le voisinage. En ce sens, elle rejoint la méthode proposée au chapitre 4.1 pour le transfert d'information à partir d'un couloir d'apprentissage au travers de la mise à jour du *prior* sur les paramètres du modèle d'intensité.

8.1.4. De l'aléa de référence à une approche décisionnelle

Dans le dimensionnement d'ouvrages de protection paravalanches, le caractère monovarié de la période de retour est particulièrement problématique car on est en général amené à rechercher une protection contre au moins deux quantités : le volume de neige et l'énergie de l'écoulement. Cette difficulté a été rencontrée de façon concrète dans le cadre de l'expertise menée sur le site de Taconnaz

rapidement présentée au chapitre 7.1 et pour laquelle une approche par modélisation POT a été utilisée. Cette expertise a entraîné une seconde inflexion marquante du sujet initial. La modélisation de l'aléa a en effet été inscrite dans un cadre décisionnel permettant le dimensionnement sous incertitude par la minimisation de la perte moyenne espérée.

D'un point de vue méthodologique, le modèle d'aléa a de ce fait été complété par une quantification des dommages prévisionnels permettant le calcul de risque. En cohérence avec le reste du travail de thèse, une attention particulière a été donnée à la répercussion de l'erreur d'estimation sur la décision au travers du calcul de la fonction de risque bayésienne. Par rapport aux autres approches économiques du dimensionnement d'ouvrages paravalanches qui sont principalement de type coût-bénéfice, cette thèse a donc introduit la prise en compte pour la décision de la variabilité du phénomène et de la rareté de l'information.

Chaque type d'ouvrage paravalanche ayant des effets particuliers sur les différents types d'écoulements avalancheux, on s'est restreint au problème de l'optimisation de la hauteur d'une digue paravalanche impactée par des avalanches denses. Les calculs ont été menés au chapitre 7, tout d'abord dans un cadre analytique simplifié puis dans un cadre plus réaliste avec le modèle stochastique d'aléa « Voellmy bloc glissant » et une fonction de coût dépendant de la vitesse au point d'impact. Le modèle décisionnel analytique proposé permet une quantification rapide de la perte maximale attendue tandis que le modèle numérique proposé permet une quantification plus précise du bénéfice apporté par l'ouvrage. Les facteurs prépondérants influençant la hauteur optimale de l'ouvrage ont été étudiés au travers d'une étude paramétrique portant sur l'intensité et la fréquence de l'aléa, le rapport entre coûts de construction et de destruction ainsi que le pouvoir dissipatif de la digue. De même, le caractère systématiquement pessimiste de l'optimum bayésien qui devrait inciter le décideur à privilégier la prudence en contexte d'information peu abondante a été établi. Les résultats obtenus ont ainsi montré le potentiel d'une approche décisionnelle du dimensionnement dans un contexte où les financements publics se doivent d'être rigoureusement justifiés.

8.1.5. Limites opérationnelles des outils développés

L'approche proposée dans le cadre de cette thèse possède un certain potentiel opérationnel. Néanmoins, outre la prudence de rigueur vis-à-vis d'outils et de méthodes jeunes et par conséquent insuffisamment validés, deux limites systématiques sont à considérer. La première est liée au caractère délicat des algorithmes de calibration développés pour les modèles de prédétermination locaux. Il est en effet apparu que le réglage des fonctions d'exploration devait se faire au cas par cas et requérait une certaine habileté de la part de l'utilisateur pour obtenir la convergence recherchée. Une utilisation de type « presse bouton » pour la prédétermination des avalanches ne peut donc pour l'instant être envisagée.

Une seconde limite encore plus évidente est liée à la qualité des données EPA dont l'insuffisance a été abondamment illustrée. Pour le modèle spatial, un traitement automatisé rudimentaire a pu être utilisé. Pour le développement des modèles de prédétermination locaux, la présélection manuelle d'un petit panel de sites a été préférée du fait de la multiplicité des sources d'erreur à considérer. Les chroniques ont donc été analysées au cas par cas sur les sites choisis afin de vérifier qu'elles étaient utilisables (cf Annexe). Un travail de recoupement entre la base informatisée et les carnets forestiers contenant les avis d'avalanche originaux a en outre été effectué pour retrouver les volumes de dépôt des avalanches anciennes. Si une telle approche est possible pour une expertise ponctuelle au prix de beaucoup de temps, elle est évidemment rédhitoire pour des applications plus larges. Une utilisation généralisée des méthodes proposées dans cette thèse à l'échelle, par exemple, de tous les sites sensibles d'une vallée ne pourra donc être entreprise qu'une fois que des travaux importants permettant de s'assurer préalablement de la fiabilité des données auront été accomplis. Pour ce faire, le développement de filtres automatisés plus complets que ceux utilisés pour le modèle spatial. Plus généralement, la démarche qualité engagée par le Cemagref en lien avec les observateurs de terrain devra être approfondie.

8.2. Avalanches et période de retour

Bien que très utilisée pour l'ensemble des risques naturels, la période de retour possède des limites qui remettent actuellement en cause son utilisation opérationnelle systématique, en particulier pour la gestion du risque d'inondation. Dans le domaine des avalanches, l'emploi de la notion de période de

retour est de plus particulièrement galvaudé, de sorte que son applicabilité à la prédétermination des avalanches mérite une discussion particulière.

8.2.1. Caractère monovarié : privilégier la distance d'arrêt ou une approche décisionnelle

Le premier obstacle à l'utilisation de la période de retour dans le domaine des avalanches est son caractère fondamentalement monovarié s'opposant à la nature multivariée du risque avalancheux. Contrairement aux problèmes hydrologiques standards pour lesquels les ingénieurs peuvent se focaliser sur le calcul d'un débit de projet, le risque avalancheux est en effet caractérisé par la distance d'arrêt, les champs de vitesse et de pression, le volume de dépôt, etc., avec une fonction de répartition jointe de type $F(x_{stop}, v_x, pr_x, v_{stop}) = P(X_{stop} \leq x_{stop}, V_x \leq v_x, PR_x \leq pr_x, V_{stop} \leq v_{stop})$. Il n'y a donc pas, comme l'a illustré l'exemple de Tacconnaz, de lien univoque entre un quantile $(x_{stop}, v_x, pr_x, v_{stop})$ et une période de retour T puisque deux événements différents $(x_{stop}, v_x, pr_x, v_{stop})$ et $(x'_{stop}, v'_x, pr'_x, v'_{stop})$ peuvent vérifier $F(x_{stop}, v_x, pr_x, v_{stop}) = F(x'_{stop}, v'_x, pr'_x, v'_{stop})$.

En outre, les variables marginales peuvent être très mal corrélées (Ancey et al., 2004), de sorte qu'une période de retour calculée sur l'une des variables marginales ne peut pas être transposée directement aux autres. La directive suisse (Salm et al., 1990) est de ce fait une méthode standardisée de calcul d'avalanches rares et extrêmes plutôt qu'une méthode de prédétermination d'avalanches de période de retour 30 et 300 ans, la période de retour calculée correspondant au cumul de neige uniquement et pas aux variables d'intérêt en zone d'arrêt.

La période de retour n'a donc du sens que pour une variable marginale et à condition d'être évaluée à partir de la loi de probabilité marginale correspondante. Pour contourner la difficulté, deux solutions ont été explorées au cours de cette thèse. La première est de purement et simplement renoncer à la notion de période de retour en faveur d'une approche décisionnelle. La seconde, également préconisée par d'autres auteurs, est de limiter la notion de période de retour à la seule distance d'arrêt et de l'évaluer à partir de sa loi de probabilité marginale $F(x_{stop}) = P(X_{stop} \leq x_{stop}, V_x \leq +\infty, PR_x \leq +\infty, V_{stop} \leq +\infty)$. Un argument pragmatique en faveur de ce choix est que la distance d'arrêt est la donnée la plus uniformément collectée et que des mesures de vitesse ou de pression ne sont en général pas disponibles pour une analyse statistique. Un argument plus théorique est que la distance d'arrêt est la variable conduisant à la période de retour la plus

pessimiste. En effet si l'on calcule la période de retour $T_1 = \frac{1}{1 - F(x_{stop})}$ associée à la distance d'arrêt

x_{stop} et, par exemple, la période de retour $T_2 = \frac{1}{1 - F(v_x)}$ associée à la vitesse v à la même abscisse,

nécessairement $T_2 \leq T_1$ parce que la distance d'arrêt x_{stop} doit être dépassée pour que l'événement $v_x > 0$ soit réalisé. En d'autres mots, pour une période de retour donnée, une zone d'aléa délimitée grâce à la distance d'arrêt inclut toutes les zones d'aléa qui peuvent être délimitées en utilisant d'autres variables. Evidemment, n'appliquer la période de retour qu'à la distance d'arrêt ne dispense cependant pas de quantifier sérieusement la distribution des autres variables nécessaires au zonage et au dimensionnement...

8.2.2. Processus à temps discret : combiner intensité et fréquence moyenne

Une autre difficulté est liée au caractère intrinsèquement discret du phénomène avalancheux. Celui-ci implique que les lois de probabilité estimées à partir des données sont conditionnelles au déclenchement plutôt qu'annuelles. Pour le calcul de périodes de retour, la combinaison de la loi de probabilité des distances d'arrêt sachant le nombre de déclenchements annuels et de la loi de probabilité du nombre de déclenchements annuels est donc nécessaire.

Dans cette thèse, plutôt que d'effectuer tous les calculs conditionnels, l'hypothèse de l'indépendance entre intensité des avalanches et nombre de déclenchements annuels a systématiquement été effectuée.

Le calcul de période de retour s'est alors ramené à une pondération de la probabilité de la distance d'arrêt considérée par la fréquence avalancheuse moyenne conformément à l'eq. 3.16. L'hypothèse d'indépendance intensité/fréquence est donc bien pratique. Mais si elle semble justifiable pour des couloirs avalancheux peu actifs, elle mériterait d'être étudiée plus sérieusement dans le cas de couloirs enregistrant de nombreux événements chaque hiver.

8.2.3. Voir la période de retour comme un outil standardisé

La période de retour a, sans parler des erreurs d'interprétations grossières qui l'assimileraient à une périodicité, une connotation forte en termes d'intervalle de récurrence moyen des événements de référence. Pourtant, l'hypothèse de stationnarité des aléas naturels sur le long terme en contexte de changement climatique semble difficilement justifiable. C'est particulièrement vrai pour les avalanches qui, même si la quantification des effets du changement climatique sur l'activité avalancheuse reste peu claire, sont directement influencées par la variabilité du climat.

A l'instar de ce qui se produit dans le domaine des inondations où les crues centennales utilisées pour le zonage sont remises en cause suite aux catastrophes répétées des années 1985-2002, des aléas avalancheux définis en termes de périodes de retour pourraient être discutés sous prétexte que les hypothèses de calcul sont trop hasardeuses pour justifier que la période de retour proposée est bien l'intervalle de temps moyen de dépassement de la valeur de projet. Le point de vue pragmatique défendu par cette thèse est que le calcul de période de retour doit être vu avant tout comme un outil standardisé permettant l'uniformisation de la prédétermination, et ce à l'instar des dimensionnement décennaux utilisés en hydrologie auxquels aucun hydrologue sérieux n'a jamais réellement prêté de signification temporelle trop précise.

8.3. L'importance de la topographie

8.3.1. Modéliser explicitement la propagation pour extrapoler

Un débit de crue caractérise entièrement la dynamique du phénomène hydraulique et est de ce fait une variable dite physique ou intrinsèque. Par conséquent, il est possible d'ajuster une loi de probabilité sur un jeu de données et d'obtenir simplement les quantiles correspondant à des périodes de retour élevées. La qualité de l'estimation dépend bien entendu du quantile considéré et de la taille du jeu de données utilisé pour l'inférence mais le pari moyen est toujours valide.

Pour les avalanches, au contraire, la distance d'arrêt dépend fortement de la topographie. Il n'est donc pas possible en général d'extrapoler la distribution empirique ou un modèle statistique simple ajusté sur les données au delà de la plus forte valeur observée (Meunier et al., 2004), et ce même si un effort théorique important est effectué pour la modélisation (Keylock, 2005). Pour s'en convaincre, il suffit d'imaginer deux couloirs dont la géométrie ne diverge que pour de très grandes périodes de retour. Si aucune avalanche extrême n'est présente dans le jeu de données, ajuster un modèle statistique sur les deux sites conduit, pour des périodes de retour élevées, aux mêmes valeurs de projet, ce qui ne correspond pas à la réalité du terrain.

Cette différence entre hydrologie stochastique et modélisation stochastique des avalanches a été bien mise en lumière au chapitre 7 en comparant les distributions des distances d'arrêt obtenues avec un modèle statistique simple de type POT et l'un des modèles stochastiques proposés. Le modèle statistique simple s'est en effet révélé adapté à l'extrémité de la queue de distribution correspondant aux avalanches les plus rares s'arrêtant sur une pente très régulière. Il a en revanche été incapable de prendre en compte les effets des changements de concavité, même modestes, sur la distribution des distances d'arrêt. S'inspirant des approches dites « statistiques dynamiques », cette thèse défend donc le point de vue que l'association explicite de la statistique et d'un modèle de propagation est nécessaire de façon à simuler la distribution des distances d'arrêt en prenant en compte explicitement la topographie.

8.3.2. Malgré tout une très forte influence

Au-delà du seul problème de l'extrapolation, les résultats obtenus avec l'ensemble des approches développées dans cette thèse vont dans le sens d'une forte influence de la topographie sur la distribution des variables d'intérêt pour le zonage. La forte sensibilité des résultats au pas de

discrétisation utilisé étant bien connue en modélisation numérique des avalanches, cela n'est toutefois guère étonnant.

Ainsi le modèle stochastique dit de Coulomb ne peut être utilisé que pour les couloirs les plus simples, ce qui est normal puisque l'inférence statistique y est déconnectée de la propagation. Pour étendre son utilisation à des couloirs un peu moins simples géométriquement, il a fallu ajouter un terme d'erreur et surtout imposer une vitesse initiale non nulle à l'avalanche pour permettre la propagation des événements caractérisés par les coefficients de frottement les plus extrêmes. Si c'est justifiable d'un point de vue opérationnel, ça l'est beaucoup moins d'un point de vue physique et plaide fortement pour l'emploi d'un des deux autres modèles de prédétermination proposés dès lors que la géométrie du couloir est un peu irrégulière.

Les deux autres modèles stochastiques dépendent néanmoins eux aussi de la topographie, ne serait-ce qu'au travers de l'existence d'une vitesse limite avec le modèle de Voellmy qui contraint fortement l'allure des distributions de vitesses. Ainsi, au chapitre 7, est-il apparu que, pour un enjeu situé loin d'un obstacle le long d'une pente faible quasi constante, la probabilité d'atteinte de l'enjeu décroissait logiquement avec l'abscisse mais pas la vitesse des événements qui l'atteignaient. De même, les distributions des hauteurs, nombres de Froude et pression se sont révélées très sensibles à la topographie locale, avec par exemple des nombres de Froude plus forts sur le cas d'étude proposé à l'abscisse centennale qu'à l'abscisse trentennale du fait d'un ralentissement de la dynamique de l'écoulement au niveau de l'abscisse trentennale. La sensibilité à la topographie est même apparue dans les calculs décisionnels puisque la forme de la fonction de risque obtenue au chapitre 7.3 s'est avérée dépendante du profil topographique dans la zone d'arrêt.

8.4. Pourquoi une approche bayésienne ?

8.4.1. Faciliter l'assemblage des connaissances

Le fonctionnement d'un couloir avalancheux étant complexe, la construction des modèles proposés s'est effectuée par l'assemblage de sous-modèles représentant les processus élémentaires de déclenchement, propagation et arrêt concourant à la genèse du phénomène que l'on cherchait à appréhender de façon globale. Ces sous-modèles simples se sont emboîtés « comme des briques de Lego » (Parent et Bernier, 2007), les distributions de probabilité des variables en sortie d'un sous-modèle constituant les distributions de probabilité utilisées en entrée par le sous-modèle suivant.

De façon technique, l'assemblage des sous-modèles s'est appuyé sur la conditionnalité : plutôt que de caractériser la loi jointe des variables, ce sont les lois conditionnelles qui ont été modélisées, avec une formulation explicite des éventuelles corrélations. Ainsi, dans les modèles de prédétermination locaux proposés, des modèles de mélange ont été employés pour décrire la variabilité de la hauteur de neige sachant la position de départ, et du coefficient de frottement latent sachant les autres variables. De même, au niveau régional, l'emploi de couches hiérarchiques a permis la modélisation de la ressemblance entre les différents sites avalancheux et le transfert d'information d'une commune à l'autre. Un rôle clé a donc été joué par la notion de variable latente, c'est-à-dire de grandeur inobservable ayant un statut de paramètre par rapport aux observables et de variable par rapport aux paramètres de la couche supérieure du modèle.

Une fois le modèle construit, le formalisme bayésien a permis l'assemblage des connaissances par la combinaison explicite des données et la connaissance naturaliste. Dans le cas des modèles d'intensité, la dynamique de la propagation de l'avalanche a également été incluse, de sorte que le théorème de Bayes a permis en outre l'incorporation de la connaissance physique dans la modélisation de la variabilité du phénomène. Enfin, lorsque certaines observations ponctuelles étaient manquantes, elles ont simplement été traitées et inférées comme des paramètres supplémentaires du modèle.

D'un point de vue plus pratique, l'assemblage a encore été facilité par l'existence des algorithmes d'estimation bayésiens permettant d'obtenir la loi jointe *a posteriori* des inconnues du problème, et ce quelle que soit la complexité du modèle considéré. Ainsi un algorithme de Metropolis-Hastings a permis la calibration des deux modèles de propagation non inversibles utilisés pour la prédétermination locale. De même, l'inférence du modèle spatial et spatio-temporel a pu être réalisée au moyen d'un algorithme de Gibbs, ce dernier s'appuyant d'ailleurs justement sur la structure conditionnelle à la base de la construction du modèle.

8.4.2. Distinguer clairement les sources d'incertitudes

Le formalisme bayésien a permis de distinguer clairement les sources d'incertitudes, en particulier la variabilité intrinsèque du phénomène et l'incertitude de connaissance dite épistémique. Cette distinction s'est incarnée dans la différence entre paramètre et variable latente : l'estimateur du paramètre tend asymptotiquement vers une valeur unique lorsque la quantité de données augmente tandis que la variable latente garde une dispersion correspondant à la variabilité du phénomène modélisé. Ainsi en est-il des coefficients de frottement caractérisant le comportement de la neige en écoulement. L'un d'entre eux, le coefficient unique μ_{coul} du modèle de Coulomb et le coefficient μ_{voel} de la loi de Voellmy, a été modélisé par une variable latente puisque le frottement varie d'un événement à l'autre en fonction de la qualité de la neige. L'autre coefficient de la loi de frottement de Voellmy a par contre été considéré comme un paramètre conformément à l'interprétation qu'en font la majorité des praticiens en termes de rugosité du couloir.

Cette distinction des sources d'incertitudes a permis une prise en compte de la pénurie d'information locale pour la prédétermination. L'EPA, même si elle constitue une source de données unique en son genre, ne recense en effet au mieux que quelques dizaines d'événements pour un couloir donné, ce qui est bien peu comparativement aux longues chroniques de débit usuellement utilisées pour caler les modèles hydrologiques. L'erreur d'estimation inhérente à cet état de fait et se traduisant par des intervalles de crédibilité assez larges a pu être répercutée sur les valeurs de projet grâce au calcul de leurs lois prédictives. Les distributions ainsi obtenues peuvent être communiquées au décideur afin qu'il puisse éventuellement faire le choix d'un niveau de sécurité en accord avec l'incertitude entourant l'aléa de référence choisi.

8.4.3. Offrir un prolongement décisionnel direct

Dans le cadre de la problématique de gestion du risque avalanche, le décideur est confronté à des choix difficiles entre les impératifs de développement économique et la sécurité des biens et des personnes. La notion d'aléa de référence n'offre, comme cela a été illustré par l'exemple de Tacconnaz, que des solutions limitées du fait de la difficulté à prendre en compte un aléa multivarié.

La seule réponse réellement satisfaisante est donc de prendre la décision qui minimise la perte moyenne attendue au regard de la connaissance disponible au moment de la prise de décision (Krysztofowicz, 1983). Une telle approche s'effectue naturellement dans le cadre bayésien en associant une fonction de coût à la loi *a posteriori* de façon à sélectionner de façon rigoureuse une décision optimale sous incertitude (Berger, 1985). Ainsi, les deux modèles proposés pour le dimensionnement d'une digue paravalanche ont montré que le problème de l'inférence pouvait ne pas être déconnecté de celui de la prise de décision de façon à répercuter l'état de (mé)connaissance locale sur la hauteur de projet.

8.4.4. Revendiquer explicitement une certaine subjectivité

Jusqu'ici, la vieille controverse classique-bayésien a été évacuée au profit de justifications essentiellement pratiques du choix bayésien, notamment l'existence d'algorithmes permettant la mise en œuvre de l'inférence. Néanmoins, la notion de probabilité subjective au cœur de la démarche bayésienne s'applique plutôt bien à la prédétermination des avalanches où les événements évalués sont si peu fréquents qu'ils peuvent difficilement être assimilés à des expériences répétables, surtout si, étant donné la variabilité du climat, on interprète la période de retour plus comme un outil de normalisation que comme un intervalle de récurrence moyen.

De même, la notion de subjectivité a été au cœur du travail de modélisation présenté dans cette thèse : choix des structures de modèles, des *priors*, formalisation des coûts..., etc. Bien évidemment, revendiquer explicitement que construire un modèle c'est, sauf à disposer de justifications mathématiques comme celles de la théorie des valeurs extrêmes, faire des choix subjectifs ne veut pas dire faire n'importe quoi. Il s'agit au contraire de prévenir l'ingénieur et le décideur que, même si la démarche qui lui a été proposée cherche à traduire le plus honnêtement possible l'état des connaissances disponibles en calibrant de manière rigoureuse un modèle sur des données, de nombreuses hypothèses discutables restent nécessaires à l'obtention des résultats opérationnels. Par conséquent les outils proposés doivent être utilisés avec prudence et réserves en n'hésitant pas, par exemple, à tester la robustesse des résultats obtenus. Pour ce faire, des analyses de sensibilité et des comparaisons de modèles ont été mises en œuvre.

8.5. Comment situer l'approche proposée ?

8.5.1. Par rapport à la théorie des valeurs extrêmes

8.5.1.1. *Un cadre de travail probabiliste commun*

La modélisation stochastique rencontre un grand succès pour la modélisation des risques, par exemple dans le domaine des assurances, de la finance ou dans le secteur de la médecine. Dans le domaine des risques naturels, en particulier celui des risques en montagne, elle a pourtant connu un développement assez faible, principalement à cause de la difficulté à recueillir des données en quantité suffisante concernant des événements par définition rares. Au milieu de ce tableau, l'hydrologie stochastique fait figure d'exception, puisque la théorie des valeurs extrêmes y est employée de façon routinière pour le zonage et le dimensionnement d'ouvrages.

L'approche proposée dans cette thèse partage avec l'hydrologie stochastique son cadre probabiliste distinguant les différentes sources d'incertitudes, son vocabulaire -en particulier la distinction observables, paramètres, variables latentes- et sa possible inscription dans une perspective décisionnelle.

8.5.1.2. *Justifications mathématiques contre connaissance physique*

En hydrologie stochastique, la théorie des valeurs extrêmes justifie le modèle par des propriétés mathématiques de convergence. En particulier, le modèle POT indique que les débits de crue les plus élevés sont distribués selon une loi de Pareto généralisée (Pickands, 1975). Les connaissances de l'hydrologue ne sont donc guère mobilisées pour la caractérisation de l'aléa, excepté pour choisir un seuil suffisamment haut pour garantir l'indépendance des pics de crue et la convergence vers le modèle limite.

Pour une approche stochastique de la prédétermination des avalanches du type de celle proposée dans cette thèse, les outils statistiques ne sont au contraire pas suffisants à cause de l'influence de la topographie sur la dynamique de l'écoulement. Ainsi, aucun « vrai » modèle ne s'impose pour encoder le comportement asymptotique et le modélisateur doit faire preuve d'imagination pour trouver une structure de modèle « intelligente » mais forcément subjective. En cela, il est aidé par une connaissance physique relativement abondante mettant à sa disposition différents modèles physiques de complexité variable. Son rôle est de trouver un bon compromis entre le réalisme physique de la description, un temps de calcul acceptable par l'ingénieur et les difficultés techniques de l'inférence. Cette thèse a de ce fait proposé trois modèles de prédétermination construits autour de modèles de propagation dont le niveau de complexité va du très faible (bloc glissant inversible) jusqu'au standard actuel de la modélisation déterministe dans les pays européens (modèle fluide avec loi de frottement de type Voellmy, Christen et al., 2002) afin de permettre de traiter différents cas de figure. Ces modèles ont été comparés afin de guider le choix de l'ingénieur intéressé par leur utilisation.

8.5.1.3. *Intensité et fréquence*

Les modèles POT ont généralisé dans la communauté hydrologique l'idée que, pour le calcul de périodes de retour, intensité et fréquence des événements doivent être modélisées. Cela n'est au contraire pas encore complètement admis dans le domaine des avalanches où des "périodes de retour" sont toujours calculées sans que le problème de l'évaluation de la fréquence avalancheuse moyenne locale ne soit abordé. De ce point de vue, l'approche proposée dans cette thèse est extrêmement proche de celle des modèles POT puisque l'activité avalancheuse est représentée par un modèle de Poisson dans l'ensemble des travaux effectués. La seule petite différence est que ce sont presque toujours l'ensemble des occurrences avalancheuses qui sont modélisées plutôt que seulement des dépassements, l'exception étant le modèle de dimensionnement de digue pour lequel on se ramène à un modèle POT en ne retenant que les événements qui dépassent l'abscisse de l'ouvrage. L'intérêt d'effectuer l'inférence sur toutes les données plutôt que seulement sur les dépassements est d'obtenir une estimation plus précise de la fréquence avalancheuse moyenne sachant que, de toute façon, la convergence asymptotique n'est pas recherchée.

8.5.2. Par rapport aux approches déterministes de type modèles hydrologiques à réservoirs

8.5.2.1. *Un cadre systémique commun mais différentes échelles de description*

La comparaison avec l'hydrologie stochastique étant insuffisante pour situer l'approche de la prédétermination des avalanches proposée dans cette thèse, une comparaison avec l'hydrologie déterministe « à réservoirs » s'impose également. La prédétermination stochastique des avalanches partage avec la modélisation hydrologique déterministe un cadre systémique commun avec une unité spatiale bien définie correspondant à un couloir avalancheux, qui au bassin versant.

Pour la prédétermination des avalanches, l'échelle spatiale de description n'est pas un vrai sujet de débat : l'avalanche est assimilée à un flux bi ou tridimensionnel dont les équations du mouvement sont résolues sur une grille de discrétisation. Dans cette thèse, une approximation bidimensionnelle a toujours été effectuée afin de réduire le temps calcul, mais l'approche proposée pourrait être menée en prenant en compte les variations de largeur de l'écoulement à l'aide d'un modèle de propagation plus élaboré (Gruber et al., 1998 ; Keylock et Barbolini, 2001). Pour la modalisation hydrologique déterministe, une variété beaucoup plus grande d'échelles d'études est couramment utilisée. En ingénierie, les modèles les plus populaires sont de type globaux : HBV (Bergström, 1995 ; Lindström et al., 1997), GR3J (Edijatno et al., 1999), UBC (Micovic et Quick, 1999)... Cependant, lorsque le bassin étudié comprend une large gamme d'altitudes et est affecté par des précipitations neigeuses significatives, un découpage en bandes d'altitude est souvent préféré (Kokkonen et al., 2001). De plus, quand une description réellement physique de l'ensemble des processus élémentaires est entreprise, une discrétisation complète sur grille tridimensionnelle est nécessaire (Abbott et al., 1986).

8.5.2.2. *Equations physiques contre équations empiriques*

En hydrologie déterministe, la distinction entre des modèles à base physiques et des modèles dits "conceptuels" est relativement claire. Les modèles conceptuels sont généralement globaux et totalement empiriques avec des opérateurs mathématiques représentant les échanges entre différents compartiments hypothétiques (sol, atmosphère, rivière..., etc.), tandis que les modèles physiques sont généralement distribués et basés sur la résolution des équations élémentaires (conservation de la masse et de l'énergie, perméabilité, etc.). Il est depuis longtemps admis dans ce domaine (Beven, 1989) que les modèles empiriques sont plus efficaces en ingénierie que les modèles physiques du fait de difficultés de calibration bien moindres. Il a même été montré que des modèles globaux très parcimonieux pouvaient être aussi efficaces que des modèles bien plus paramétrisés (Perrin et al., 2001). L'explication est que les processus hydrologiques qui interagissent pour générer les débits observés sont si nombreux et si imparfaitement maîtrisés qu'il est quasiment impossible d'intégrer de façon adéquate les équations élémentaires à l'échelle du bassin versant.

Le débat est plus ouvert dans la communauté avalancheuse. La rhéologie de la neige reste partiellement inconnue, de sorte que les coefficients de frottement doivent être calibrés. De plus, ils dépendent également du modèle considéré et ne peuvent de ce fait être considérés comme des caractéristiques uniquement physiques de l'écoulement. Du fait de ce caractère semi-empirique, Meunier et al. (2004) ont estimé que les modèles de propagation des avalanches sont de type conceptuel. Néanmoins tous les modèles de propagation usuels (Harbitz et al., 1998), depuis les blocs glissants jusqu'aux modèles fluides les plus sophistiqués, restent à base physique puisque les équations du mouvement y sont intégrées le long du chenal d'écoulement. L'explication est peut-être que les spécialistes des avalanches ont pour l'instant peu recherché à développer des modèles purement empiriques. Mais le point de vue défendu dans cette thèse est plutôt que la base physique de l'ensemble des modèles d'avalanches est liée au caractère relativement simple du phénomène étudié. Les processus physiques prépondérants sont en effet très peu nombreux, à savoir la gravité et le frottement, de sorte qu'il est bien plus simple qu'en hydrologie des bassins versants de prendre en compte la connaissance physique, même partielle, dans la modélisation. Ainsi, les trois modèles d'intensité proposés dans la thèse sont effectivement basés sur les équations du mouvement. Mais ils sont utilisés comme des modèles empiriques puisqu'un apprentissage sur données est effectué pour calibrer la loi de frottement. On est donc bien "quelque part" entre le modèle physique et le modèle purement empirique.

D'un point de vue plus pratique, la prédétermination des avalanches et l'hydrologie déterministe se rejoignent sur l'idée qu'un modèle très simple peut être suffisant pour l'ingénierie. Ainsi le modèle

“Voellmy bloc glissant” autour duquel est construit le second modèle de prédétermination proposé dans cette thèse est utilisé par les praticiens depuis 50 ans avec succès. Quant au modèle de Coulomb, il est encore plus grossier.

8.5.2.3. *Principales sources de variabilité*

La pratique courante en hydrologie déterministe voit l’ajustement des débits prévus sur les débits observés par des méthodes de type moindres carrés. Pourtant, selon Kavetski et al. (2002), un cadre de travail systémique doit traduire le fait que les variables d’entrée introduisent bien plus de variabilité dans le système que les erreurs de sortie. Les spécialistes des avalanches partagent généralement ce point de vue en considérant les caractéristiques de la zone de départ au moment du déclenchement comme la source prépondérante de variabilité (Maggioni, 2005). Les résultats obtenus au cours de cette thèse vont dans ce sens en montrant que la variabilité des distances d’arrêt dans les modèles de prédétermination est essentiellement liée à celle du coefficient de frottement latent traduisant la qualité du matériau mis en mouvement. Par contre, cette thèse souligne aussi, tout comme Kavetski et al. (2002), qu’une prise en compte aussi exacte que possible des erreurs de sortie est nécessaire pour inférer sans biais au travers d’un modèle non inversible certains paramètres décrivant les lois des variables d’entrée.

8.5.2.4. *Signification des inconnues*

En dynamique des avalanches, les coefficients de frottement ont un statut hybride. Du fait de la nature changeante du fluide en écoulement, l’un d’entre eux au moins varie d’une avalanche à l’autre et est de ce fait une variable, latente en général, mais observable si un modèle inversible comme celui de Coulomb est utilisé. Les autres coefficients de frottement peuvent, selon les choix du modélisateur, être également des variables ou ne refléter que les caractéristiques topographiques et être ainsi des paramètres au sens mathématique du terme. C’est le cas du paramètre dit « turbulent » dans les deux modèles présentés utilisant une loi de frottement de type Voellmy. En hydrologie déterministe, au contraire, tous les paramètres reflètent les propriétés du bassin versant. Mais du fait du caractère empirique des modèles utilisés, aucune relation déterministe évidente n’existe entre les paramètres et des covariables observables quantifiant la topographie, la géologie ou le couvert forestier du bassin versant. Le problème du transfert d’information aux bassins non instrumentés discuté au paragraphe 8.1.3 pour les avalanches est donc également délicat en hydrologie.

8.6. Conclusion et perspectives

8.6.1. Une approche méthodologique pour des outils opérationnels

Au cours de cette thèse, trois modèles de prédétermination locaux correspondant à des compromis différents entre faisabilité de l’inférence, réalisme physique de la description, disponibilité des données et temps de calcul ont été développés. Comme le prévoyait le projet initial, la thèse a donc été jusqu’à la mise au point d’outils à caractère opérationnel pour la prédétermination locale.

De plus, deux élargissements de la thématique initiale ont été abordés de façon à répondre à différents problèmes rencontrés par l’ingénieur : le passage à une approche régionale et/ou non stationnaire de la modélisation des occurrences avalancheuses et l’analyse décisionnelle du dimensionnement d’un ouvrage. Chacune des hypothèses de modélisation effectuées reste cependant discutable et il est certain que de nombreuses améliorations pourront encore être apportées à chacun des modèles proposés une fois que davantage d’expérience concernant leur utilisation pour des cas concrets aura été acquise.

D’un point de vue moins pratique, cette thèse a posé, grâce à la redéfinition d’un cadre formel, des jalons permettant d’avancer dans le sens d’une prédétermination plus rigoureuse des avalanches incluant connaissance naturaliste, modélisation physique et données historiques. Pour l’ensemble des travaux, les outils bayésiens d’inférence et de quantification des incertitudes ont été largement mis à profit pour surmonter les difficultés techniques. Ainsi, même si la transposition dans la pratique en ingénierie des outils proposés n’est pas pour tout de suite, la thèse propose au moins une base méthodologique illustrée par des cas d’étude concrets pour le développement futur de la modélisation stochastique des avalanches.

8.6.2. Pour aller plus loin...

8.6.2.1. *Exploiter plus largement les données historiques par des approches régionales*

Les bases de données avalanches françaises ont été très peu exploitées, à part pour des expertises locales, avant l'an 2000 car leur qualité était jugée insuffisante pour des analyses au niveau régional. Elles constituent pourtant une source d'information quasi unique au monde permettant d'apprendre sur un phénomène naturel complexe à l'échelle de massifs montagneux entiers. La demande est donc maintenant forte pour une valorisation plus systématique de ces données, à la fois pour la prédétermination, la prévision et la « connaissance pure ». D'autant plus que le travail de rénovation des bases de données engagé depuis 2001 ainsi que l'initiation au sein du Cemagref d'une démarche qualité renforcent la crédibilité de l'information disponible.

Cette thèse ayant contribué à montrer qu'il était possible de s'appuyer sur l'EPA pour illustrer des développements méthodologiques et extraire de l'information à l'échelle d'un département entier, de telles approches régionales sont certainement amenées à se multiplier dans les prochaines années. Les pistes prometteuses sont potentiellement très nombreuses : relation météorologie-déclenchement, variation spatiales et spatio-temporelles des pics d'activité, évaluation de la vulnérabilité...etc. L'effet du changement climatique sur la fréquence mais aussi l'intensité des avalanches figurant dans l'EPA constitue en particulier un axe de recherche prioritaire.

8.6.2.2. *Approfondir la prédétermination locale*

Les modèles stochastiques locaux proposés dans cette thèse restent des représentations assez grossières du fonctionnement d'un couloir avalancheux. A terme, des efforts d'amélioration de ces modèles seront certainement nécessaires. Il s'agira en particulier d'introduire de nouvelles variables pour aboutir à une représentation plus réaliste des conditions en zone de départ. Dans cette optique, les données météorologiques récemment acquises par le Cemagref pourront utilement compléter les informations contenues dans l'EPA.

De même, la hauteur de départ a été considérée abusivement comme observée dans les deux modèles avec une loi de frottement de Voellmy, et ce alors que la valeur utilisée a été obtenue après plusieurs transformations déterministes des dimensions du dépôt. Une amélioration naturelle des modèles proposés serait donc d'inclure cette transformation dans un opérateur stochastique supplémentaire traduisant l'erreur ainsi effectuée.

Le corollaire naturel de la complexification des modèles est l'amélioration des techniques d'inférence afin de limiter le temps de calcul. Deux directions mériteraient d'être explorées : soit le remplacement des algorithmes itératifs utilisés dans cette thèse par des algorithmes de type échantillonnage par importance qui ont l'avantage de ne pas nécessiter de période de chauffe avant convergence, soit une approche de type repro-modelling où, dans l'algorithme de calibration, le modèle de propagation est approximé par un modèle statistique plus simple grâce à des méthodes dérivées de la géostatistique.

De même, en prédiction, des méthodes de simulation de type FORM plus efficaces qu'une approche de Monte Carlo « brutale » et permettant de ce fait de réduire la quantité d'itérations du modèle de propagation nécessaires à l'obtention d'estimateurs précis des valeurs de projet mériteraient d'être essayées. Quant à l'effet des ouvrages de protection sur la distribution des aléas, il mériterait d'être discuté de façon plus approfondie, afin, sans nécessairement chercher à optimiser l'ouvrage, d'évaluer précisément son influence éventuelle sur le zonage.

Enfin, la prédétermination locale pourrait être améliorée par une meilleure prise en compte de l'expertise naturaliste grâce à un effort d'élicitation plus important que celui accompli au cours de la thèse.

8.6.2.3. *Hiérarchiser le modèle d'intensité*

Cette thèse s'est focalisée sur le développement d'outils de prédétermination locaux adaptés aux couloirs bien documentés. Pour le modèle fréquentiel, un élargissement spatial et spatio-temporel a toutefois été proposé. Son intérêt est de poser de premiers jalons vers ce qui constitue le principal défi pour la poursuite de cette thématique de recherche, c'est-à-dire la prédétermination de l'aléa avalancheux sur les couloirs non documentés, et ce dans un contexte éventuellement non stationnaire.

De futurs travaux devront donc se focaliser sur le développement pour le modèle d'intensité d'une méthode d'interpolation analogue à celle déjà développée pour le modèle fréquentiel. La modélisation bayésienne hiérarchique pourra continuer à être mise à profit en utilisant notamment les potentialités

des modèles spatiaux pour la mesure de la proximité topographique et nivo-météorologique inter-couloirs.

8.6.2.4. Développer les approches décisionnelles et la quantification de la vulnérabilité

Du fait des limites d'un aléa de référence défini en période de retour pour le dimensionnement d'un ouvrage paravalanche, cette thèse a proposé le passage à une approche décisionnelle. Néanmoins, la quantification de la vulnérabilité qui a été utilisée reste simpliste. De plus, la méthode développée reste limitée au seul cas des digues verticales impactées par des avalanches denses.

La systématisation du recours à une approche décisionnelle plutôt qu'à un aléa de référence passe donc par le développement d'autres outils adaptés aux différents cas de figure rencontrés dans la pratique en ingénierie : digues déviatrices, avalanches aérosol, etc. Ce faisant, des méthodes d'optimisation sous incertitude plus performantes que celle employée dans la thèse mériteraient sans doute d'être utilisées afin de permettre notamment l'optimisation jointe par rapport à plusieurs variables. Mais pour que les résultats des approches décisionnelles soient réellement jugés fiables, des progrès dans la quantification de la pression d'impact et de la vulnérabilité des structures aux avalanches devront au préalable être effectués.

Références

- Abbott, M. B., J. C. Bathurst, et al. (1986). An introduction to the European Hydrological System - Système Hydrologique Européen, 'SHE', 1: History and philosophy of a physically-based distributed modelling system. *Journal of Hydrology*. 87. pp 45-59.
- Adjel, G. (1995). Méthodes statistiques pour la détermination de la distance d'arrêt maximale des avalanches. *La Houille Blanche*. 7. pp 100-104.
- Adjel, G. (1996). Méthodes statistiques pour la détermination de la distance d'arrêt maximale des avalanches. Thèse de doctorat. Université Joseph Fourier. Grenoble. France. 156p.
- Amzal, B. (2004). Optimisation Bayésienne de décisions et de plans d'expériences par algorithmes particuliers. Thèse de doctorat en mathématiques de la décision. Université Paris-Dauphine. 220p.
- Amzal, B., Bois, F., Y., Parent, E., Robert, C.P. (2006). Bayesian-Optimal Design via Interacting Particle Systems. *Journal of the American Statistical Association*, Volume 101, Number 474. pp. 773-785.
- Ancey, C. (1998). Guide Neige et Avalanches. Edisud, Aix-en-Provence, 2nde édition.
- Ancey, C., Rapin, F., Martin, E., Coleou, C. et al. (2000). L'avalanche de Pécclerey du 9 février 1999. *La Houille Blanche*. 5. pp 45-53.
- Ancey, C., Richard, D. (2000). Détermination de l'aléa de référence. Rapport Cemagref /Météo France à la Direction de la Prévention des Pollutions et des Risques. 176 p.
- Ancey, C., Meunier, M., Richard, D. (2003). The inverse problem for avalanche dynamics models. *Water Resources Research* 39. 10.1029/2002wr001749.
- Ancey, C., Gervasoni, C., Meunier, M. (2004). Computing extreme avalanches. *Cold Regions Science and Technology* 39. pp 161-184.
- Ancey, C., Meunier, M. (2004). Estimating bulk rheological properties of flowing snow avalanches from field data. *Journal of Geophysical Research*. 109. (doi:10.1029/2003JF000036).
- Ancey, C. (2005). Monte Carlo calibration of avalanches described as Coulomb fluid flows. *Philosophical Transactions of the Royal Society of London A*. 363. pp 1529-1550. (doi: 10.1098/rsta.2005.1593).
- Ancey, C. dir. et al. (2006). Dynamique des avalanches. Presses Polytechniques et Universitaires Romandes. 334p.
- Anderson, C. W., Coles, S., Hüslér, J (1997). Maxima of Poisson-like variables and related triangular arrays. *The Annals of Applied Probability*. Vol 7. No 4. pp 953-971.
- Arnalds, P., Jonasson, K., Sigurdson, S. T. (2004). Avalanche hazard zoning in Iceland based on individual risk. *Annals of glaciology*. 38. pp 285-290.
- Atkinson, A., Donev, A. (1992). Optimum Experimental Designs. Oxford Press ed. 333p.
- Bacro, J. N. (2006). Modélisations stochastiques et valeurs extrêmes. Congrès SHF "Valeurs rares et extrêmes de débit...". Lyon. Mars 2006. pp 21-27.
- Bakkehoi, S., Dormaas, U., Lied, K. (1983). Calculation of snow avalanche runout distance. *Annals of Glaciology*. 4. pp 24-29.
- Banerjee, S., Carlin, B., Gelfand, A. E. (2003). Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall ed. 472p.

- Barbolini, M., Gruber, U., Keylock, C.J., Naaim, M., Savi, F. (2000). Application of statistical and hydraulic-continuum dense-snow avalanche models to five real European sites. *Cold Regions Science and Technology*. 31. pp 133-149.
- Barbolini, M., Savi, F. (2001). Estimate of uncertainties in avalanche hazard mapping. *Annals of Glaciology*. 32. pp 299-305.
- Barbolini, M., Keylock, C.J. (2002). A new method for avalanche hazard mapping using a combination of statistical and deterministic models. *Natural Hazard and Earth System Sciences*. 2. pp 239-245.
- Barbolini, M., Cappabianca, F. (2003). A new method for the estimation of avalanche distance exceeded probabilities. *Surveys in Geophysics*. 24. pp 587-601.
- Barbolini, M., Cappabianca, F., Savi, S. (2004a). Risk assessment in avalanche-prone areas. *Annals of Glaciology*. 38. pp 115-122.
- Barbolini, M., Cappabianca, F., Sailer, R. (2004b). Empirical estimate of vulnerability relations for use in snow avalanche risk assessment. Acte du congrès "Risk Analysis 2004". 27-29 Septembre. Rhodes. Grèce. pp. 533-542. C. A. Brebbia ed.
- Bardou, E. (2006). Réglementation en Suisse. In *Dynamique des Avalanches*. C. Ancey dir. Presses Polytechniques et Universitaires Romandes. 334p.
- Bartelt, P., Salm, B., Gruber, U. (1999). Calculating dense-snow avalanche runout using a Voellmy-fluid model with active/passive longitudinal straining. *Journal of Glaciology*. 45. pp 242-254.
- Bayes T. (1763). Essay Towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London*. 53. pp 370-418 et 54. pp 296-325.
- Beghin, P., Hopfinger, E. J., Britter, R. E. (1981). Gravitational convection from instantaneous sources on inclined boundaries. *Journal of Fluid Mechanics*. 107. pp 407-422.
- Bélangier, L. (2004). La rénovation de l'observation des avalanches : l'EPA et la CLPA. La lettre de l'IPGR, vol. supplémentaire n° 31. 4 p.
- Bélangier, L., Cassayre, Y. (2004). Projects for past avalanche observation and zoning in France, after 1999 catastrophic avalanches. Proceedings of the International Snow Survey Workshop. 19-24 September 2004. Jackson Hole, Wyoming. pp 416-422.
- Bertrand, D. (2007). Vulnérabilité des structures face aux aléas avalancheux : Synthèse préliminaire. Rapport interne. Cemagref ETNA. Grenoble France. 47p.
- Benestad, R. E. (2004). Record-values, non stationarity tests and extreme value distributions. *Global and Planetary Change*. 44. pp 11-26.
- Bergström, S. (1995). The HBV Model. Computer models in watershed hydrology. V. P. Singh ed. Water Resources Publications, Colorado, USA. pp 443-476.
- Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Second edition. Springer-Verlag ed. 617 p.
- Berliner, M. (2003). Physical-statistical modelling in geophysics. *Journal of Geophysical Research* 108(D24). (doi:10.1029/2002JD002865).
- Bernardo, J. M., Smith, A. F. M. (1994). *Bayesian Theory*. Wiley ed.
- Bernier, J., Parent, E., Boreux, J.J. (2000). *Statistiques pour l'environnement. Traitement bayésien de l'incertitude*. TEC&DOC ed. 303p.
- Bernier, J. (2003). Décisions et comportements des décideurs face au risque. *Journal des Sciences Hydrologiques* 48(3). pp 301-316.

- Berthet Rambaud, P. (2004). Structures rigides soumises aux avalanches et chutes de blocs : modélisation du comportement mécanique et caractérisation de l'interaction phénomène-ouvrage. Doctorat sciences et géographie, spécialité: mécanique et génie-civil. Université Joseph Fourier. Grenoble. 285 p.
- Berthier, B. (1986). Evaluation statistique des limites maximales atteintes par les avalanches à partir des données topographiques. Rapport Cemagref-RTM, 26p.
- Besag, J., York, J. and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*. 43, 1-59, (with discussion).
- Besson, L. (1996). Les risques naturels en Montagne - Traitement, prévention et surveillance, Artès-Publiaip ed. 438p.
- Best, N., Richardson, S., Thomson, A. (2004). A comparison of Bayesian Spatial Models for Disease Mapping. *Statistical Methods in Medical Research*, Vol14, 1. pp. 35-59.
- Beven K., (1989). Changing ideas in hydrology – The case of physically-based models. *Journal of Hydrology* 105. pp 157-172.
- Blagovechshenskiy, V., Eglit, M., Naaim, M. (2002). The calibration of an avalanche mathematical model using field data. *Natural Hazard and Earth System Science* 2. pp 217-220.
- Boreux, J.J. (2005). Modélisation statistique des valeurs extrêmes sous le paradigme bayésien. Note de cours. Département des Sciences et de Gestion de l'Environnement. Université de Liège. 26p.
- Borrel, G. (1993). Les cartes d'avalanches et leurs utilisations. *Neige et avalanches*. 64.
- Borrel, G. (1994). La carte de localisation probable des avalanches. *Mappemonde*. 4. pp 17-19.
- Bouchet, A. (2003). Etude expérimentale des avalanches denses de neige sèche. Thèse de doctorat. Université Joseph Fourier, Grenoble, France. 132 p.
- Bouchet, A., Naaim, M., Bellot, H., Ousset, F. (2004). Experimental study of dense snow avalanches: velocity profiles in steady and fully developed flows. *Annals of Glaciology*. Vol. 38. pp. 30-34.
- Box, G. E. P., Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Addison-Wesley ed.
- Bozhinskiy, A.N. (1992). Constructing distribution functions for the dynamic parameters of avalanches. *Bulletin Moscow State University. Geography*. 5.
- Bozhinskiy, A.N., Andreyev, Y. B. (1993). Distribution laws for parameters of natural processes with limited data. *Bulletin Moscow State University. Geography*. 3.
- Bozhinskiy, A.N., Nazarov, A.N., Chernouss, P.A. (2001). Avalanches: a probabilistic approach to modelling. *Annals of Glaciology*. 32. pp 255-258.
- Bozhinskiy, A.N. (2004). The Monte Carlo simulation of avalanche-type processes. *Annals of Glaciology* 38. pp 351-356.
- Brooks, S. P. (1998). Markov Chain Monte Carlo Method and its application. *The Statistician*. Vol 47, 1. pp 69-100.
- Brooks, S. P., Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*. Vol 7, 4. pp 434-455.
- Brooks, S. P., Roberts, G.O. (1998). Assessing convergence of Markov chain Monte Carlo algorithms. *Statistics and Computing*. 8. pp 319-335.
- Brugnot, G., Pochat, R. (1981). Numerical simulation of avalanches. *Journal of Glaciology*. Vol 27, 95. pp 77-88.
- Burnet, R., Marti, G. (1998). Avalanche maps and databases in Europe. Deliverable D1 of the EU project SAME.

- Buser, O., Frutiger, H. (1980). Observed maximum runout distance of snow avalanches and the determination of the friction coefficients μ and ξ . *Journal of Glaciology* 26. pp 121-130.
- Cabos, S. (2003). Quantification du bâti dans les emprises de la Carte de Localisation des Phénomènes d'Avalanche. *Maîtrise Géographie*. Institut de Géographie. Grenoble. France. 135 p.
- Cabos, S., Bélanger, L. (2003). Note exploratoire sur le nombre de constructions situées dans les emprises d'avalanches CLPA en France. 11p.
- Casassa, G., Narita, H., Maeno, N. (1989). Measurements of friction coefficients of snow blocks. *Annals of Glaciology*. 13 (40). pp 40-44.
- Casassa, G., Narita, H., Maeno, N. (1990). Shear cell experiments of snow and ice friction. *Journal of Applied Physics*. 69 (6). pp 3745-3755.
- Carlin, J. B., Chib, S. (1995). Bayesian model choice via Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society, series B*. 57. pp 473-484.
- Cemagref division Nivologie (1983). Activité de la division « nivologie » en matière de cartographie et d'analyse des risques d'avalanches. *Informations techniques*. Cahier 50, N°2. 4p.
- Cemagref ETNA (2000). Commune de Chamonix Mont Blanc, projet de Centre de secours principal près des Pèlerins, étude du risque d'avalanche. Rapport d'expertise, F. Rapin coord.
- Cemagref ETNA (2002). Ligne ferroviaire Pau-Canfranc. Etude du risque d'avalanche. Etude préalable a la réouverture. 44p.
- Cemagref ETNA (2004). Bilan de la campagne 2003/2004 de l'Enquête Permanente sur les Avalanches. Available online at <http://www.avalanches.fr/>. 35 p.
- Cemagref ETNA (2005). Bilan de la campagne 2004/2005 de l'Enquête Permanente sur les Avalanches. Available online at <http://www.avalanches.fr/>. 35 p.
- Cemagref ETNA (2006). Bilan de la campagne 2005/2006 de l'Enquête Permanente sur les Avalanches. Convention MEDD-ONF-Cemagref. N. Eckert, S. Garcia et L. Bélanger red. Available online at <http://www.avalanches.fr/>. 15p.
- Cemagref, Office National des Forêts, Ministère de l'Ecologie et du développement durable (2003). L'Enquête permanente sur les Avalanches : Manuel à l'attention des observateurs. 54 p.
- CTGREF division Nivologie (1974). L'archivage des données concernant les avalanches. L' « Observation Permanente des Avalanches ». Cahier 17, N°3. 2p.
- CTGREF division Nivologie (1975). La cartographie à échelle moyenne du risque d'avalanche. Les « cartes de localisation probable des avalanches ». Cahier 15, N°1. 3p.
- CEN (2000). Modélisation numérique du manteau neigeux, des risques d'avalanches et des principaux paramètres météorologiques en zone montagneuse. Note de centre. 15p.
- CEN, Cemagref ETNA (2006). Climatologie des situations nivo-météorologiques associées aux périodes avalancheuses importantes de ces 45 dernières années en Isère. Rapport de synthèse. J.P. Navarre coord. Projet financé par le Conseil Général de l'Isère. 22p.
- Chernouss, P. A., Fedorenko, Y. (2001). Application of statistical simulation for avalanche risk evaluation. *Annals of Glaciology*. 32. pp 182-186.
- Christen, M., Bartelt, P., Gruber, U. (2002). Aval-1D: an avalanche dynamics program for the practice. *Proceedings of the INTERPRAEVENT congress, Matsumoto, Japan*. Congress publication 2. pp 715-725.
- Clark., J.S. (2005). Why environmental scientists are becoming bayesians. *Ecology Letters*. 8. pp 2-14.

- Coeur, D., Lang, M., Nallet, R., Burnet, R., Strazzeri, D (1998). Histoire et connaissance des phénomènes naturels extrêmes. Ingénieries-EAT. Risques Naturels. pp 15-27.
- Coles, S. (2001). An introduction to statistical modelling of extreme values. Springer ed. 208p.
- Congdon, P. (2001). Bayesian Statistical Modelling. Wiley & Sons ed. 531p.
- Cosse, J. M., Vauterin, P., Bélanger, L., Garcia, S. (2005). Regards croisés sur l'épisode neigeux de janvier 2004 dans les Hautes Alpes. Neige et Avalanches 109.
- Cressie, N. (1993). Statistics for spatial data. Wiley ed. 928p.
- De Crecy, L. (1980). Avalanche zoning in France-Regulation and technical bases. Journal of Glaciology. 26-94. pp 325-330.
- De Finetti, B. (1937). La prévision : ses lois logiques, ses sources subjectives. Annales de l'Institut Henri Poincaré. 7. pp 86-133.
- De Marsily, G. (1994). Quelques réflexions sur l'utilisation des modèles en hydrologie. Revue des Sciences de l'Eau 7 (3). pp 219- 234.
- Dent, J. D., Lang, T. E. (1980). Modelling of snow flow. Journal of Glaciology. 26-94. pp 131-140.
- Dent, J. D. (1993). The dynamic friction characteristics of a rapidly sheared granular material applied to the motion of snow avalanches. Annals of glaciology. 18. pp 215-220.
- Diggle, P. J., Ribeiro, P. J. (2007). Model-Based Geostatistics. Springer ed. 228p.
- Diongue, S. (2006). Caractérisation des crues avalancheuses à différentes échelles de temps et d'espace. Rapport de stage de master professionnel Ingénierie Statistique. Université Joseph Fourier. Grenoble. France. 49p.
- Diongue, S., Eckert, N. (2006). Caractérisation des crues avalancheuses affectant les départements français. Rapport Cemagref ETNA diffusé aux services RTM. 16p.
- Domaas, U., Harbitz, C. B., Bakkehoi, K. (2002). The EU CADZIE database for extreme and deflected snow avalanches. Natural Hazard and Earth System Science. 2. pp 227-238.
- Durand, Y., Laternser, M., Giraud, G., Etchevers, P. et al. (2006). 44 years of climate reanalyses in the French Alps (1958 – 2002). Part I: methodology, model validation and temperature-precipitation-snowfall trends. Submitted to International Journal of Applied Meteorology and Climatology.
- Eckert, N., Parent, (2006). Monte Carlo calibration of a dynamic model for avalanche predetermination using the Metropolis-Hastings algorithm. Proceedings of the 3rd international conference on avalanches and related subjects. 2-9 septembre 2006. Kirovsk. Russia. 8p.
- Eckert, N., Parent, E., Naaim, M., Richard, D. (2007a). Bayesian stochastic modelling for avalanche predetermination: from a general system framework to return period computations. Stochastic Environmental Research and Risk Assessment. Accepted for publication. DOI: 10.1007/s00477-007-0107-4.
- Eckert, N., Parent, E., Richard, D. (2007b). Revisiting statistical-topographical methods for avalanche predetermination: Bayesian modelling for runout distance predictive distribution. Cold Regions Science and Technology. 49. pp 88-107. DOI: 10.1016/j.coldregions.2007.01.005.
- Eckert, N., Parent, E., Belanger, L. et al. (2007c). Hierarchical modelling for spatial analysis of the number of avalanche occurrences at the scale of the township. Cold Regions Science and Technology. 50. pp 97-112. DOI: 10.1016/j.coldregions.2007.01.008.
- Eckert, N., Parent, E., Faug, T. (2007d). Optimal design under uncertainty of a passive defense structure against snow avalanches: from a general Bayesian framework to a simple analytical model. In preparation.
- Eckert, N., Parent, E., Faug, T., Naaim, M. (2007e). Bayesian optimal design of an avalanche dam using a multivariate numerical avalanche model. In preparation.

- Edijatno, Nascimento, N. d. O., Yang, X., Makhlof, Z., Michel, C. (1999). GR3J: a daily watershed model with three free parameters. *Hydrological Sciences*. 4. pp 263-277.
- Elliott, P., Wakefield, J., Best, N., Briggs, D. (2000). *Spatial Epidemiology, Methods and Applications*. Oxford ed. 475p.
- Etienne, D. (2006). Situations météorologiques synoptiques et risques d'avalanche. L'exemple de Vallorcine. Master 2 Evaluation et Gestion de l'Environnement et des Paysages de Montagne. Université Joseph Fourier - Grenoble 1 - Institut de Géographie Alpine. 84p.
- Faug, T., Naaim, M., Bertrand, D., Lachamp, P. Naaim-Bouvet, F. (2003). Varying dam height to shorten the run-out of dense avalanche flows: developing a scaling law from laboratory experiments. *Surveys in Geophysics*, 24(5/6). pp 555-568.
- Faug, T. (2004). Simulation sur modèle réduit de l'influence d'un obstacle sur un écoulement à surface libre. Thèse de doctorat en Sciences de la Terre, de l'Univers et de l'Environnement. Université Joseph Fourier. Grenoble. France. 175p.
- Faug, T., Naaim, M., Naaim-Bouvet, F. (2004). An equation for spreading length, center of mass and maximum run-out shortenings of dense avalanche flows by vertical obstacles. *Cold Regions Science and Technology*. 39. pp 141-151.
- Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N., Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*. 40. (W01101).
- Fuchs, S., Bründl, M., Stötter, J. (2004). Development of avalanche risk between 1950 and 2000 in the municipality of Davos, Switzerland. *Natural Hazards and Earth System Sciences*. 4. pp 263-275.
- Fuchs, S., Bründl, M. (2005). Damage potential and losses resulting from snow avalanches in settlements in the Canton of Grisons, Switzerland. *Natural Hazards*. 34. pp 53-69.
- Fuchs, S., Keiler, M., Zischg, A. et al. (2005). The long-term development of avalanche risk in settlements considering the temporal variability of damage potential. *Natural Hazards and Earth System Sciences*. 5. pp 893-901.
- Fuchs, S., McAlpin, M. C. (2005). The net benefit of public expenditures on avalanche defense structures in the municipality of Davos, Switzerland. *Natural Hazards and Earth System Sciences*. 5. pp 319-330.
- Fujizawa, K., Tsunaki, R., Kamishi, I. (1993). Estimating snow avalanche runout distances from topographic data. *Annals of Glaciology*. 18. pp 239-244.
- Fyffe, B., Zaiser, M. (2004). The effects of snow variability on slab avalanche release. *Cold Regions Science and Technology*. 40. pp 229-242.
- Garcia, S., Bélanger, L. (2002). Analyse de la régularité des observations de l'Enquête Permanente sur les Avalanches. DESS Ingénierie Mathématique. Université Joseph Fourier, Grenoble. France. 101 p.
- Gassner, M., Brabec, B. (2002). Nearest neighbour models for local and regional avalanche forecasting. *Natural Hazards and Earth System Sciences*. 2. pp 247-253.
- Gauer, P., Kristensen, K. (2005). Avalanche Studies and Model Validation in Europe, SATSIE; Ryggfonn measurements: Overview and dam interaction. NGI Report 20021048-10. Sognsveien 72, N-0806 Oslo, Norwegian Geotechnical Institute.
- Gelman, A., Carlin, J. B., Stern, H. S., Rubin D. B. (1995). *Bayesian Data Analysis*. Chapman & Hall ed. 526p.
- Gelman, A. (1996). Inference and monitoring convergence. *Markov Chain Monte Carlo in Practice*. pp 131-143. Chapman & Hall ed.
- Geman, S., Geman, D. (1984). Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol Pami-6, 6. pp 721-741.

- Ghini, A., Chang, C. J. (2005). STARTER: a statistical GIS-based model for the prediction of snow avalanche susceptibility using terrain features – application to Alta Val Badia, Italian Dolomites. *Geomorphology*. 66. pp 305-325.
- Gilks, W. R., Richardson, S., Spiegelhalter, D. J. (2001). *Markov Chain Monte Carlo in Practice*. Chapman & Hall ed. 486 p.
- Girard, P., Parent, E. (2004). The deductive phase of statistical analysis via predictive simulations : test, validation and control of a linear model with autocorrelated errors representing a food process. *Journal of Statistical Planning and Inference*. 124. pp 99-120.
- Glass, B., Huet, P., Rat, M., Tordjeman, R. (2000). Retour d'expérience sur l'avalanche du 9 février 1999 à Montroc, commune de Chamonix. Rapport d'expertise. Inspection Générale de l'Environnement, France. Available online at <http://www.avalanches.fr/>. 65p.
- Grandjean, E. (2000). Analyse statistique des conditions de déclenchement d'une avalanche. Rapport de DESS Ingénierie mathématique – UFR IMA . Université Joseph Fourier. Grenoble. France. 29 p.
- Grêt-Regamey, A., Straub, D. (2006). Spatially explicit avalanche risk assessment linking Bayesian networks to a GIS. *Natural Hazards and Earth System Sciences*. 6. pp 911-926.
- Gruber, U., Bartelt, P., Haefner, H. (1998). Avalanche hazard mapping using numerical Voellmy-fluid models. *Proceedings of the Anniversary Conference 25 years of snow avalanche research*. Voss, Norway. pp 117-121.
- Gruber, U., Bartelt, P. (2007). Snow avalanche hazard modelling of large areas using shallow water numerical methods and GIS. *Environmental Modelling and Software*. 22. pp 1472-1481.
- Gubler, H., Hiller, M., Klausegger, G., Suter, U. (1986). Messungen an Fliesslawinen. Zwischenbericht. *Mitteilungen des Eigenössischen Institutes für Schnee und Lawinenforschung*. 41. 69p.
- Guillou, G., Naveau, P., Dieboldt, J., Ribereau, P. (2006). Return Level Bounds for Extreme Values. Submitted.
- Harbitz, C, Issler, D, Keylock, C. J. (1998). Conclusions from a recent survey of avalanche computational models. *Proceedings of the anniversary conference 25 years of snow avalanche research*, Voss, 12-16 May. Norwegian Geotechnical Institute Publication. 203. pp 128-139.
- Harbitz, K. (1999). A survey of computational models for snow avalanche motion. Fourth European Framework Programme (ENV4-CT96-0258) *Avalanche Modelling, Mapping and Warning In Europe*. 130 p.
- Harbitz, C., Harbitz, A., Farrokh, N. (2001). On probability analysis in snow avalanche hazard zoning. *Annals of Glaciology*. 32. pp 290-298.
- Hakonardottir, K. M., Johannesson, T., Tiefenbacher, F., Kern., M. (2001). A laboratory study of the retarding effect of breaking mounds in 3, 6 and 9 m long chutes. IMO Report 01007. VEDURSTOFA ISLANDS.
- Hastings, W. K. (1970). Monte Carlo Sampling Methods using Markov Chains and their applications. *Biometrika*. 57. pp 97-109.
- Heffernan, J. E., Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society. Series B*. 66. pp 497-546
- Hopf, J. (1998). An overview of natural hazard zoning with special reference to avalanches: 25 years of experience and results in Austria. *Proceedings of the anniversary conference 25 years of snow avalanche research*, Voss, 12-16 May. Norwegian Geotechnical Institute publication. 203. pp 128-139.
- Hutter, K., Koch, T., Plüss, C. and Savage, S. (1995). The dynamics of avalanches of granular materials from initiation to runout, part II : Laboratory Experiments. *Acta Mechanica*. 109. pp 127-165.
- Issler, D. (1997). Modelling of snow entrainment and deposition in powder snow avalanches. *Annals of Glaciology*. 26. pp 253-258.

- Jamard, A. L., Bélanger, L. (2002). Etude de la similarité de fonctionnement des sites avalanches. DESS Ingénierie Mathématique option Statistique. Université Joseph Fourier. Grenoble. France. 29 p.
- Jamard, A. L., Garcia, S., Bélanger, L. (2002). L'enquête permanente sur les Avalanches (EPA). Statistique descriptive générale des événements et des sites. DESS Ingénierie Mathématique option Statistique. Université Joseph Fourier. Grenoble. France. Available online at <http://www.avalanches.fr/>. 101 p.
- Jamison, B., Stethem, C. (2002). Snow avalanche Hazards and Management in Canada: Challenges and Progresses. *Natural Hazards*. 26. pp 35-53.
- Jarry, F., Ancey, C. (2006). Contexte réglementaire du zonage d'avalanche en France. In *Dynamique des avalanches*. C. Ancey dir. Presses Polytechniques et Universitaires Romandes. 334 p.
- Jibson, R. W., Harp, E. L., Michael, J. A. (2000). A method for producing digital probabilistic seismic landslide hazard maps. *Engineering Geology*. 58. pp 271-299.
- Johannesson, T. (1998). Icelandic avalanche runout models compared with topographical models used in other countries. Proceedings of the anniversary conference 25 years of snow avalanche research, Voss, 12-16 May. Norwegian Geotechnical Institute publication. 203. pp 43-52.
- Jomelli, V., Delval, C., Grancher, D. et al. (2007). Probabilistic analysis of recent snow avalanche activity and climate in the French Alps. *Cold Regions Science and Technology*. 47. pp 180-192.
- Jonasson, K., Sigurdson, S. T., Arnalds, P. (1999). Estimation of Avalanche Risk. *Vedurstofu Islands*. Reykjavik. VI-R99001-UR01. 44 pp.
- Jordaan, I. (2005). *Decisions under Uncertainty. Probabilistic Analysis for Engineering Decisions*. Cambridge University Press.
- Kaas, R. E., Raftery, A. E. (1994). Bayes Factor. Technical report 254. Dep. of Statistics. University of Washington and Technical report. 571. Dep. of Statistics. Carnegie-Mellon University.
- Kadane, J. B., Wolfson, D. B. (1998). Experiences in elicitation. *The Statistician*. 47. pp 1-20.
- Karniadakis, G. E., Glimm, J. (2006). Uncertainty quantification in simulation science-Preface. *Journal of Computational Physics* 217. pp 1-4.
- Kavetski, D., Franks, S. W. and G. Kuczera (2002). Confronting input uncertainty in environmental modelling. In Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N. and R. Turcotte eds. *Calibration of Watershed Models*, AGU Water Science and Applications Series. 6. pp. 49-68.
- Keiler, M. (2004). Development of the damage potential resulting from avalanche risk in the period 1950–2000, case study Galtür. *Natural Hazards and Earth System Sciences*. 4. pp 249–256.
- Keiler, M., Zischg, A., Fuchs et al. (2005). Avalanche related damage potential – changes of persons and mobile values since the mid-twentieth century, case study Galtür. *Natural Hazards and Earth System Sciences*. 5. pp 49–58.
- Keylock, C. J., Barbolini, M. (2001). Snow avalanche impact pressure - vulnerability relations for use in risk assessment. *Canadian Geotechnical Journal*. 38. pp 227-238.
- Keylock, C. J., McClung, D., Magnusson M. (1999). Avalanche risk mapping by simulation. *Journal of Glaciology* 45 (150). pp 303-314.
- Keylock, C. (2005). An alternative form for the statistical distribution of extreme avalanche runout distances. *Cold Regions Science and Technology*. 42. pp 185-193.
- Kies, R. (2007). Modélisation spatiale et spatio-temporelle des occurrences avalanches à partir des données historiques continues dans l'Enquête Permanente sur les Avalanches. Rapport de stage de master professionnel Ingénierie Statistique. Université Joseph Fourier. Grenoble. France. 59 p.

- Kleemayr, K. (2004). Modelling and simulation in snow science. *Mathematics and Computers in Simulation*. 66. pp 129-153.0
- Kokkonen, T., Koivusalo, H., Karvonen, T. (2001). A semi-distributed approach to rainfall-runoff modelling: a case study in a snow affected catchment. *Environmental Modelling and Software* 16. pp 481-493.
- Krzysztofowicz, R. (1983). Why should a forecaster and a decision maker use Bayes Theorem. *Water Resources Research*. Vol 19, 2. pp 327-336.
- Krzysztofowicz, R. (2001a). The case of probabilistic forecasting in hydrology. *Journal of Hydrology*. 249. pp 2-9.
- Krzysztofowicz, R. (2001b). Integrator of uncertainties for probabilistic river stage forecasting: precipitation-dependent model. *Journal of Hydrology*. 249. pp 69-85.
- Kuczera G., Parent E. (1998). Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology*. 211, 1-4. pp. 69-85.
- Lang, M. (2000). Notes sommaire sur l'estimation des paramètres de la distribution des crues, 14 p.
- Lauritzen, S. (1996). *Graphical Models*. Oxford University Press.
- Lee, P. M. (1997). *Bayesian Statistics, an introduction*. Second Edition. Wiley & Smith ed. 343 p.
- Lied, K., Bakkehoi, S. (1980). Empirical calculations of snow-avalanche run-out distances based on topographic parameters. *Journal of Glaciology*. 26 (94). pp 165-176.
- Lied, K., Moe, A., Kristensen, K. et al. (2001). Ryggfonn. Full scale avalanche test site and the effect of the catching dam. *Proceedings of the symposium Snow and avalanche test sites*. 22-23 novembre 2001. Grenoble. France.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* 201. pp 272-288.
- Llamas, J. (1993). Neige: formation, propriétés physiques et fonte. *Hydrologie Générale, Principes et applications* chapitre 10. G. Morin ed., Québec. Canada. 527 p.
- Maggioni, M., Gruber, U. (2003). The influence of topographic parameters on avalanche release and frequency. *Cold Regions Science and Technology* 37. pp 407-419.
- Maggioni, M. (2005). *Avalanche release Areas and their influence on uncertainty in avalanche hazard mapping*. PhD Thesis. University of Zurich.
- Marchal, L., Bélanger, L., Garcia, S. (2003). *Corrélations météorologiques avec l'EPA (Enquête Permanente sur les Avalanches)*. Rapport IUP GMI option Mathématiques Appliquées et Industrielles. Université Joseph Fourier. Grenoble, France. 69 p.
- Martinelli, M., Lang, T. E., Mears, I. (1980). Calculation of avalanche friction coefficients from field data. *Journal of Glaciology*. 26 (94). pp 109-119.
- Mases Cobero, M. (1997). *Analyse et modélisation de la répartition spatiale de la neige transportée par le vent sur des sites d'avalanches et sur des domaines skiables*. Thèse de doctorat. 180 p.
- MATE (1997). *Plans de Prévention des Risques Naturels (PPR), Guide général*. La documentation française ed. 4p.
- MEDD, METLTM (2004). *Construire en montagne, la prise en compte du risque d'avalanche*. M. Givry et P. Perfettini coords.
- MEDD, METLTM, MAAPAR, MISILL (2004). *Plans de Prévention des Risques Naturels, risques d'avalanches*. Guide Méthodologique. Version provisoire. 106 p.

- McClung, D. (1981). Fracture mechanical models of dry slab avalanche releases. *Journal of Geophysical Research*. 86(B11). pp 10783-10790.
- McClung, D., Lied, K. (1987). Statistical and geometrical definition of snow-avalanche runout. *Cold Regions Science and Technology*. 13. pp 107-119.
- McClung, D., Mears, A. I., Schearer, P. (1989). Extreme avalanche run-out: Data from four mountain ranges. *Annals of Glaciology*. 13. pp 180-185.
- McClung, D., Schearer, P. (1993). *The avalanche handbook*. 2nd edition. Mountaineers books ed. 271p
- McClung, D., Tweedy, J. (1994). Numerical avalanche prediction: Kootenay Pass, British Columbia, Canada. *Journal of Glaciology* 40 (135). pp 350-358.
- McClung, D. (2000). Extreme avalanche runout in space and time. *Canadian Geotechnical Journal*. 37. pp 161-170
- McClung, D. (2001). Extreme avalanche runout: a comparison of empirical models. *Canadian Geotechnical Journal*. 38. pp 1254-1265.
- McCollister, C., Birkeland, K., Hansen, K. et al. (2003). Exploring multi-scale spatial patterns in historical avalanche data, Jackson Hole Mountain Resort, Wyoming. *Cold Regions Science and Technology*. 37. pp 299-313.
- Mengersen, K.L., Robert, C.P., Guihenec-Jouyaux, C. (1999). MCMC Convergence Diagnostics : A Review (with Discussion). Pages 415–440 of : Bernardo, J. M., Berger, J. O., Dawid, A. P. , Smith, A. F. M. eds, *Bayesian Statistics*. Oxford University Press.
- Merz, B., Kreibich, H. Thielen, A., Schmidtke, R. (2004). Estimation uncertainty of direct monetary flood damage to buildings. *Natural Hazards and Earth System Sciences*. 4. pp 153–163.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M.N., Teller, A. H., Teller, E. (1953). Equation of State Calculations by Fast Computing Machine. *Journal of Chemical Physics*. 21. pp 1087-1091.
- Meunier, M., Ancy, C., Naaim, M. (2001). Mise au point d'une méthode de prédétermination statistique des cotes d'arrêt d'avalanches. *La Houille Blanche*. 67. pp 92-98.
- Meunier, M., Ancy, C., Taillandier, J. M. (2003). Fitting dynamic avalanche models with documented events at the Col du Lautaret site (France) using the conceptual approach. *Cold Regions Science and Technology*. 39. pp 55-66.
- Meunier, M., Ancy, C. (2004). Towards a conceptual approach to predetermining high-return-period avalanche run-out distances. *Journal of Glaciology*. 50-169. pp 268-278.
- Meunier, M., Ancy, C., Richard, D. (2004). *L'approche conceptuelle pour l'étude des avalanches*. Cemagref ed. 245 p.
- Micovic, Z., Quick, M. C. (1999). A rainfall and snowmelt runoff modelling approach to flow estimation at ungauged sites in British Columbia. *Journal of Hydrology*. 226. pp 101-120.
- Mollié, A. (1996). Bayesian mapping of disease. In *Markov Chain Monte Carlo in Practice*. 2nd edition (2001). W.R. Gilks, S. Richardson and D.J. Spiegelhalter eds. Chapman & Hall. pp 359-379.
- Monfort, A. (1982). *Cours de statistique mathématique*. Economica ed. 317 p.
- Mougin, P. (1922). *Les avalanches en Savoie*. Ministère de l'Agriculture, Direction Générale des Eaux et Forêts, Service des Grandes Forces Hydrauliques, Paris. pp 175–317.
- Müller, P. (1999). Simulation-based optimal design. *Bayesian Statistics* 6. pp 459-474.
- Munier, B., Rousseau, J-M. (1995). La rationalité face au risque. *Revue d'Economie Politique*.

- Naaim, M. (1998). Contribution to snow drift and avalanches flows modelling. Habilitation thesis, University Joseph Fourier. Grenoble. France. 180 p.
- Naaim, M., Naaim-Bouvet, F., Faug, T., Bouchet, A. (2004). Dense snow avalanche modelling: flow, erosion, deposition and obstacle effects. *Cold Regions Science and Technology*. 39. pp 193-204.
- Naaim, M. (2007). A propos de la pression dans le zonage de l'aléa avalanche. Note technique à destination des Services de Restauration des Terrains en Montagne. 11 p.
- Naveau, P., Nogaj, M., Amman, C. et al. (2005). Statistical methods for the analysis of climate extremes. *CRAS Geoscience*. 337. pp 1013-1022.
- Nespor, V., Sevruk, B. (1998) Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. *Journal of Atmospheric and Oceanic Technology*. 16. pp 450-464.
- Nishimura, K., Maeno, N. (1989). Contribution of viscous forces to avalanche dynamics. *Annals of Glaciology*. 13.
- Nuiry, L. (2004). Elaboration de multiples éléments de synthèse sur les données du projet EPA-CLPA, Licence Professionnelle Biostatistique UJF-UPMF. 87p.
- Oakley, J., O'Hagan, A. (2004a). Bayesian inference for the uncertainty distribution of computer model outputs. *Biometrika* 89. pp 769-784.
- Oakley, J., O'Hagan, A. (2004b). Probabilistic sensitivity analysis of complex models: a Bayesian approach. *Journal of the Royal Statistical Society. Series B*. 66. pp 651-659.
- O'Hagan, A., Oakley, J. (2004). Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering and System Safety*. 85. pp 239-248.
- Parent, E., Bernier, J. (2001). Méthodes Bayésiennes et modélisation des risques géophysiques extrêmes. *La revue du Modulad. INRIA*. 26. pp 1-26.
- Parent, E., Bernier, J. (2003a). Bayesian P.O.T. modelling for historical data. *Journal of Hydrology*. 274. pp 95-108.
- Parent, E., Bernier, J. (2003b). Encoding *prior* experts judgments to improve risk analysis of extreme hydrological events via POT modelling. *Journal of Hydrology*. 283. pp 1-18.
- Parent, E., Bernier, J. (2007). *Le raisonnement bayésien : modélisation et inférence*. Springer ed. 380 p.
- Perla, R., Cheng, T., McClung, D. (1980). A two-parameter model of snow-avalanche motion. *Journal of Glaciology*. 26. pp 197-207.
- Perla, R., Lied, K., Kristensen, K. (1984). Particle simulation of snow avalanche motion. *Cold Regions Science and Technology*. 9. pp 191-202.
- Perreault, L. (2000). Analyse bayésienne rétrospective d'une rupture dans les séquences de variables aléatoires hydrologiques. Thèse de doctorat. ENGREF. Paris. France.
- Perrin, C., Michel, C., Andreassian, V. (2001). Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*. 242. pp 275-301.
- Pickhands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*. 3. pp 11-130.
- Primus, M., Naaim-Bouvet, F., Naaim, M., Faug, T. (2004). Physical modelling of the interaction between mounds or deflecting dams and powder snow avalanches. *Cold Regions Science and Technology*. 39. pp 257-267.
- Rapin, F., Ancey, C. (2000). Occurrence conditions of two catastrophic avalanches at Chamonix, France. *Proceedings of the International Snow Science Workshop*. Bozeman Montana, USA, Octobre 2000.

- Rapin, F., Meunier, M., Bolognesi, R. (2004). Une nouvelle méthode d'identification des sites à haut risque d'avalanche. Ingénieries-EAT. 39. pp 43-53.
- Rapin, F., Bélanger, L., Hurand, A., Bernard, J. M. (2006). Sensitive avalanche paths : using a new method for inventory and classification of risk. Proceedings of the International Snow Science Workshop. Telluride, Colorado, 1-6 Octobre 2006.
- Renard, B., Lang, M., Bois, P. et al. (2006). Evolution des extrêmes hydrométriques en France à partir des données observées. Congrès SHF "Valeurs extrêmes de débit...". Lyon. Mars 2006. pp 21-27.
- Rivot, E. (2003). Investigations bayésiennes de la dynamique des populations de Saumon Atlantique. Des observations de terrain à la construction du modèle statistique pour apprendre et gérer. Thèse de doctorat, ENSAR.
- Robert C. (1992). L'analyse statistique bayésienne. Economica ed. 393 p.
- Robert, C. (1996). Méthodes de Monte Carlo par Chaînes de Markov. Economica ed. 340p.
- Salm, B. Burkard, A. Et Gubler, H.U. (1990). Calcul des avalanches : une méthode pour le praticien avec des exemples. Communication de l'institut fédéral suisse pour l'étude de la neige et des avalanches, C. Ancey traducteur. 38p.
- Salm, B. (1993). Flow, flow transition and runout distances of flowing avalanches. Annals of Glaciology. 18. pp 221-226.
- Salm, B. (2004). A short and personal history of snow avalanche dynamics. Cold Regions Science and Technology. 39. pp 83-92.
- Sampl, P., Naaim-Bouvet, F., Naaim, M. (2004). Interaction between dams and powder avalanches : determination of simple friction laws for shallow water avalanche models. Cold Regions Science and Technology 39. pp 115-131.
- Saporta, G. (1990). Analyse des données et statistique. Technip ed. 439p.
- Saunier, J. (1970). Rapport de la commission interministérielle d'enquête sur la catastrophe de Val d'Isère le 10 février 1970 et Rapport de la mission interministérielle d'étude sur la sécurité des stations de montagne. Secrétariat général du comité interministériel pour l'information. Available online at <http://www.avalanches.fr/>. 71p.
- Savage, S.B., Hutter, K. (1989). The motion of a finite mass of granular material down a rough incline. Journal of Fluid Mechanics. 199. pp 177-215.
- Savage, S. et Hutter, K. (1991). The dynamics of avalanches of granular materials from initiation to runout, part 1 : Analysis. Acta Mechanica. 86. pp 201-233.
- Schweizer, J. (1999). Review of dry snow slab avalanche release. Cold Regions Science and Technology. 30. pp 43-57.
- Sigurdsson, S., Jonasson, K., Arnalds, K. (1998). Transferring avalanches between paths. Proceedings of the anniversary conference 25 years of snow avalanche research, Voss, 12-16 May. Norwegian Geotechnical Institute publication 203. pp 259-263.
- Skaggs, T. H., Barry, D. A. (1997). The First-Order Reliability Method of predicting cumulative mass flux in heterogeneous porous formations. Water Resources Research. Vol. 33, 6. pp 1485-1494.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris. 8. pp. 229-231.
- SLF Davos (2000). Der Lawinenwinter 1999. Ereignisanalyse. 588 p.

- Straub, D. Grêt-Regamey, A. (2006). A Bayesian probabilistic framework for avalanche modelling based on observations. *Cold Regions Science and Technology*. 46. pp 192-203.
- Sovilla, B., Somavilla, F. and Tomaselli, A. (2001). Measurements of mass balance in dense snow avalanche events. *Annals of Glaciology*, 32. pp 230-236.
- Sovilla, B., Bartelt, P. (2002). Observations and modelling of snow avalanche entertainment. *Natural Hazards and Earth System Sciences*. 2. pp 168-179.
- Sovilla, B. (2004). Field experiments and numerical modelling of mass entrainment and deposition processes in snow avalanches. Ph.D Thesis, ETH Zurich, Switzerland.
- Sovilla, B., Schaer, M., Kern, M., Bartelt, P. (2007). Impact pressures and flow regimes in dense snow avalanches observed at the Vallé de la Sionne test site. Submitted to *Journal of Geophysical Research*.
- Spiegelhalter, D. J., Thomas, A., Best, N.G. (1996). Computation on Bayesian Graphical Models. *Bayesian Statistics*. 5. Calendon Press ed. pp 407-425.
- Spiegelhalter, D. J., Thomas, A., Best, N. G. (2000). WinBUGS Version 1.3 User Manual. MRC Biostatistics Unit.
- Spiegelhalter, D.J., Best, N., Carlin, B., Van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society. Series B*. 64. pp 583-640.
- Strazzeri, D., Manche, Y. (1998). L'enquête permanente sur les avalanches. *Revue de Géographie Alpine*. Vol. 86, n° 2. pp 45-52.
- Takeuchi, Y., Yamanoi, K., Endo, Y., Murakami, S., Izumi, K. (2003). Velocities for the dry and wet snow avalanches at Makunosawa valley in Myoko, Japan. *Cold Regions Science and Technology*. 37. pp 483-486.
- Tanner, M. H. (1992). *Tools for Statistical Inference: Observed Data and Data Augmentation Methods*. Springer-Verlag ed. 220p.
- Thomas, A., Best, N., Lunn, D. et al. (2004). *GeoBUGS Version 1.2 User Manual*.
- Torre F., Boreux J.-J. & Parent E. (2001).- The Metropolis-Hastings Algorithm, a handy tool for the practice of Environmental Model Estimation: illustration with Biochemical Oxygen Demand Data. *Cybergéogé*, article #187, 34 pp. available at <http://www.cybergeogeo.presse.fr/>
- Voellmy, A. (1955). Über die Zerstörungskraft von Lawinen. *Schweizerische Bauzeitung*, Jahrg. 73, 12. pp 159-162, 212-217, 246-249 and 280-285.
- Von Neumann, J., Morgenstern, O. (1953). *Theory of Games and Economic Behaviour*. New Jersey, USA. Princeton University Press.
- Von Mises, R., Geiringer, H. (1984). *The mathematical Theory of Probability and Statistics*. Academic Press ed.
- Wakefield, J. (1994). An Expected Loss Approach to the Design of Dosage Regimens Via Sampling-Based Methods. *The Statistician*, Vol. 43, No. 1. Special Issue: Conference on Practical Bayesian Statistics. pp. 13-29.
- Wilhelm, C. (1997) Wirtschaftlichkeit im Lawinenschutz. Methodik und Erhebungen zur Beurteilung von Schutzmassnahmen mittels quantitativer Risikoanalyse und ökonomischer Bewertung. *Mitt.Eidgenöss. Inst. Schnee- Lawinenforsch.*, 54, 1-309.
- Wilhelm, C. (1999). *Kosten-Wirksamkeit von Lawinenschutzmassnahmen an Verkehrsachsen*, BUWAL, Bern (Vollzug Umwelt, Praxishilfe).
- Yue, S., Pilon, P., Cavadias, G. (2002). Power of the Mann-Kendall and Spearman's rho test for detecting monotonic trends in hydrological series. *Journal of Hydrology*. 259. pp 254-271.
- Zambito, S., Bélanger, L., Garcia, S. (2002). Etude de corrélation sur les rangs de l'Enquête Permanente sur les Avalanches. *Rapport de Maîtrise d'IUP GMI spécialité MAI*. 24p.

Zwinger, T., (2000). Dynamik einer Trockenschneelawine auf beliebig geformten Berghängen. PhD-Thesis, Technische Universität Wien.

Et sur le web :

<http://www.avalanches.fr/>

<http://www.arena.org/>

Liste des Figures

Figure 1.1 : Catastrophe de Montroc.....	10
Figure 2.1 : Avalanche dense sur site expérimental.....	14
Figure 2.2 : Avalanche aérosol atteignant la forêt.....	15
Figure 2.3 : Dégâts dus à une avalanche aérosol.....	15
Figure 2.4 : Extrait de la CLPA.....	20
Figure 2.5 : Exemple d'un site EPA.....	22
Figure 2.6 : Site EPA sur fond topographique.....	22
Figure 2.7 : Altitude d'arrêt des avalanches sur sites EPA.....	23
Figure 2.8 : Volumes de dépôt des avalanches sur sites EPA durant l'hiver 2006/07.....	24
Figure 2.9 : Conditions météorologiques de déclenchement sur sites EPA durant l'hiver 2006/07.....	24
Figure 2.10 : Distribution temporelle des avalanches sur sites EPA durant l'hiver 2005/06.....	26
Figure 2.11 : Distribution spatiale des avalanches sur sites EPA durant l'hiver 2005/06.....	27
Figure 2.12 : Principe des méthodes Norvégiennes.....	29
Figure 2.13 : Filets paravalanches retenant la neige dans la zone de départ.....	30
Figure 2.14 : Protection paravalanche individuelle.....	31
Figure 2.15 : Principe des méthodes MCMC et mise en œuvre grâce à l'algorithme de Gibbs.....	36
Figure 3.1 : Cadre intégrateur général pour la prédétermination des avalanches.....	41
Figure 3.2 : Description systémique du fonctionnement d'un site avalancheux.....	46
Figure 3.3 : Effet d'une digue verticale sur un écoulement avalancheux.....	59
Figure 3.4 : Adéquation entre loi expérimentale et données.....	60
Figure 4.1 : DAG du modèle de Coulomb.....	66
Figure 4.2 : Cas d'étude pour le modèle de Coulomb.....	68
Figure 4.3 : Construction de la loi a priori sur le couloir d'apprentissage.....	69
Figure 4.4 : Prior et posterior pour le modèle de Coulomb.....	70
Figure 4.5 : Distributions moyennes a posteriori des variables d'entrée du modèle de Coulomb.....	71
Figure 4.6 : Adéquation entre données et moyenne a posteriori du modèle de Coulomb.....	72
Figure 4.7 : Cas d'étude pour le modèle de Coulomb bruité.....	73
Figure 4.8 : Profil de pente du cas d'étude pour le modèle de Coulomb bruité.....	73
Figure 4.9 : Distances d'arrêt observées et calculées pour différentes vitesses initiales.....	74
Figure 4.10 : Effet de la vitesse initiale sur les profils de vitesses avec le modèle de Coulomb.....	75
Figure 4.11 : DAG du modèle de Coulomb bruité.....	76
Figure 4.12 : Scatter plot des données d'entrée pour le modèle de Coulomb avec terme d'erreur.....	77
Figure 4.13 : Loi a posteriori des paramètres du modèle de Coulomb bruité.....	79
Figure 4.14 : Distributions moyennes des variables du modèle de Coulomb bruité.....	80
Figure 4.15 : Adéquation entre données et moyenne a posteriori du modèle de Coulomb bruité.....	81
Figure 4.16 : DAG des modèles de Voellmy bloc glissant et Saint-Venant.....	83
Figure 4.17 : Valeurs successives de deux chaînes avec les données simulées.....	87
Figure 4.18 : Convergence des variables latentes avec les données simulées.....	87
Figure 4.19 : Différence entre moyenne a posteriori des variables latentes et vraies valeurs.....	88
Figure 4.20 : Adéquation entre modèle et données simulées.....	89
Figure 4.21 : Cas d'étude pour les modèles de Voellmy bloc glissant et Saint Venant.....	90
Figure 4.22 : Scatter plot des données d'entrée du modèle de Voellmy bloc glissant.....	90
Figure 4.23 : Valeurs successives de deux chaînes avec les vraies données.....	92
Figure 4.24 : Lois a priori et a posteriori pour le modèle de Voellmy bloc glissant.....	93
Figure 4.25 : Corrélations a posteriori pour le modèle de Voellmy bloc glissant.....	94
Figure 4.26 : Loi a posteriori du coefficient de frottement latent du modèle Voellmy bloc glissant.....	95
Figure 4.27 : Comparaison entre modèle et loi a posteriori du coefficient de frottement latent.....	95
Figure 4.28 : Loi a posteriori de la distance d'arrêt latente avec le modèle Voellmy bloc glissant.....	96
Figure 4.29 : Distributions moyennes a posteriori des variables du modèle Voellmy bloc glissant.....	97
Figure 4.30 : Fonction de répartition des distances d'arrêt avec les données simulées.....	99
Figure 4.31 : Prior et posterior sur le paramètre ξ_{voel}	100

Figure 4.32 : Effet du prior sur ξ_{voel} sur les coefficients de frottement latents estimés.....	100
Figure 4.33 : Variables latentes estimées avec ξ_{voel} fixé.....	101
Figure 4.34 : Scatter plot des données d'entrée du modèle de Saint-Venant.....	104
Figure 4.35 : Lois a priori et a posteriori pour le modèle de Saint-Venant.....	106
Figure 4.36 : Corrélations a posteriori pour le modèle de Saint-Venant.....	107
Figure 4.37 : Loi a posteriori du coefficient de frottement latent du modèle de Saint-Venant.....	107
Figure 4.38 : Distribution a posteriori des distances d'arrêt latentes du modèle de Saint-Venant.....	108
Figure 4.39 : Distributions moyennes des variables d'entrée du modèle de Saint-Venant (1).....	109
Figure 4.40 : Distributions moyennes des variables d'entrée du modèle de Saint-Venant (2).....	110
Figure 5.1 : Nombres d'avalanches observées sur le couloir EPA 5 de Bessans.....	113
Figure 5.2 : Nombres d'avalanches observées sur le couloir EPA 13 de Bessans.....	113
Figure 5.3 : Inférence sur le couloir EPA 5 de Bessans.....	114
Figure 5.4 : Inférence sur le couloir EPA 13 de Bessans.....	114
Figure 5.5 : DAG du modèle spatial de base.....	117
Figure 5.6 : Nombres totaux d'avalanches en Savoie.....	119
Figure 5.7 : Lois marginales a posteriori pour le modèle spatial de base.....	120
Figure 5.8 : Moyenne a posteriori des fréquences avalancheuses communales moyennes.....	121
Figure 5.9 : Moyenne a posteriori des risques relatifs.....	122
Figure 5.10 : Moyenne a posteriori de la composante spatiale.....	123
Figure 5.11 : Prédiction moyenne pour la fréquence avalancheuse annuelle.....	124
Figure 5.12 : Adéquation entre données et modèle spatial de base.....	125
Figure 5.13 : Moyenne a posteriori de la composante locale sans et avec la composante spatiale.....	126
Figure 5.14 : Nombres d'événements par hiver sur la commune de Sixt Fer à Cheval.....	128
Figure 5.15 : Composante spatiale après prétraitement des données.....	130
Figure 5.16 : Moyenne prédictive sans et avec prétraitement des données.....	130
Figure 5.17 : Ajustement d'un variogramme exponentiel sur les précipitations décennales.....	131
Figure 5.18 : Prédiction de la hauteur de précipitation décennale par krigeage.....	132
Figure 5.19 : Nombres attendus avant et après introduction des covariables.....	132
Figure 5.20 : Composantes locale et spatiale après introduction des covariables.....	133
Figure 5.21 : DAG du modèle spatio-temporel de base.....	134
Figure 5.22 : Box plot du terme temporel avec le modèle spatio-temporel de base.....	135
Figure 5.23 : Composante locale et composante spatiale avec le modèle spatio-temporel de base.....	136
Figure 5.24 : Composante spatiale et bruit cumulés avec le CAR propre et le modèle de krigeage... ..	137
Figure 5.25 : DAG du modèle avec trois régimes.....	140
Figure 5.26 : Lois a posteriori des régimes annuels pour le modèle à trois régimes.....	140
Figure 5.27 : Box plot du terme temporel avec le modèle à 3 régimes.....	141
Figure 5.28 : DAG du modèle final avec trois régimes et structure du second ordre.....	142
Figure 5.29 : Bruit local avec le modèle final.....	143
Figure 5.30 : Structure spatiale du premier et du second ordre avec le modèle final.....	144
Figure 5.31 : Structure temporelle du second ordre avec le modèle final.....	145
Figure 5.32 : Box plot du terme temporel avec le modèle final.....	145
Figure 5.33 : Médiane des fréquences prédites avec le modèle final.....	147
Figure 5.34 : Variation des risques relatifs en Savoie et Haute-Savoie.....	148
Figure 5.35 : Composante spatiale dans les Alpes du Sud.....	148
Figure 5.36 : Composante temporelle dans les Alpes du Sud.....	149
Figure 5.37 : Composante spatiale en Isère.....	150
Figure 5.38 : Comparaison des composantes temporelles entre les différentes régions.....	150
Figure 6.1 : Distribution des distances d'arrêt avec le modèle de Coulomb.....	152
Figure 6.2 : Période de retour avec le modèle de Coulomb.....	153
Figure 6.3 : Distances d'arrêt et période de retour avec le modèle de Coulomb bruité.....	155
Figure 6.4 : Distances d'arrêt et période de retour avec le modèle de Voellmy bloc glissant.....	156
Figure 6.5 : Distribution des vitesses avec le modèle de Voellmy bloc glissant.....	157
Figure 6.6 : Vitesses et période de retour avec le modèle de Voellmy bloc glissant.....	158
Figure 6.7 : Frottement latent et période de retour avec le modèle de Voellmy bloc glissant.....	159
Figure 6.8 : Distance d'arrêt et période de retour avec le modèle de Saint-Venant.....	160
Figure 6.9 : Distribution des vitesses et hauteurs maximales avec le modèle de Saint-Venant.....	162

Figure 6.10 : Hauteurs maximales et période de retour avec le modèle de Saint-Venant.....	162
Figure 6.11 : Vitesses maximales et période de retour avec le modèle de Saint-Venant.....	163
Figure 6.12 : Frottement latent et période de retour avec le modèle de Saint-Venant.....	164
Figure 6.13 : Nombre de Froude maximal et période de retour avec le modèle de Saint-Venant.....	165
Figure 6.14 : Lois prédictives des variables d'entrée du modèle de Coulomb.....	166
Figure 6.15 : Périodes de retour moyennes et prédictives.....	167
Figure 6.16 : Lois prédictives de différents quantiles avec le modèle de Coulomb.....	168
Figure 6.17 : Lois prédictives de différents quantiles annuels avec le modèle de Coulomb.....	169
Figure 6.18 : Lois prédictives de différentes périodes de retour avec le modèle de Coulomb.....	169
Figure 6.19 : Lois prédictives des quantiles de distance d'arrêt avec le modèle de Coulomb bruité..	170
Figure 6.20 : Lois prédictives des quantiles annuels avec le modèle de Coulomb bruité.....	171
Figure 6.21 : Lois prédictives des quantiles annuels avec le modèle de Voellmy.....	172
Figure 6.22 : Sensibilité au prior sur ξ_{voel} de la moyenne a posteriori des distances d'arrêt.....	174
Figure 6.23 : Fonction de répartition des distances d'arrêt avec ξ_{voel} estimé ou fixé.....	175
Figure 6.24 : Distribution des vitesses avec ξ_{voel} estimé ou fixé.....	176
Figure 6.25 : Sensibilité de la loi prédictive des distances d'arrêt au modèle de mélange.....	178
Figure 6.26 : Sensibilité aux variables d'entrée avec le modèle de Coulomb bruité.....	179
Figure 6.27 : Sensibilité aux variables d'entrée avec le modèle de Voellmy bloc glissant.....	180
Figure 6.28 : Sensibilité de la distribution des distances d'arrêt à la quantité de données.....	183
Figure 6.29 : Fonction de répartition des distances d'arrêt avec les trois modèles d'intensité.....	184
Figure 6.30 : Période de retour avec les trois modèles d'intensité.....	184
Figure 6.31 : Fonctions de répartition à quantité d'information équivalente pour l'inférence.....	185
Figure 6.32 : Période de retour à quantité d'information équivalente pour l'inférence.....	185
Figure 6.33 : Répartition des distances d'arrêt, modèles bloc glissant Voellmy et Saint-Venant.....	186
Figure 6.34 : Coefficients de frottement avec les modèles bloc glissant Voellmy et Saint-Venant....	187
Figure 6.35 : Vitesses maximales avec les modèles de bloc glissant Voellmy et Saint-Venant.....	187
Figure 6.36 : Pression théorique en écoulement libre avec le modèle de Voellmy bloc glissant.....	189
Figure 6.37 : Pression maximale théorique en écoulement libre avec le modèle de Saint-Venant.....	189
Figure 6.38 : C_x à une abscisse décennale.....	191
Figure 6.39 : Pression décennale maximale tenant compte de la rhéologie.....	192
Figure 6.40 : Pression centennale maximale tenant compte de la rhéologie.....	194
Figure 6.41 : Loi prédictive par rapport à la densité de la pression maximale décennale.....	194
Figure 7.1 : Dépôt de l'avalanche d'avril 2006 sur le site de Taconnaz.....	198
Figure 7.2 : Adéquation entre données et modèle pour les volumes de dépôt.....	202
Figure 7.3 : Distribution prédictive des quantiles annuels pour les volumes de dépôt.....	202
Figure 7.4 : Courbes enveloppes pour différentes périodes de retour.....	204
Figure 7.5 : Cas d'étude pour le modèle décisionnel analytique.....	207
Figure 7.6 : Inférence du modèle fréquentiel.....	208
Figure 7.7 : Inférence du modèle d'intensité analytique.....	209
Figure 7.8 : Risques analytiques pour un bâtiment situé à une abscisse centennale.....	210
Figure 7.9 : Sensibilité du risque bayésien analytique à la position de l'enjeu.....	211
Figure 7.10 : Sensibilité du risque bayésien analytique à la fréquence moyenne de dépassement.....	212
Figure 7.11 : Sensibilité du risque bayésien analytique au coût de destruction.....	213
Figure 7.12 : Sensibilité du risque bayésien analytique au paramètre α	214
Figure 7.13 : Cas d'étude pour le modèle décisionnel complet.....	216
Figure 7.14 : Distribution des variables d'entrée des avalanches de référence.....	217
Figure 7.15 : Risques numériques pour un bâtiment situé à une abscisse centennale.....	218
Figure 7.16 : Risques bayésiens numériques pour des abscisses décennale et trentennale.....	220
Figure 7.17 : Sensibilité du risque bayésien à la dépendance en vitesse du coût de destruction.....	221
Figure 7.18 : Distribution des distances d'arrêt sans et avec ouvrage.....	222
Figure 7.19 : Distribution des vitesses à une abscisse centennale sans et avec ouvrage.....	224

Liste des Tableaux

Tableau 4.1 : Statistiques descriptives des données d'entrée pour le modèle de Coulomb.....	68
Tableau 4.2 : Lois a priori marginales pour le modèle de Coulomb	69
Tableau 4.3 : Statistiques descriptives des lois a posteriori marginales pour le modèle de Coulomb ..	70
Tableau 4.4 : Distributions moyennes a posteriori des variables du modèle de Coulomb.....	71
Tableau 4.5 : Données d'entrée pour l'application du modèle de Coulomb bruité.....	72
Tableau 4.6 : Corrélations totales et partielles entre variables du modèle de Coulomb bruité	78
Tableau 4.7 : Lois a priori pour le modèle de Coulomb bruité	78
Tableau 4.8 : Statistiques descriptives des lois a posteriori pour le modèle de Coulomb bruité.....	79
Tableau 4.9 : Fonctions d'exploration pour les données simulées.....	85
Tableau 4.10 : Taux d'acceptation pour les données simulées.....	85
Tableau 4.11 : Lois a priori pour les données simulées	86
Tableau 4.12 : Test de convergence de Gelman et Rubin avec les données simulées	86
Tableau 4.13 : Lois a posteriori et vraies valeurs pour les données simulées	88
Tableau 4.14 : Données pour l'application du modèle de Voellmy bloc glissant.....	89
Tableau 4.15 : Corrélations empiriques entre variables du modèle de Voellmy bloc glissant.....	91
Tableau 4.16 : Lois a priori utilisées avec le modèle Voellmy bloc glissant	91
Tableau 4.17 : Fonctions d'exploration pour le modèle Voellmy bloc glissant.....	91
Tableau 4.18 : Taux d'acceptation avec le modèle Voellmy bloc glissant	92
Tableau 4.19 : Test de convergence de Gelman et Rubin avec le modèle Voellmy bloc glissant	92
Tableau 4.20 : Statistiques descriptives des lois a posteriori pour, modèle de Voellmy bloc glissant..	93
Tableau 4.21 : Corrélation a posteriori des paramètres du modèle de Voellmy bloc glissant	94
Tableau 4.22 : Distributions moyennes a posteriori des variables, modèle de Voellmy bloc glissant..	97
Tableau 4.23 : Corrélations a posteriori entre les variables du modèle de Voellmy bloc glissant.....	97
Tableau 4.24 : Statistiques descriptives des lois a posteriori avec les données simulées bruitées	98
Tableau 4.25 : Distances d'arrêt de référence avec les données simulées bruitées ou non.....	99
Tableau 4.26 : Sensibilité au prior de la loi a posteriori de ξ_{voel}	99
Tableau 4.27 : Statistiques descriptives des lois a posteriori avec ξ_{voel} fixé	102
Tableau 4.28 : Données pour l'application du modèle de Saint-Venant.....	103
Tableau 4.29 : Corrélations empiriques entre variables du modèle de Saint-Venant.....	104
Tableau 4.30 : Fonctions d'exploration avec le modèle de Saint-Venant.....	105
Tableau 4.31 : Taux d'acceptation avec le modèle de Saint-Venant.....	105
Tableau 4.32 : Statistiques descriptives des lois a posteriori pour le modèle de Saint-Venant.....	105
Tableau 4.33 : Corrélation a posteriori des paramètres du modèle de Saint-Venant	106
Tableau 4.34 : Distributions moyennes a posteriori des variables du modèle de Saint-Venant.....	110
Tableau 4.35 : Corrélation a posteriori pour les variables du modèle de Saint-Venant.....	110
Tableau 5.1 : Activité avalancheuse observée sur les deux sites étudiés	113
Tableau 5.2 : Lois a posteriori sur les deux sites étudiés	113
Tableau 5.3 : Modèle spatial de base, paramètres et variables latentes pour les communes 1 et 43..	121
Tableau 5.4 : DIC, modèle spatial de base M_0 contre les modèles simplifiés M_1 et M_2	126
Tableau 5.5 : Rapport spatial pour les différents modèles (Savoie-Haute-Savoie).....	129
Tableau 5.6 : Paramètres du modèle spatio-temporel de base (Savoie-Haute-Savoie).....	136
Tableau 5.7 : Paramètres globaux de différents modèles (Savoie-Haute-Savoie).....	137
Tableau 5.8 : Paramètres du modèle à trois régimes (Savoie-Haute-Savoie)	141
Tableau 5.9 : Paramètres du modèle final (Savoie-Haute-Savoie).....	144
Tableau 5.10 : DIC pour les différents modèles spatio-temporels testés (Savoie-Haute-Savoie).....	146
Tableau 6.1 : Distance d'arrêt et période de retour avec le modèle de Coulomb.....	154
Tableau 6.2 : Comparaison entre données et moyenne a posteriori du modèle de Coulomb bruité....	154
Tableau 6.3 : Comparaison entre données et moyenne du modèle de Voellmy bloc glissant.....	157
Tableau 6.4 : Vitesses des événements de référence avec le modèle de Voellmy bloc glissant	158
Tableau 6.5 : Frottement μ_{voel} des événements de référence, modèle Voellmy bloc glissant.....	159
Tableau 6.6 : Comparaison entre données et moyenne du modèle de Saint-Venant.....	161

Tableau 6.7 : Erreur de Monte Carlo avec le modèle de Saint-Venant	161
Tableau 6.8 : Hauteurs maximales des événements de référence avec le modèle de Saint-Venant	163
Tableau 6.9 : Vitesses maximales des événements de référence avec le modèle de Saint-Venant	164
Tableau 6.10 : Frottement μ_{voel} des événements de référence avec le modèle de Saint-Venant.....	164
Tableau 6.11 : Nombres de Froude des événements de référence avec le modèle de Saint-Venant..	165
Tableau 6.12 : Lois prédictives des variables du modèle de Coulomb	166
Tableau 6.13 : Comparaison entre moyenne et loi prédictive avec le modèle de Coulomb.....	167
Tableau 6.14 : Quantiles de distance d'arrêt avec le modèle de Coulomb.....	167
Tableau 6.15 : Quantiles annuels avec le modèle de Coulomb	168
Tableau 6.16 : Périodes de retour avec le modèle de Coulomb	170
Tableau 6.17 : Quantiles de distance d'arrêt avec le modèle de Coulomb bruité	171
Tableau 6.18 : Quantiles annuels avec le modèle de Coulomb bruité.....	172
Tableau 6.19 : Quantiles annuels avec le modèle de Voellmy bloc glissant.....	172
Tableau 6.20 : Statistiques descriptives des lois a posteriori, prior vague, modèle de Coulomb.....	173
Tableau 6.21 : Quantiles annuels avec un prior vague pour le modèle de Coulomb	174
Tableau 6.22 : Quantiles annuels avec un prior $N(700,50)$ sur ξ_{voel}	175
Tableau 6.23 : Quantiles annuels avec un prior $N(1700,200)$ sur ξ_{voel}	175
Tableau 6.24 : Quantiles annuels avec ξ_{voel} fixé à 1200.....	176
Tableau 6.25 : Facteurs de Bayes pour le choix du modèle de Coulomb bruité	177
Tableau 6.26 : Loi prédictive des distances d'arrêt et modèle de mélange	177
Tableau 6.27 : Sensibilité aux variables d'entrée du modèle de Coulomb bruité	179
Tableau 6.28 : Sensibilité aux variables d'entrée du modèle Voellmy bloc glissant	180
Tableau 6.29 : Statistiques descriptives des lois a posteriori, modèle bruité, 26 données	181
Tableau 6.30 : Distribution moyenne a posteriori des distances d'arrêt, modèle bruité, 26 données .	181
Tableau 6.31 : Périodes de retour avec le modèle bruité et 26 données.....	182
Tableau 6.32 : Lois a posteriori avec le modèle de Saint-Venant et 13 données	182
Tableau 6.33 : Période de retour avec le modèle de Saint-Venant et 13 données.....	182
Tableau 6.34 : Comparaison des vitesses entre les modèles de Voellmy et Saint-Venant.....	188
Tableau 6.35 : Pression en écoulement libre et période de retour avec le modèle de bloc glissant ...	188
Tableau 6.36 : Pression en écoulement libre et période de retour avec le modèle de Saint-Venant ...	190
Tableau 6.37 : Cx à une abscisse décennale	191
Tableau 6.38 : Cx à une abscisse trentennale	192
Tableau 6.39 : Cx à une abscisse centennale	192
Tableau 6.40 : Pression d'impact à une abscisse décennale.....	193
Tableau 6.41 : Pression d'impact à une abscisse trentennale.....	193
Tableau 6.42 : Pression d'impact à une abscisse centennale.....	193
Tableau 6.43 : Bilan sur la pression d'impact à une abscisse décennale.....	195
Tableau 6.44 : Bilan sur les modèles de prédétermination proposés.....	196
Tableau 7.1 : Statistiques descriptives des séries de dépassement	201
Tableau 7.2 : Facteurs de Bayes pour les volumes de dépôt.....	201
Tableau 7.3 : Lois a posteriori des paramètres du modèle POT pour les volumes de dépôt.....	202
Tableau 7.4 : Quantiles annuels pour les volumes de dépôt.....	203
Tableau 7.5 : Lois a posteriori des paramètres du modèle POT pour les nombres de Froude	203
Tableau 7.6 : Quantiles annuels pour les nombres de Froude	203
Tableau 7.7 : Scénarii centennaux retenus	204
Tableau 7.8 : Principales caractéristiques du cas d'étude pour le modèle décisionnel analytique.....	208
Tableau 7.9 : Hauteurs optimales avec le modèle décisionnel analytique	210
Tableau 7.10 : Bénéfices associés aux hauteurs optimales avec le modèle décisionnel analytique....	211
Tableau 7.11 : Hauteurs optimales avec le modèle décisionnel numérique	219
Tableau 7.12 : Bénéfices associés aux hauteurs optimales avec le modèle décisionnel numérique ...	220
Tableau 7.13 : Augmentation de la période de retour avec la hauteur d'obstacle optimale.....	223

Liste des Variables et notations

Alphabet latin :

A_{ct} : actualisation totale

$a_t, a_{obs}, a_{jt}, a_j, a_{tR}$: nombre d'avalanche de l'hiver t , nombre d'avalanche observé durant l'hiver t , nombre d'avalanche dans la commune j l'hiver t , nombre d'avalanche moyen dans la commune j , nombre d'avalanche total de l'hiver t à l'échelle régionale

a_1, a_2 : paramètres de la loi des abscisses de départ normalisées

$a_\rho, b_\rho / a'_\rho, b'_\rho$: paramètres *a priori* / *a posteriori* de la loi Gamma du paramètre ρ

$a_\lambda, b_\lambda / a'_\lambda, b'_\lambda$: paramètres *a priori* / *a posteriori* de la loi Gamma du paramètre λ

b_c^*, b_B^* : bénéfice attendu classique et bayésien pour une hauteur de digue optimale

B_{12} : facteur de Bayes entre les modèles concurrents M_1 et M_2

$Beta(\)$: loi Beta

b_1, b_2, σ_h : paramètres de la loi des hauteurs d'écoulement ou de départ

c, d, e, σ : paramètres de la loi du coefficient de frottement latent μ_{voel}

c, d, e, f, σ : paramètres de la loi du coefficient de frottement latent μ_{coul}

$C(\)$, [€] : coût total

$CAR(\)$: modèle conditionnel autoregressif

$C_o(\)$, [€] : coût de construction total

$C_1(\)$, [€] : coût de destruction total

c_j : nombre de sites EPA dans la commune j

ch : longueur de la période de chauffe pour les algorithmes MCMC

Cx : coefficient de proportionnalité entre énergie cinétique par unité de volume et pression d'impact

$data, data_1, data_2$: données utilisées pour l'inférence, données utilisées pour l'inférence du modèle fréquentiel, données utilisées pour l'inférence du modèle d'intensité

d : variable de décision

d_o , [m] : diamètre d'obstacle pour le calcul de pression d'impact

D_{chap} : log vraisemblance calculée pour la moyenne a posteriori des paramètres

D_{moy} : log vraisemblance moyenne

$dmulti(\)$: loi multinomiale

$E [\]$: espérance mathématique

$E[a|\theta_F]$, $E_d[a|\theta_F, \theta_M]$: fréquence avalancheuse moyenne, fréquence moyenne de dépassement de l'abscisse de la digue

e_j, e'_j : nombre d'avalanches attendu dans la commune j sans et avec normalisation par les covariables topographiques et météorologiques

F, F_Z, \hat{F}, F^{-1} : Fonction de répartition, Fonction de répartition de la variable aléatoire Z , Fonction de répartition estimée, Fonction de répartition inverse

- Fr, Fr_T : nombre de Froude, nombre de Froude de période de retour T
- $Frot, F_{voel}, F_{coul}$, [m/s²] : écriture générique du terme de frottement, frottement de Voellmy, frottement de Coulomb
- f_j : fréquence avalancheuse moyenne dans la commune j
- g : constante de gravité, 9.81 m/s²
- $\text{Gamma}(\)$: loi Gamma
- $G(\)$: modèle de propagation déterministe
- $GPD(\)$: loi de Pareto Généralisée
- G_t : champ latent représentant la variabilité interannuelle
- H_{jt} : champ latent représentant la variabilité du second ordre supposée décomposable en H_{1j} et H_{2t} , respectivement variabilité spatiale et temporelle du second ordre
- $h, h_x, h_{xt}, h_x^{\max}, h_o$, [m] : hauteur en écoulement, hauteur en écoulement à l'abscisse x , en écoulement à l'abscisse x et au temps t , hauteur maximale à l'abscisse x , hauteur en écoulement de référence
- h_d , [m] : hauteur de la digue
- h_C^* , [m] : hauteur de la digue optimale classique
- h_B^* , [m] : hauteur de la digue optimale bayésienne
- $h_{plaque}, L_{plaque}, l_{plaque}$, [m] : hauteur, longueur et largeur de la plaque de départ
- $h_{stop}, L_{stop}, l_{stop}$, [m] : hauteur, longueur et largeur du dépôt
- i_t : taux d'intérêt de l'année t
- $I_{\{ \}}$: fonction indicatrice
- $l(\)$: modèle de vraisemblance
- l_1, l_2 : paramètres de régression des log risques relatifs sur les variables topographiques et météorologiques
- M : nombre de sous unités spatiales (communes)
- M_i : Modèle i
- $N(\)$: loi normale
- N, N_1, N_2, N_T : nombre de répétitions du phénomène étudié
- n, n_1, n_2 : nombre de tirage de Monte Carlo effectuées
- $neige_{T_{10}}$, [mm d'eau] : précipitation hivernale décennale
- $p(\)$: distribution de probabilité
- p_1, p_2, p_3 : probabilités respectives des trois régimes climatiques
- $P(\)$: probabilité ou loi de Poisson
- P : nombre total d'événements au cours des T_{obs} années d'observation
- pD : nombre effectif de paramètres du modèle
- Pr , [Pa] : pression
- $q, q_{\alpha_c}, q_{N_{\alpha_c}}$: quantile, quantile correspondant au niveau de confiance α_c , quantile de loi normale correspondant au niveau de confiance α_c
- $R_t, m_g(R_t), \sigma_g^2(R_t)$: régime de l'année t , moyenne et variance du régime de l'année t
- $R_B(\), R_B'(\)$: risque bayésien, risque bayésien « complet » intégrant la variabilité du paramètre α

R_c () : risque classique

RR_j : risque relatif dans la commune j

r_s : rapport spatial entre variabilité spatiale et variabilité locale

$S(N)$: somme des dépassements correspondant aux N événements

t_d : instant du déclenchement

$T, T_N, T_{x_{stop}}, T_1, T_2$, [années] : période de retour, période de retour de N années, période de retour de l'abscisse x_{stop} , période de retour sans ouvrage de protection, période de retour à la même abscisse avec ouvrage de protection

T_{obs} : nombre d'années d'observation

u (), [€] : utilité

U_j : champ latent représentant la structure spatiale

$VAR[]$: variance

V_j : champ latent représentant le bruit local

$vlat$: notation générique pour les variables latentes d'un modèle

$v, v_x, v_{x_o}, v_x(h_d)$, [m/s] : vitesse, vitesse à l'abscisse x , vitesse de référence à l'abscisse x , vitesse à l'abscisse x conditionnellement à la hauteur de digue h_d

v_b, v_m, v_M , [m/s] : vitesse à l'abscisse de l'enjeu, vitesses seuils minimale et maximale pour le calcul du dommage infligé à l'enjeu

$v_{start_{eq}}$, [m³] : volume de départ équivalent

v_{stop}, v_{stop_T} , [m³] : volume de dépôt, volume de dépôt de période de retour T

v_{zero} , [m/s] : vitesse initiale

W_j : champ latent représentant la somme de la structure spatiale et du bruit local

x : abscisse ([m]) ou notation générique pour les variables d'entrée du modèle de propagation

x_{obs}, y_{obs} : intensité de l'avalanche observée (entrées et sorties)

x', x'' : séparation des variables d'entrée du modèles de propagation en entrées observables et variables latentes (coefficients de frottement)

x_b , [m] : abscisse du bâtiment ou plus généralement de l'enjeu considéré

x_d , [m] : abscisse de la digue

x_{min}, x_{max} , [m] : abscisse minimale et maximale de la zone de départ

x_{start}, x_{startn} , [m] : abscisse de départ, abscisse de départ normalisée

$x_{stop}, x_{stop_{cal}}, x_{stop_o}, x_{stop_q}, x_{stop_T}, x_{stop}(h_d)$, [m] : abscisse d'arrêt, abscisse d'arrêt calculée, abscisse d'arrêt de référence, abscisse d'arrêt d'ordre q correspondant à la probabilité de non dépassement $1 - q/100$, abscisse d'arrêt de période de retour T , abscisse d'arrêt conditionnellement à la hauteur de digue h_d

$x_o, x_o(t_d)$: forçage climatique, forçage climatique instantané à l'instant t_d du déclenchement

$y, y_o, y(h_d)$: notation générique pour les variables de sortie du modèle de propagation, variables de sortie du modèle de propagation de référence, variables de sortie du modèle de propagation modifiées par une digue de hauteur h_d

z : altitude ([m]) ou notation générique pour une variable aléatoire

z_{moy} , [m] : altitude moyenne

z^{pred} : loi prédictive de la variable aléatoire z
 z_{start} , [m] : altitude de départ
 z_{stop} , z_{min} , [m] : altitude d'arrêt, altitude minimale d'un site (fond de vallée)
 z_T : réalisation de la variable z correspondant à la période de retour

Alphabet grec :

α : pouvoir dissipatif de la digue
 α_c : niveau de confiance pour un test ou un intervalle de confiance
 α_{moy} : risque relatif moyen dans la région étudiée
 α_{sv}, k_{sv} : paramètres des équations de Saint-Venant intervenant dans l'intégration sur la verticale
 $\alpha / \beta, L_\alpha, L_\beta, H_\alpha, H_\beta$: grandeurs utilisées dans les méthodes norvégiennes, le « point β » correspond à l'abscisse du couloir où la pente passe en dessous de 10 degrés.
 $\delta(\)$: fonction de Dirac
 $\delta_{h^*}, \delta_{b^*}$: différence entre hauteur optimale bayésienne et classique, différence entre bénéfice à la hauteur optimale bayésienne et classique
 ϕ , [deg] : pente locale
 ϕ_o , [deg] : pente moyenne dans la zone de départ
 ϕ_b , [deg] : pente moyenne dans la zone d'arrêt
 ϕ_{exp} , [m⁻¹] : paramètre de portée du modèle de krigeage bayésien
 γ_{CAR} : paramètre de structure spatiale du modèle CAR propre
 $\Gamma(\)$: fonction Gamma
 κ , [m⁻¹] : frottement turbulent tenant compte de la hauteur de l'écoulement
 λ : paramètre de la distribution du modèle fréquentiel
 $\lambda_{v_{stop}}, \beta_{v_{stop}}, \rho_{v_{stop}}, s_{v_{stop}}$: paramètres et seuil du modèle POT pour le volume de dépôt v_{stop}
 $\lambda_{Fr}, \beta_{Fr}, \rho_{Fr}, s_{Fr}$: paramètres et seuil du modèle POT pour le nombre de Froude Fr
 μ_{voel} : coefficient de frottement basal de la loi de Voellmy
 μ_{coul} : coefficient de frottement basal de la loi de Coulomb
 ν, ν_1, ν_2 : paramètres de la loi empirique reliant la pression d'impact au nombre de Froude
 π : notation générique pour les inconnues (paramètres et variables latentes)
 ρ : paramètre de forme du modèle d'aléa du modèle décisionnel analytique
 ρ_{auto} : paramètre d'autorégression du modèle AR1
 ρ_N [kg/m³] : masse volumique de la neige en écoulement
 σ_{num} : écart type pour l'erreur numérique
 $\sigma_u, \sigma_v, \sigma_h, \sigma_{h1}, \sigma_g$: écart type pour les champs latents U, V, H, H_1, G
 σ_λ : écart type représentant la variabilité intercommunale avec un modèle binomial négatif
 θ : notation générique pour les paramètres d'un modèle
 θ_G : notation générique pour les paramètres du modèle déterministe de propagation
 θ_o : notation générique pour les paramètres du modèle d'entrée
 $\theta_x, \theta_y, \theta_A$: notation générique pour les paramètres des modèles d'erreur
 $\hat{\theta}$: estimateur ponctuel de θ

$\theta^{(k)}$: k ième valeur d'un échantillon de la loi de θ

θ_M : notation générique pour les paramètres du modèle d'intensité

θ_F : notation générique pour les paramètres du modèle fréquentiel

$\theta_{FR}, \theta_{F1}, \theta_{F2}, \theta_{F3}, \theta_{F4}, \theta_{FR1}, \theta_{FR2}, \theta_{FT1}, \theta_{FT2}$: notation générique pour les hyperparamètres du modèle fréquentiel régional spatial et spatio-temporel

ξ_{voel} [m/s²] : coefficient de frottement turbulent de la loi de Voellmy

ω_{jk}, ω_{j+} : matrice de distance, distance totale entre la commune j et ses voisines

Liste des abréviations et sigles

AgroParisTech : Institut des sciences et industries du vivant et de l'environnement.
AIC : Critère d'Information d'Akaike
CADZIE: (projet) Catastrophic Avalanche: Defense Structures and Zoning In Europe
CAR : (modèle) Conditionnel Auto Régressif
CEN : Centre d'Etude de la Neige
Cemagref : Institut de recherche pour l'ingénierie de l'agriculture et de l'environnement
CLPA : Carte de Localisation des Phénomènes d'Avalanche
CTGREF : Centre Technique du Génie Rural, des Eaux et Forêts
DAG : Direct Acyclic Graph
DIC : Critère d'Information de Déviance
EPA : Enquête Permanente sur les Avalanches
ENGREF : Ecole Nationale du Génie Rural, des Eaux et Forêts, membre d'AgroParisTech
ETNA : (unité de recherche) Erosion Torrentielle, Neige et Avalanches
FORM : First Order Reliability Method
GEV : Loi Généralisée des Valeurs Extrêmes
GREF : (corps du) Génie Rural, des Eaux et Forêts
GRN : (école doctorale) Géosciences et Ressources Naturelles
HDR : Habilitation à Diriger des Recherches
IGREF : Ingénieur du Génie Rural, des Eaux et Forêts
ICGREF : Ingénieur en Chef du Génie Rural, des Eaux et Forêts
IGGREF : Ingénieur Général du Génie Rural, des Eaux et Forêts
INRS : Institut National de la Recherche Scientifique (Québec)
MCMC : (méthodes de) Monte Carlo par Chaînes de Markov
MAP, MATE, MAAPAR : Ministère chargé de l'Agriculture
Moy. : Moyenne
MEDD, MEDAD : Ministère chargé de l'Ecologie et du Développement Durable
METLTM : Ministère chargé de l'Équipement
MISILL : Ministère chargé de l'Intérieur
MORSE : (équipe de recherche) MODélisation et Risque en Statistique Environnementale
ONF : Office National des Forêts
PPR : Plan de Prévention des Risques
PER : Plan d'Exposition aux Risques naturels prévisibles
POT : (modèle) Peak Over Threshold
PZEA : Plan de Zones Exposées aux Avalanches
RR : Risque Relatif
RTM : (service de) Restauration des Terrains en Montagne
SAME : (projet) Snow Avalanche Modelling and Warning in Europe
SD : Standard Deviation (écart-type)
Skew. : Asymétrie
SLF : (Eidgenössischen Institut für) Schnee und LawinenForschung
SSA : Sites Sensibles d'Avalanches
SIG : Système d'Information Géographique

Annexe : jeux de données utilisés

A1. Site EPA numéro 5 de Bessans

Carte d'identité du site et origine des informations

- Pays : France
- Département : Savoie (73)
- Commune : Bessans
- EPA : Couloir numéro 5 de la commune de Bessans, site toileté
- Topographie : Carte scannée en noir et blanc 1/25000, pas de MNT

Topographie du site

- *Description générale*

Le site avalancheux est situé sur le versant sud ouest de l'Ouille Alegra. La zone de départ potentielle débute par une forte cassure de pente. Elle est située sur un versant assez large à la pente moyenne et régulière. Le site est ensuite très peu chenalisé avec un dénivelé total moyen. Il aboutit dans la vallée du torrent d'Avéroles entre le hameau de la Goulaz et l'oratoire Sainte Catherine. La route secondaire Bessans-Avéroles est directement menacée.

- *Photographie*

Une photo dans le CD-rom EPA millésime 2004 (Figure A1.1).

Figure A1.1 : Site EPA 5 de la commune de Bessans, photographie officielle



- *Profil 2D*

- altitude maximale: 2510 m
- altitude minimale : 1760 m
- dénivelé total : 850 m
- longueur du couloir : 1200 m du point le plus haut au torrent d'Avéroles
- pente moyenne : 32 degrés
- chenalisation : nulle
- sinuosité : quasi rectiligne malgré l'impression un peu trompeuse donnée par la Figure A1.1
- profil de pente/concavité : profil relativement régulier jusqu'à x=850 m, la pente décroît ensuite rapidement jusqu'à la rivière (Figure 4.2)
- longueur approximative de zone de départ potentielle : Au maximum 800 m selon un critère de pense simple (pente toujours au dessus de 30 degrés), mais vraisemblablement nettement moins (l'abscisse de déclenchement la plus basse enregistrée correspond à x=560 m)
- zone de dépôt : zone de pente faible en bas du couloir de part et d'autre de la route secondaire Bessans-Avéroles (x=1078 m) jusqu'au torrent d'Avéroles (x=1186.8 m) au-delà duquel la pente remonte rapidement

- *Fonctionnement présumé de la zone de départ*

Le début du couloir est marqué par une forte rupture de pente. La pente reste ensuite supérieure à 30 degrés sur 500 mètres avant de rester constante autour de cette valeur sur environ 350 mètres. La fracture se produit donc préférentiellement en tout début de couloir au niveau de la rupture de pente mais peut également se produire plus bas. La largeur de la zone de départ est peu contrainte, de sorte que des avalanches de volumes très différents sont susceptibles d'être générées.

- *Zone d'arrêt, caractéristiques et enjeux*

La zone d'arrêt est relativement simple : la pente faiblit nettement à partir de $x=850$ m et s'atténue progressivement jusqu'au torrent d'Avéroles qui peut vraisemblablement être traversé par de grandes avalanches. L'unique enjeu présent dans l'emprise de ce site avalancheux est la route secondaire Bessans-Avéroles régulièrement traversée par des avalanches.

Activité avalancheuse enregistrée :

- *Sources*

Base de données EPA (1901-2003) et carnets forestiers de la commune de Bessans (1901-1976). Aucune information particulière concernant la fiabilité des données n'est disponible. Point positif : Bessans est l'une des communes les mieux renseignées de l'EPA. Points négatifs : fréquents changements d'observateurs entre 1920 et 1970, non recouvrements fréquents entre les carnets et la base EPA, présence de quelques valeurs aberrantes.

Figure A1.2 : site EPA 5 de la commune de Bessans, activité avalancheuse enregistrée

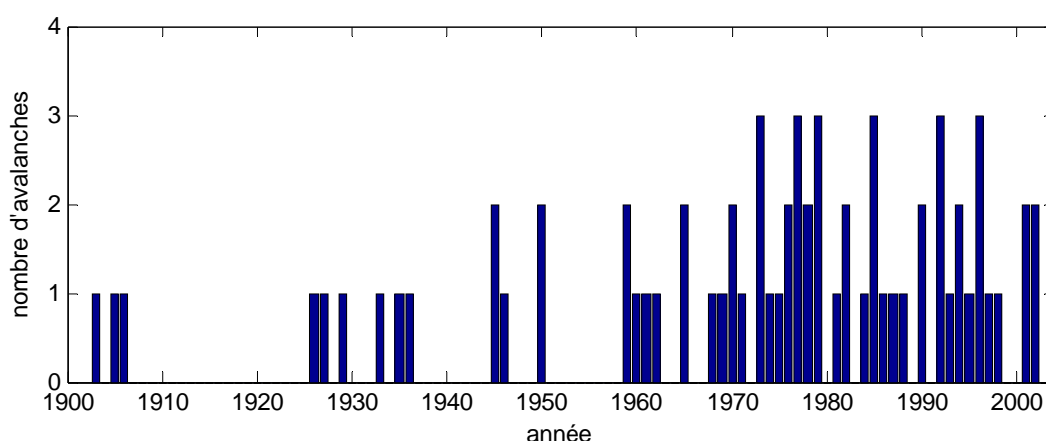


Tableau A1.1 : site EPA 5 de Bessans, test de stationnarité de l'activité avalancheuse enregistrée

Stationnarité	période	Nb. Camp.	Nb. Aval.	Test	Résultat	p-value
Activité annuelle	1904-2003	101	69	Spearman	1	0.0000002
Activité annuelle	1960-2003	44	53	Spearman	0	0.4

- *Activité et régularité de l'observation*

Au total, 69 avalanches ont été observées en 101 campagnes. 31 sont enregistrées dans la base et les carnets, toutes les autres, récentes, ne sont présentes que dans la base.

De zéro à trois avalanches ont été observées au cours d'une même campagne. L'observation a certainement été intermittente jusqu'en 1960, avec de nombreuses années consécutives sans avalanches (Figure A1.2). Logiquement, un test de stationnarité de la série des nombres d'avalanches par campagne est négatif pour l'ensemble des 101 campagnes. En revanche, l'hypothèse de stationnarité des occurrences avalancheuses ne peut être rejetée pour la période 1960-2003 sur laquelle la probabilité de déclenchement annuelle peut être évaluée. Il n'existe toutefois aucune preuve de l'exhaustivité de la collecte sur cette période (Tableau A1.1).

- *Altitudes de départ et d'arrêt*

Pour les altitudes de départ et d'arrêt, un désaccord classique existe entre la base EPA et les carnets forestiers. Une première correction conduit à privilégier les valeurs issues des carnets et à obtenir ainsi 66 valeurs d'altitudes de départ et d'arrivée. Un assez grand nombre de valeurs enregistrées est cependant incompatible avec la topographie du site : altitudes d'arrêt inférieures au fond de vallée et altitudes de départ supérieures au début du couloir. On peut en effet difficilement envisager que des avalanches se déclenchent au dessus de 2510 m d'altitude car les pentes y sont très faibles. De même, si des altitudes de 1750 m peuvent effectivement correspondre au fond de vallée (en réalité à 1760 m), des altitudes de 1700 m sont ininterprétables. Au final, ne sont donc conservées aucune des avalanches présentant une ou deux altitudes irréalistes, ce qui conduit à ne garder que 34 événements (Tableau A1.2)

Tableau A1.2 : site EPA 5 de la commune de Bessans, variables quantitatives enregistrées

	Nb. valeurs	Moy.	SD.	Min	Max	Skew.	q ₂₅	q ₇₅
z_{start} (m)	34	2313	145	2010	2500	-0.15	2200	2470
z_{stop} (m)	34	1783	30	1760	1900	2.1	1760	1800
v_{stop} (m ³)	11	15 273	8893	2900	29250	0.47	9150	23688

- *Dimensions des avalanches enregistrées*

22 dimensions de dépôts (longueur, largeur et hauteur) sont consignées dans les carnets forestiers. Elles correspondent à des avalanches enregistrées durant les campagnes 1905 à 1971. Si l'on se limite aux 34 avalanches sélectionnées à partir des altitudes de départ et d'arrêt, il reste cependant seulement 11 dimensions de dépôt. On ne travaillera donc pas avec les volumes de dépôt sur ce cas d'étude.

- *Indépendance des événements*

Des événements se produisant durant deux campagnes différentes peuvent être considérés comme physiquement indépendants. Lorsque plusieurs événements se produisent au cours d'une même campagne, l'indépendance n'est par contre pas garantie. C'est ici le cas pour 12 doublets et 6 triplets. Le délai séparant deux avalanches est alors à considérer : il est ici 5 fois inférieur à une semaine et dix fois d'une semaine à un mois, de sorte qu'un nombre significatif de couples d'avalanches ne sont vraisemblablement pas physiquement indépendants. Par contre, l'hypothèse d'indépendance statistique mutuelle des observations est tenable pour les altitudes de départ et d'arrêt, mais avec des niveaux de confiance assez faibles (Tableau A1.3).

Tableau A1.3 : site EPA 5 de Bessans, test d'indépendance des variables quantitatives enregistrées

Stationnarité	Nb. données	Test	Résultat	p-value
z_{start}	34	Spearman	0	0.09
z_{stop}	34	Spearman	0	0.06

Bilan et jeu de données utilisé

Le site EPA 5 de Bessans propose une information d'assez bonne qualité par rapport à la moyenne de ce que fournit l'EPA : longue série d'observations, recoupement base/carnets et visualisation du site sous SIG possible. De plus, la topographie du site est très régulière. Malheureusement, la série de volumes de dépôts est un peu courte. Ce cas d'étude est donc bien adapté à la mise au point du modèle de prédétermination le plus simple dit de Coulomb. Les jeux de données sélectionnés sont les suivants :

- une période assez longue, avec 55 avalanches en 44 campagnes pour l'évaluation de l'activité annuelle
- 34 altitudes de départ et d'arrivée pour la calibration du modèle de Coulomb. Les 34 doublets peuvent être considérés comme statistiquement indépendants.

A2. Site EPA numéro 13 de Bessans

Carte d'identité du site et origine des informations

- Pays : France
- Département : Savoie (73)
- Commune : Bessans
- EPA : Couloir numéro 13 de la commune de Bessans, site toileté
- Topographie : Carte scannée en noir et blanc 1/25000, pas de MNT

Topographie du site

- *Description générale*

Le site avalancheux est situé sur le versant sud-sud est de la pointe de Claret. La zone de départ potentielle est longue mais de superficie relativement modeste. Le site est ensuite presque rectiligne et peu chenalisé avec plusieurs petits ressauts successifs. Son dénivelé total est très fort. Il aboutit dans la vallée de l'Arc en amont du village de Bessans au niveau de la chapelle Saint Sébastien. Aucun enjeu notable ne semble à première vue pouvoir être identifié.

- *Photographie*

Une photographie officielle dans le CD-Rom EPA millésime 2004 (Figure A2.1).

Figure A2.1 : Site EPA 13 de la commune de Bessans, photographie officielle



- *Profil 2D*

- altitude maximale : 3200 m (Figure 4.7)
- altitude minimale : 1700 m
- dénivelé total : 1500 m
- longueur du couloir : 2278 m du point le plus haut jusqu'à l'Arc
- pente moyenne : 34.2 degrés
- chenalisation : faible
- sinuosité : site quasi rectiligne, légère courbure juste avant l'arrivée dans la vallée de l'Arc
- profil de pente/concavité : la pente décroît globalement très lentement. On observe cependant plusieurs petits ressauts, le dernier autour de x=1800 m : la pente qui était descendue aux alentours de 25 degrés remonte brutalement à plus de 40 degrés (Figure 4.8)
- longueur approximative de zone de départ potentielle : 1900 m selon un critère de pente simple
- zone de dépôt : courte zone de pente faible autour de x=2000 m puis arrivée rapide dans la vallée de l'Arc large et plane. Versant opposé à x=3200 m.

- *Fonctionnement présumé de la zone de départ*

Les fortes pentes sur la majeure partie du couloir autorisent une large variabilité des altitudes de départ. Suivant la longueur de la zone de départ, une large gamme de volumes est de ce fait envisageable malgré la largeur réduite de la surface de départ potentielle.

- *Zone d'arrêt, caractéristiques et enjeux*

La zone d'arrêt semble assez bien délimitée : il s'agit d'une zone de pentes faibles en sortie de couloir allant jusqu'au lit de l'Arc. Il est peu probable que des avalanches aillent au-delà de la rivière. Seule la chapelle Saint Sébastien et deux ou trois maisons sont dans l'axe du couloir mais *a priori* suffisamment loin de son exutoire pour être hors de portée d'avalanches même rares.

- *Ouvrages de protection*

Pas de paravalanches à signaler. Présence d'une digue en rive gauche de l'Arc (date de construction inconnue), *a priori* plutôt destinée à canaliser les crues mais qui joue probablement un rôle pour les avalanches extrêmes.

Données historiques

- *Sources :*

Base de données EPA (1901-2003) et Carnets Forestiers de la commune de Bessans (1901-1976). Aucune information particulière concernant la fiabilité des données n'est disponible. Point positif : Bessans est l'une des communes les mieux renseignées de l'EPA. Points négatifs : fréquents changements d'observateurs entre 1920 et 1970 et non recoupements fréquents entre les carnets et la base EPA.

Figure A2.2 : site EPA 13 de la commune de Bessans, activité avalancheuse enregistrée

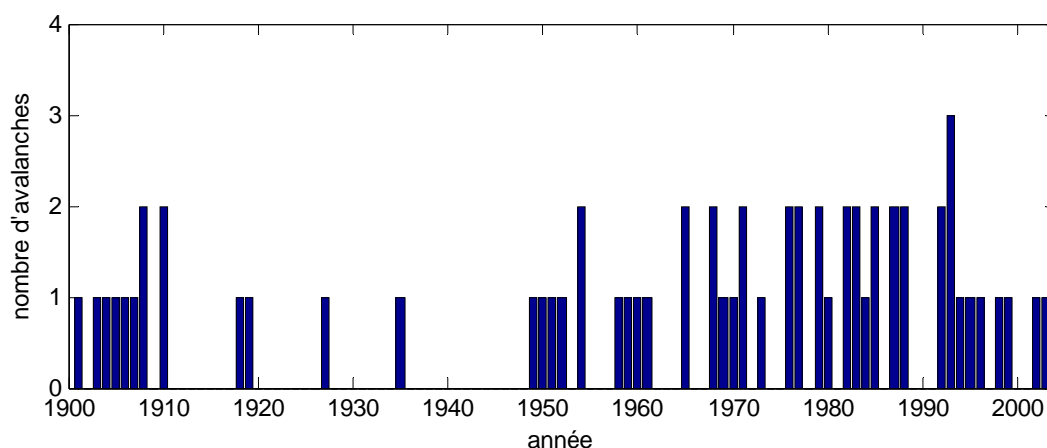


Tableau A2.1 : site EPA 13 de Bessans, test de stationnarité de l'activité avalancheuse enregistrée

Stationnarité	période	Nb. Camp.	Nb. Aval.	Test	Résultat	p-value
Activité annuelle	1901-2003	104	63	Spearman	1	0.001
Activité annuelle	1960-2003	44	41	Spearman	0	0.49

- *Activité et régularité de l'observation*

Au total, 63 avalanches ont été observées en 103 campagnes. 34 sont enregistrées dans la base et les carnets, deux ne figurent que dans les carnets et les autres, récentes, ne sont enregistrées que dans la base. D'après l'écriture dans les carnets forestiers, un même observateur a effectué le relevé de 1902 à 1920. Ensuite, de très nombreux observateurs se sont succédés jusqu'en 1977.

Au cours d'une même campagne, de zéro à trois avalanches ont été observées (Figure A2.2). L'observation a certainement été intermittente avec plusieurs années consécutives sans avalanches, notamment au moment des guerres mondiales. Logiquement, un test de stationnarité de la série des nombres d'avalanches par campagne est négatif pour l'ensemble des 103 campagnes (Tableau A2.1). En revanche, l'hypothèse de stationnarité des occurrences avalancheuses ne peut être rejetée pour la période 1960-2003 sur laquelle la probabilité de déclenchement annuelle peut être évaluée. Il n'existe toutefois aucune preuve de l'exhaustivité de la collecte sur cette période.

- *Altitudes de départ et d'arrêt*

Pour les altitudes de départ et d'arrêt, un désaccord classique existe entre la base EPA et les carnets forestiers : pour les campagnes anciennes, toutes les valeurs sont égales dans l'EPA. Ceci étant irréaliste, les valeurs issues des carnets doivent être privilégiées. Une fois la correction effectuée, 57 valeurs d'altitudes de départ et d'arrivée peuvent être obtenues, les autres étant manquantes. Aucune valeur aberrante n'est à déplorer (Tableau A2.2).

Cette information étant très abondante, on privilégiera autant que possible les événements les plus récents pour lesquels on suppose plus exacte l'enregistrement des valeurs. On appliquera donc le modèle de Coulomb uniquement sur les 40 doublets correspondant à la période 1960-2003.

Tableau A2.2 : site EPA 13 de la commune de Bessans, variables quantitatives enregistrées

	Nb. valeurs	Moy.	SD.	Min	Max	Skew.	Q ₂₅	Q ₇₅
z_{start} (m)	57	2671	307	2000	3100	-0.59	2400	2913
z_{stop} (m)	57	1777	73	1700	2000	1.51	1720	1800
v_{stop} (m ³)	26	69 000	170 000	110	705 600	2.72	7 000	210 00

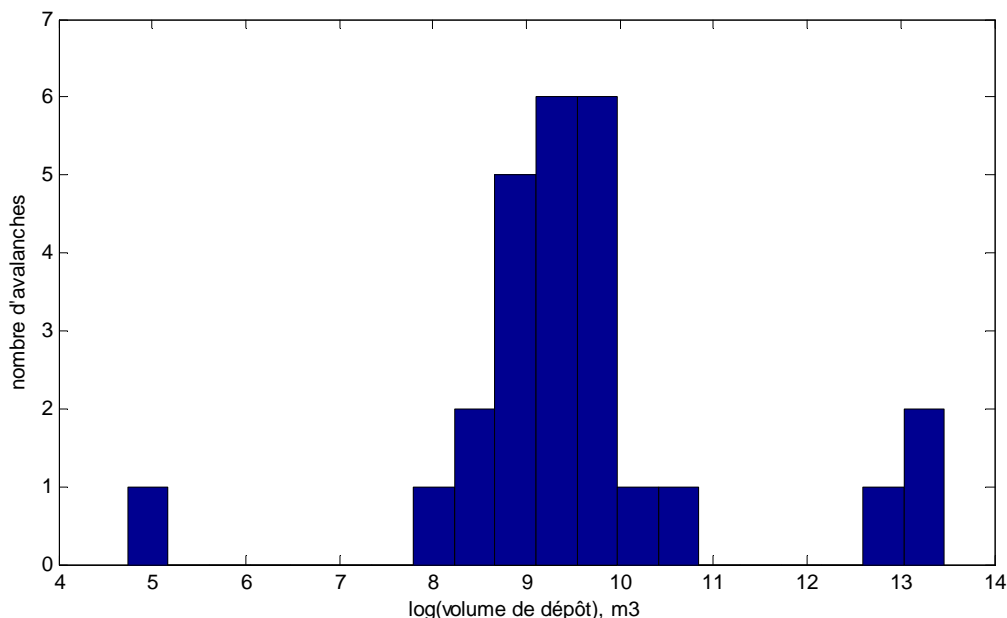
- *Dimensions des avalanches enregistrées*

28 dimensions de dépôts (longueur, largeur et hauteur) sont consignées dans les carnets forestiers. Elles correspondent à des avalanches enregistrées durant les campagnes 1902 à 1971. Toutefois, seules 26 d'entre elles correspondent à des avalanches pour lesquelles les données d'altitudes de départ et d'arrêt sont satisfaisantes. Les dimensions de dépôt peuvent être converties en volumes de dépôts avec

la formule classique du « triangle camembert », i.e. $v_{stop} = \frac{1}{2} L_{stop} \times l_{stop} \times h_{stop}$.

Les volumes d'avalanches sont très majoritairement petits à moyen (quelques milliers de m³, Figure A2.3). Néanmoins de très grosses avalanches (plus de 500 000 m³) ont été observées, la distribution empirique des observations étant très asymétrique (Tableau A2.2).

Figure A2.3 : site EPA 13 de la commune de Bessans, volumes de dépôt enregistrés



- *Autres données disponibles :*

Les données qualitatives et catégorielles classiques de l'EPA

- *Indépendance des événements :*

Des événements se produisant durant deux campagnes différentes peuvent être considérés comme physiquement indépendants. Lorsque plusieurs événements se produisent au cours d'une même campagne, l'indépendance physique n'est par contre pas garantie. C'est ici le cas pour 15 campagnes (14 fois deux avalanches, une fois trois). Le délai séparant les avalanches se produisant au cours d'une même campagne est alors à considérer : il est une fois inférieur à 24 heures (événements physiquement dépendants) ; deux fois d'une semaine à 15 jours, 3 fois de 15 jours à un mois et 11 fois supérieur à un mois (événements quasi certainement indépendants). Pour la majeure partie des événements, une hypothèse d'indépendance mutuelle semble donc raisonnable, à l'exception du doublet de la campagne 1976.

Pour plus de sûreté, un test d'indépendance statistique est possible étant donné la longueur de la série. Pour les altitudes de départ, d'arrivée, et pour les volumes de dépôt, un test classique ne permet pas de conclure à une corrélation significative entre les observations et le rang de la donnée dans la série. Dans chacune des séries d'observations, les différentes valeurs sont donc supposées statistiquement mutuellement indépendantes.

Tableau A2.3 : site EPA 13 de Bessans, test d'indépendance des variables quantitatives enregistrées

Indépendance	période	Nb. Camp.	Nb. données	Test	Résultat	p-value
z_{start}	1901-2003	103	57	Spearman	0	0.15
z_{stop}	1901-2003	103	57	Spearman	0	0.67
v_{stop}	1901-2003	103	26	Spearman	0	0.50

Bilan et jeu de données utilisé

Le site EPA numéro 13 de Bessans propose une information abondante et de qualité, en particulier par rapport à la moyenne de ce que fournit l'EPA : longue série d'observations, peu de valeurs aberrantes, nombreux volumes de dépôts, recoupement base/carnets et visualisation du site sous SIG possible. En outre, la topographie du site est « raisonnablement complexe » et l'écoulement n'est pas perturbé par la présence d'ouvrages de protection. Ce site est donc bien adapté à la mise en œuvre de l'ensemble des modèles de prédétermination locaux proposés. Les jeux de données sélectionnés sont les suivants :

- une période assez longue, avec 41 avalanches en 44 campagnes pour l'évaluation de l'activité annuelle
- 40 altitudes de départ et d'arrivée correspondant à la période 1960-2003 pour la calibration du modèle de Coulomb bruité. Les 40 doublets peuvent être considérés comme statistiquement indépendants.
- 26 volumes de dépôts accompagnés des altitudes de départ et d'arrivée pour la calibration des modèles de Voellmy bloc glissant et Saint-Venant. Les 26 triplets peuvent être considérés comme statistiquement indépendants.

A3. Site EPA numéro 30 de Bessans

Carte d'identité du site et origine des informations

- Pays : France
- Département : Savoie (73)
- Commune : Bessans
- EPA : Couloir numéro 30 de la commune de Bessans, site toileté
- Topographie : Carte scannée en noir et blanc 1/25000, pas de MNT

Topographie du site

- *Description générale*

Site avalancheux situé sur le versant nord ouest de la pointe de Soliet. Il correspond à une combe assez large drainée par le ruisseau de la Plate. La zone de départ potentielle est longue et assez large avec deux ressauts bien marqués correspondant à deux zones de pente forte. Le tracé est rectiligne et peu chenalisé. Le terrain tout d'abord nu devient peu à peu boisé. La zone de dépôt correspond essentiellement aux pentes faibles au bas du couloir et à la vallée de l'Arc atteinte au niveau de la chapelle Saint Maurice environ deux kilomètres en aval de Bessans. Quelques avalanches sont toutefois susceptibles de s'arrêter déjà avant le deuxième ressaut.

- *Photographie*

Une photographie officielle dans le CD-Rom EPA millésime 2004 (Figure A3.1).

Figure A3.1 : Site EPA 30 de la commune de Bessans, photographie officielle



- *Profil 2D*

- altitude maximale : 2850 m (Figure 7.5)
- altitude minimale : 1675
- dénivelé total : 1175 m
- longueur du couloir : 1762.8 m du point le plus haut à l'Arc
- pente moyenne : 38.2 degrés
- chenalisation : faible
- sinuosité : site quasi rectiligne,
- profil de pente/concavité : deux ressauts bien marqués débutant à x=500 m (la pente repasse de 25 à 55 degrés) et x=1200 m (la pente repasse de 25 à 45 degrés). Variations de concavité de second ordre habituelles.
- longueur approximative de zone de départ potentielle : 1550 m selon un critère de pente pure (toute pente supérieure à 30 degrés), mais en réalité on a plus vraisemblablement affaire à trois zones de départ différentes successives du fait de l'existence de deux forts ressauts.
- zone de dépôt : courte zone de pente faible autour de x=1200 m puis arrivée rapide dans la vallée de l'Arc large et plane.

- *Fonctionnement présumé de la zone de départ*

La zone de départ potentielle est assez complexe : il s'agit d'une combe assez large faite d'herbe et de barres rocheuses et caractérisée par la présence de trois zones de pente fortes séparées par des pentes plus modérées sur un dénivelé total de près de 1000 m. On peut donc envisager des types de déclenchement et des surfaces de départ extrêmement variées.

- *Zone d'arrêt, caractéristiques et enjeux*

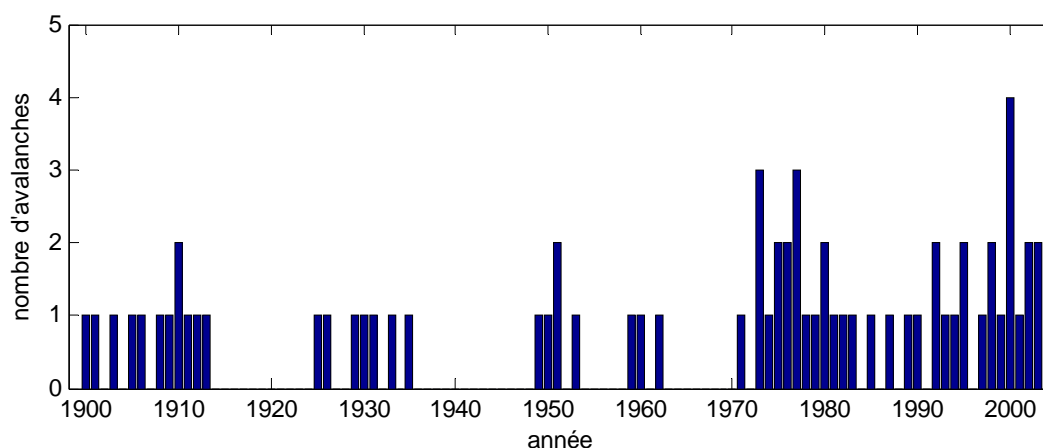
La zone d'arrêt semble assez bien délimitée : pentes moyennes avant le second ressaut pour les « petites » avalanches et zone de pente faible en sortie de couloir pour la plupart des autres. Seules de très grandes avalanches sont susceptibles de dépasser le lit de l'Arc. Pas de constructions menacées dans l'axe du couloir, le seul enjeu potentiel est lié à la présence de la RD 902 potentiellement recouverte par de très grandes avalanches.

Données historiques

- *Sources :*

Base de données EPA (1901-2003) et carnets forestiers de la commune de Bessans (1901-1976). Aucune information particulière concernant la fiabilité des données n'est disponible. Point positif : Bessans est l'une des communes les mieux renseignées de l'EPA. Points négatifs : fréquents changements d'observateurs entre 1920 et 1970 et non recoupements fréquents entre les carnets et la base EPA.

Figure A3.2 : site EPA 30 de la commune de Bessans, activité avalancheuse enregistrée



- *Activité et régularité de l'observation*

Au cours d'une même campagne, de zéro à quatre avalanches ont été observées (Figure A3.2). L'observation a certainement été intermittente avec plusieurs années consécutives sans avalanches, notamment au moment des guerres mondiales. Elle semble toutefois se stabiliser après 1970 et l'hypothèse de stationnarité du nombre d'événements par an ne peut être rejetée pour la période 1973-2003 sur laquelle la probabilité de déclenchement annuelle peut être évaluée. Il n'existe toutefois aucune preuve de l'exhaustivité de la collecte sur cette période (Tableau A3.1).

Tableau A3.1 : site EPA 30 de Bessans, test de stationnarité de l'activité avalancheuse enregistrée

Stationnarité	période	Nb. Camp.	Nb. Aval.	Test	Résultat	p-value
Activité annuelle	1901-2003	104	63	Spearman	1	0.0003
Activité annuelle	1973-2003	31	41	Spearman	0	0.83

- *Altitudes de départ et d'arrêt*

Pour les altitudes de départ et d'arrêt, un désaccord classique existe entre la base EPA et les carnets forestiers : pour les campagnes anciennes toutes les valeurs sont égales dans l'EPA. Ceci étant irréaliste, les valeurs issues des carnets doivent être privilégiées. Une fois la correction effectuée 65 valeurs d'altitudes de départ et d'arrivée peuvent être obtenues. Certaines d'entre elles sont incompatibles avec la topographie du site (altitude de départ trop élevée ou au contraire altitudes d'arrêt trop faibles) de sorte que seulement 54 événements peuvent être conservés. Et étant donné que ce cas d'étude est utilisé uniquement pour illustrer le modèle décisionnel analytique, on ne retient au final que les 28 altitudes d'arrêt inférieures à 1704 m sur la période 1973-2003 (Tableau A3.2).

Tableau A3.2 : site EPA 30 de la commune de Bessans, variables quantitatives enregistrées

	Nb. valeurs	Moy.	SD.	Min	Max	Skew.	q ₂₅	q ₇₅
z_{start} (m)	54	2412	258	1900	2800	-0.2	2200	2700
z_{stop} (m)	54	1719	31	1675	1850	1.36	1700	1745
$z_{stop} \mid z_{stop} < 1704$ (m)	28	1695	9	1675	1700	-1.4	1695	1700

- *Autres données disponibles :*

25 dimensions de dépôt ainsi que les données qualitatives et catégorielles classiques de l'EPA.

- *Indépendance des événements :*

Des événements se produisant durant deux campagnes différentes peuvent être considérés comme physiquement indépendants. Lorsque plusieurs événements se produisent au cours d'une même campagne, l'indépendance physique n'est par contre pas garantie. C'est le cas pour de nombreuses campagnes étant donné l'activité assez élevée de ce site.

Si l'on s'intéresse uniquement à la série des 28 altitudes inférieures à 1704 m, seuls les quatre événements de la campagne 2001 sont problématiques car se produisant tous en un peu plus d'un mois. Néanmoins un test statistique ne rejette pas l'indépendance mutuelle des 28 événements (Tableau A3.3).

Tableau A3.3 : site EPA 30 de Bessans, test d'indépendance des variables quantitatives enregistrées

Indépendance	période	Nb. Camp.	Nb. données	Test	Résultat	p-value
$z_{stop} \mid z_{stop} < 1704$	1973-2003	41	28	Spearman	0	0.41

Bilan et jeu de données utilisé

Le site EPA numéro 30 de Bessans propose une information assez abondante et de qualité satisfaisante, en particulier par rapport à la moyenne de ce que fournit l'EPA. Néanmoins la topographie du site est assez tourmentée avec deux ressauts marqués, de sorte qu'une limite son utilisation à l'application du modèle décisionnel simplifié. Celui-ci étant de type POT, il ne nécessite que les altitudes d'arrêt inférieures au seuil sur une période bien définie. Les jeux de données sélectionnés sont les suivants :

- 28 altitudes d'arrivée inférieures à 1704 m sur la période 1973-2003 pour la calibration du modèle d'intensité. Elles peuvent être considérées comme statistiquement indépendantes.
- une période de 31 campagnes avec 41 avalanches pour l'évaluation de la probabilité annuelle de dépassement.

A4. Site EPA numéro 9 des Houches, site dit de Tacconnaz

Carte d'identité du site et origine des informations

- Pays : France
- Département : Haute-Savoie (74)
- Commune : Les Houches
- EPA : Site numéro 9 de la commune des Houches appelé plus couramment “couloir de Tacconnaz”, site toiletté
- Topographie : Carte scannée en noir et blanc 1/25000, plusieurs MNT

Topographie du site

- *Description générale*

Site avalancheux situé dans la vallée de Chamonix et hors norme de par sa longueur, le volume des avalanches susceptibles de s'y produire, le nombre de travaux qui lui ont été consacrés et la présence d'un ouvrage paravalanche protégeant différents groupes d'habitations.

- *Photographies*

De très nombreuses photographies prises lors des avalanches récentes (Figure A4.1 par exemple), de même que des photographies du dispositif de protection (Figure 3.1). Une photographie officielle post-toilettage dans le CD-Rom EPA millésime 2004 (Figure 2.5).

Figure A4.1 : Site de Tacconnaz, avalanche d'avril 2006, photo RTM74

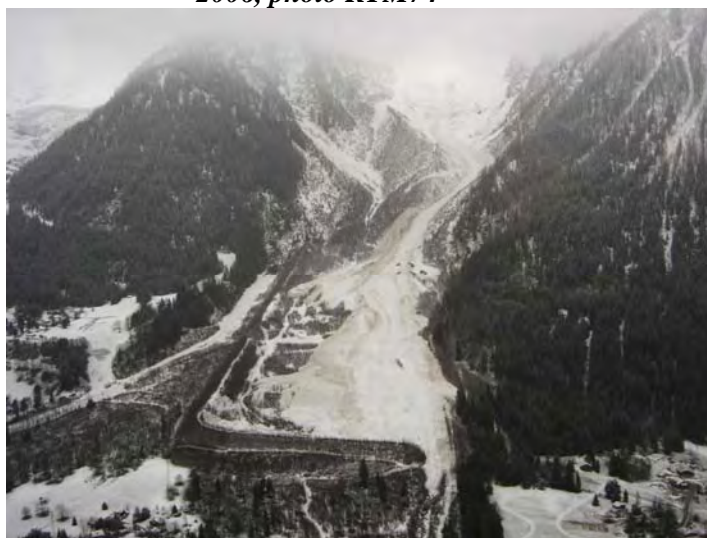
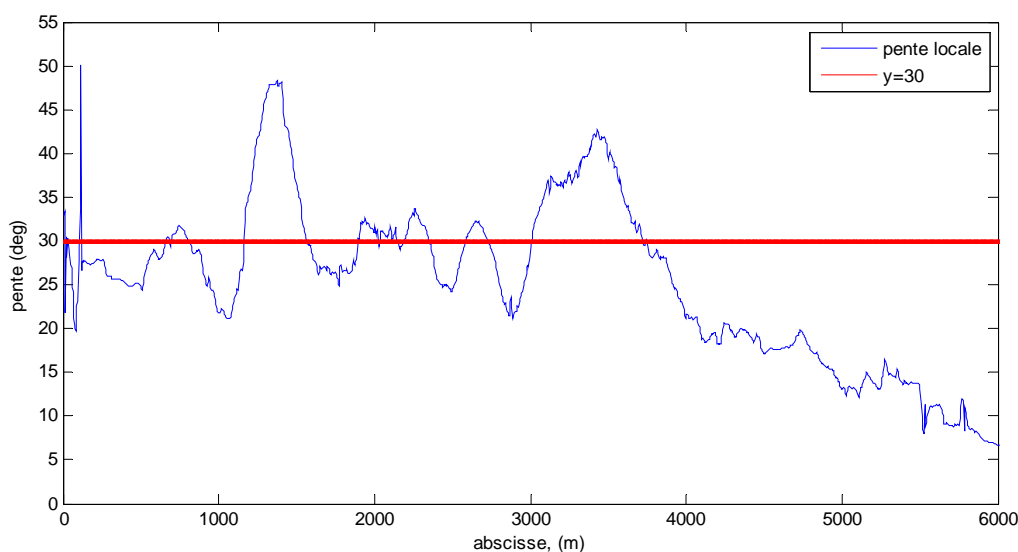


Figure A4.2 : profil de pente du site de Tacconnaz



- *Profil 2D*

- altitude maximale : 3972 m (Figure A4.4)

- altitude minimale : 1009 m
- dénivelé total : 2963 m
- longueur du couloir : 6237 m du point le plus haut à l'Arve
- pente moyenne : 25.4 degrés
- chenalisation : d'abord faible puis forte
- sinuosité : légère courbure
- profil de pente/concavité : plusieurs ressauts dans la zone de départ potentielle, notamment au niveau de la barre de séracs du glacier de Tacconnaz. Ensuite, pente assez régulière s'atténuant peu à peu (Figure A4.2)
- longueur approximative de zone de départ potentielle : 3700 m
- zone de dépôt : pente moyennes à faible au niveau de l'exutoire du couloir et jusqu'à l'Arve. Depuis les années 80, présence de différents dispositifs paravalanches successifs

- *Fonctionnement présumé de la zone de départ*

La zone de départ potentielle est extrêmement vaste. Selon un critère de pente simple, elle débute à $x=1200$ m et se termine à $x=3700$ m, ce qui correspond assez bien aux observations (les valeurs maximale et minimale enregistrées sont $x=1266$ m et $x=3455$ m). La grande majorité des déclenchements est d'origine glaciaire avec une rupture de la barre de séracs vers 3000-3200 m d'altitude ($x=1600$ m). Néanmoins des déclenchements plus classiques sont aussi possibles plus bas, voire plus haut dans le vaste bassin du glacier de Tacconnaz.

- *Zone d'arrêt, caractéristiques et enjeux*

Très longue zone de pentes moyennes à faibles au pied du couloir et jusqu'à l'Arve. Présence de plusieurs hameaux menacés et/ou touchés à différentes reprises au cours du 20^{ème} siècle.

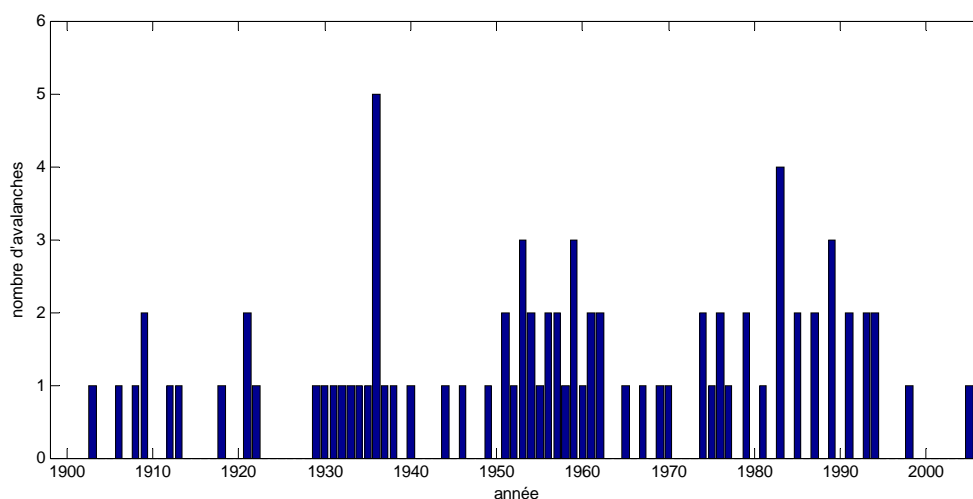
- *Ouvrages de protection*

Présence d'un très important dispositif paravalanche destiné à stocker du volume et dissiper de l'énergie. La construction s'est déroulée en plusieurs étapes à partir de 1984, sans arriver encore à une version tout à fait satisfaisante.

Activité avalancheuse enregistrée :

- *Sources*
 - Base de données EPA (1903-2006)
 - Carnets Forestiers de la commune des Houches
 - Compléments apportés par le RTM74 et le Cemagref (F. Rapin)

Figure A4.3 : site de Tacconnaz, activité avalancheuse enregistrée



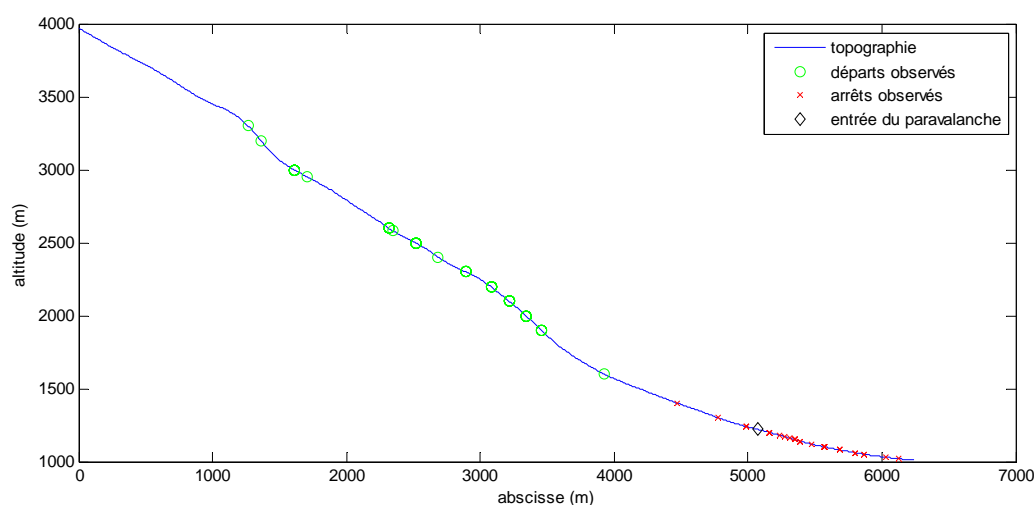
- *Activité et régularité de l'observation*

Au total 83 avalanches ont été enregistrées en 103 campagnes. Au cours d'une même campagne, de zéro à cinq avalanches ont été observées (Figure A4.3). La collecte a vraisemblablement été assez régulière (couloir bien connu avec des enjeux) et l'hypothèse de stationnarité de la collecte au cours des 103 campagnes n'est pas rejetée par un test statistique avec un niveau de confiance raisonnable (Tableau A4.1). Néanmoins certains événements ont pu être manqués, de sorte que la fréquence avalancheuse ainsi obtenue reste sujette à caution.

Tableau A4.1 : site de Tacconnaz, test de stationnarité de l'activité avalancheuse enregistrée

Stationnarité	période	Nb. Camp.	Nb. Aval.	Test	Résultat	p-value
Activité annuelle	1903-2005	103	83	Spearman	0	0.24

Figure A4.4 : site de Tacconnaz, altitudes de départ et d'arrivées projetées sur le profil 2D



- *Altitudes de départ et d'arrêt*

La série d'altitudes d'arrêt est remarquable avec 81 valeurs et une erreur d'observation de plus en plus faible : pour les avalanches les plus anciennes, les valeurs sont regroupées à 100 m près tandis que la précision est de l'ordre de 10 m pour les avalanches les plus récentes. Par contre, une dizaine d'altitudes de départ sont manquantes, principalement dans les années 70-80. Si l'on excepte de plus 3 valeurs douteuses et les données correspondant à la somme de plusieurs événements, il reste une série conséquente de 63 couples altitude de départ-altitude d'arrivée (Figure A4.4).

Les altitudes de départ observées se répartissent sur toute la surface de départ potentielle, mais correspondent très majoritairement à la barre de séracs. Les altitudes de dépôt correspondent principalement aux pentes faibles où est installé le dispositif paravalanche, même si quelques événements s'arrêtent au dessus de l'altitude correspondant à l'entrée du dispositif (Tableau A4.2).

La présence du dispositif introduit une rupture dans le fonctionnement du système puisque sans le dispositif, les avalanches qui se sont produits après sa construction seraient vraisemblablement allées plus loin. Dans la série de 63 couples, seules les deux dernières avalanches sont toutefois concernées puisque toutes les autres sont antérieures à 1984.

Tableau A4.2 : site de Tacconnaz, variables quantitatives enregistrées

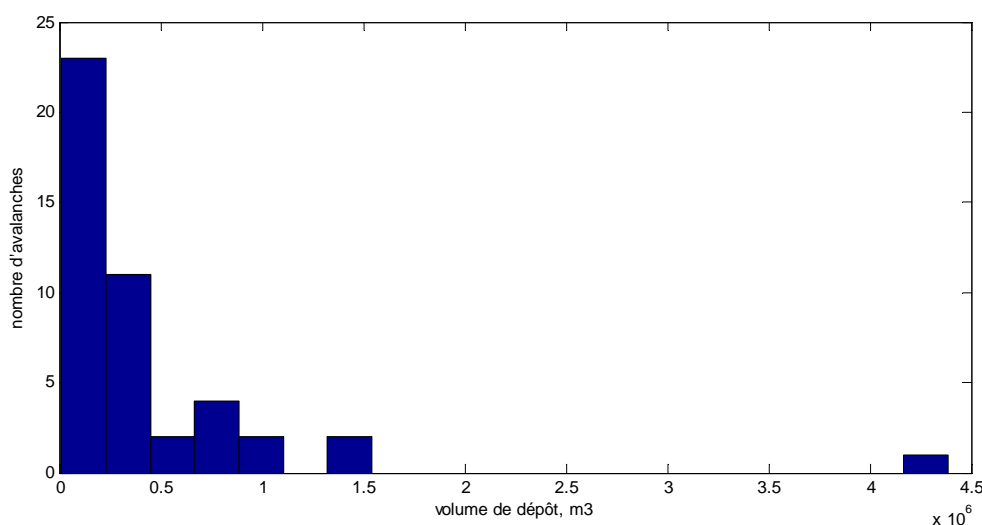
	Nb. valeurs	Moy.	SD.	Min	Max	Skew.	q ₂₅	q ₇₅
z_{start} (m)	63	2379.8	380.4	1600	3300	0.55	2100	2595
z_{stop} (m)	63	1123.8	64.6	1020	1400	1.67	1100	1200
v_{stop} (m ³)	45	422 940	704 160	11 970	4387 500	4.06	60 000	453 125

- *Dimensions des avalanches enregistrées*

49 dimensions de dépôts sont disponibles car consignées dans les carnets forestiers ou estimées par le Cemagref pour les événements récents. Elles correspondent à des avalanches enregistrées durant les campagnes 1906 à 1989. Toutefois seules 45 d'entre elles correspondent à des avalanches pour lesquelles les données d'altitudes de départ et d'arrêt sont satisfaisantes. Les dimensions de dépôt peuvent être converties en volumes de dépôts selon $v_{stop} = \frac{1}{2} L_{stop} \times l_{stop} \times h_{stop}$.

Les volumes d'avalanches sont majoritairement moyens (quelques dizaines de milliers de m³, Figure A4.7). Néanmoins plusieurs avalanches énormes (jusqu'à 1.5 millions de m³) ont été observées avec certitude et une valeur « monstrueuse » de 4.5 millions de m³ est également enregistrée. La validité de cette dernière valeur reste discutée mais le site avalancheux est si disproportionné qu'elle n'est pas totalement dénuée de vraisemblance.

Figure A4.4 : site de Tacconnaz, volumes de dépôt enregistrés



- *Rétro-analyse et modélisation POT*

Les 45 triplets altitudes de départ, altitude d'arrêt et volume de dépôt ont permis la rétro-analyse des 45 événements dans le cadre d'une étude d'ingénierie. Deux séries de 20 événements dépassant des seuils en nombre de Froude et volume de dépôt peuvent en être extraits, cf. chapitre 7.1.

- *Autres données disponibles :*

- Les données qualitatives et catégorielles classiques de l'EPA.
- Les résultats de nombreuses études d'ingénierie plus anciennes.

- *Indépendance des dépassements*

Des événements se produisant durant deux campagnes différentes peuvent être considérés comme physiquement indépendants. On accepte donc l'indépendance des deux séries de 20 dépassements.

Bilan et jeu de données utilisé

Très beau jeu de données par ailleurs déjà très étudié sur un site exceptionnel avec de forts enjeux. Les jeux de données sélectionnés sont les suivants :

- les 45 événements rétro-analysés. On ne retient cependant que les deux séries de 20 dépassements pour le développement du modèle de dimensionnement bivarié.
- Les 83 avalanches en 103 campagnes pour l'évaluation de l'activité annuelle. Toutefois, l'évaluation de la fréquence moyenne au dépassement du volume et du nombre de Froude seuil est préférée puisque les événements les plus intenses ont vraisemblablement tous été enregistrés.

A5. Données simulées sur topographie réelle

Topographie

Correspond au site EPA numéro 5 de la commune de Bessans (73), cf. Annexe A1.

Données

Jeu de 50 données (Figures A5.1 et A5.2) générées avec un modèle de type « bloc glissant » et une loi de frottement de type Voellmy selon une procédure détaillée au chapitre 4.2.3.

Figure A5.1 : altitudes de départ et d'arrivées simulées projetées sur le profil 2D

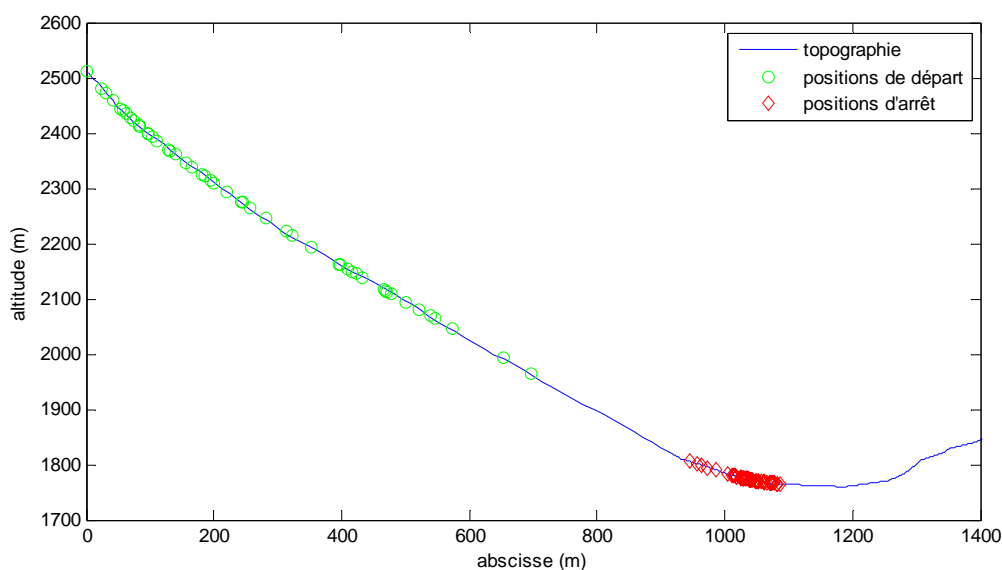
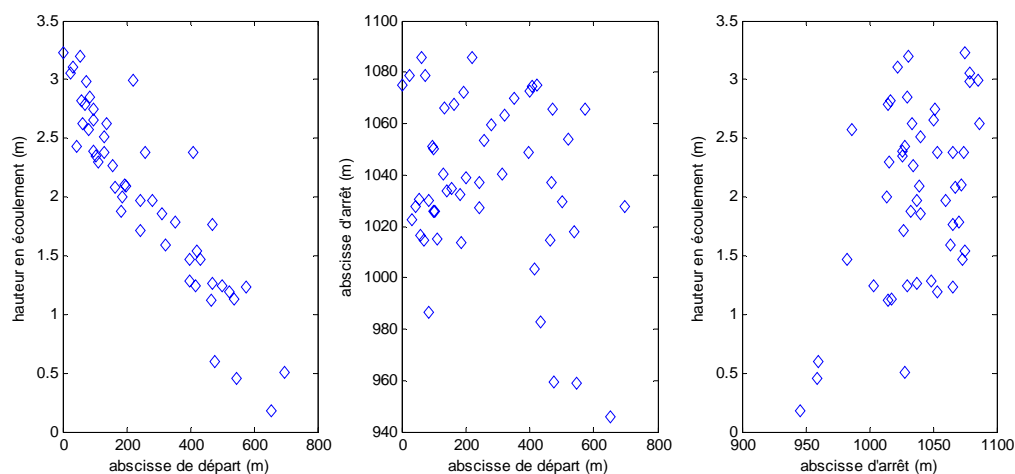


Figure A5.2 : box plots pour données simulées



Bilan et jeu de données utilisé

Ce cas d'étude est particulier car fictif. Le site a été retenu à cause de sa géométrie simple permettant de s'affranchir des problèmes posés par les singularités topographiques et des avalanches ont été simulées avec l'un des modèles stochastiques proposés. Il est utilisé afin de vérifier le bon fonctionnement des algorithmes d'estimation proposés pour les modèles de Voellmy bloc glissant et Saint-Venant.

Couplage données historiques - modélisation numérique pour la prédétermination des avalanches : une approche bayésienne

Résumé :

Cette thèse s'intéresse au problème de la prédétermination des avalanches par couplage entre connaissance naturaliste, modélisation physique et données historiques. L'assemblage des différentes sources d'information et la quantification des incertitudes s'effectuent grâce au formalisme bayésien hiérarchique. La modélisation stochastique, l'inférence du modèle et la prédiction des aléas de référence sont clairement distinguées. L'intensité des avalanches est modélisée indépendamment de leur fréquence définie comme le nombre d'occurrences par hiver. L'approche proposée est située par rapport à la pratique et aux connaissances disponibles pour d'autres risques naturels. L'applicabilité aux avalanches de la notion de période de retour est en particulier discutée.

L'échelle d'étude est tout d'abord celle d'un site avalancheux de l'Enquête Permanente sur les Avalanches. Trois jeux d'hypothèses de modélisation correspondant à différents compromis entre disponibilité des données et description de la propagation de l'avalanche sont proposés. L'inférence des modèles de prédétermination ainsi obtenus est effectuée sur des cas d'étude choisis à l'aide d'algorithmes de simulation ad hoc. L'incertitude liée à la méconnaissance locale du phénomène y est combinée avec sa variabilité intrinsèque pour obtenir la loi de probabilité prédictive de la période de retour associée à n'importe quelle distance d'arrêt. Les distributions des autres variables fournies par les différents modèles sont également étudiées et le calcul des pressions d'impact associées est mené. Une analyse de sensibilité aux différentes hypothèses de modélisation est proposée.

La prédétermination des fréquences avalancheuses sur des sites peu documentés et/ou en contexte non stationnaire est ensuite traitée à l'échelle communale avec un modèle spatio-temporel. Celui-ci permet, outre son résultat opérationnel, de quantifier les variations des occurrences avalancheuses dans les Alpes françaises au cours des soixante dernières années.

Enfin, le problème du dimensionnement d'un ouvrage de protection est abordé, tout d'abord à partir d'un aléa de référence puis dans un cadre décisionnel. Pour ce faire, le modèle stochastique est complété par une loi semi-empirique traduisant les effets sur l'écoulement avalancheux d'une digue verticale et par une fonction de coût permettant l'optimisation de la hauteur de l'ouvrage. Les calculs sont menés analytiquement grâce à des hypothèses simplificatrices puis numériquement. Le calcul du risque bayésien permet de ne pas séparer inférence et décision en prenant en compte l'erreur d'estimation pour le dimensionnement.

Mots clés :

Avalanches de Neige – Prédétermination – Cadre de Travail Stochastique – Données Historiques – Modèle de Propagation Numérique – Modélisation Bayésienne Hiérarchique – Période de Retour – Dimensionnement Optimal