



HAL
open science

Générateur stochastique de séismes en contexte de sismicité faible à modérée : des données à l'aléa. Cas de la France métropolitaine

Corentin Gouache

► **To cite this version:**

Corentin Gouache. Générateur stochastique de séismes en contexte de sismicité faible à modérée : des données à l'aléa. Cas de la France métropolitaine. Géophysique [physics.geo-ph]. Université de Lorraine, 2021. Français. NNT : 2021LORR0136 . tel-03417067

HAL Id: tel-03417067

<https://hal.univ-lorraine.fr/tel-03417067>

Submitted on 5 Nov 2021

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Générateur stochastique de séismes en
contexte de sismicité faible à modérée :
des données à l'aléa.
Cas de la France métropolitaine.

THÈSE

présentée et soutenue publiquement le 25 juin 2021

pour l'obtention du grade de

Docteur de l'Université de Lorraine

Spécialité Géosciences

par

Corentin GOUACHE

Composition du jury :

<i>Rapporteurs :</i>	Prof. Eric CALAIS	ENS, France
	Prof. Fabrice COTTON	Université de Potsdam, Allemagne
<i>Examinatrices :</i>	Prof. Françoise COURBOULEX	Université Côte d'Azur, France
	Dr. Céline BEAUVAL	Université Grenoble-Alpes, France
	Dr. Paola TRAVERSA	EDF, France
<i>Directeur de thèse :</i>	Prof. Jean-Marc MONTEL	Université de Lorraine, France
<i>Co-directeurs de thèse :</i>	Dr. François BONNEAU	Université de Lorraine, France
	Dr. Pierre TINARD	CCR, France
<i>Invité :</i>	Dr. Paul CUPILLARD	Université de Lorraine, France

GeoRessources - UMR 7359

Université de Lorraine – CNRS
École Nationale Supérieure de Géologie
2 rue du Doyen Marcel Roubault - BP 10162
54505 Vandœuvre-lès-Nancy, FRANCE

Résumé court / Short abstract

Résumé court pour le grand public

À défaut de prédire la survenance des séismes, tâche impossible à l'heure actuelle, la communauté scientifique a pris l'habitude de calculer des probabilités annuelles d'atteindre ou de dépasser un certain niveau d'aléa sismique. Pour autant, certains domaines tels que les assurances ont besoin de caractériser plus précisément les séismes qui produisent cet aléa. Une des façons de calculer ces probabilités et de définir clairement les séismes impliqués est de produire des séismes synthétiques sur de longues périodes temporelles (plusieurs centaines de milliers d'années) et d'observer leur impact sur un site donné. Pour se faire, la sismicité passée ainsi que les incertitudes associées sont analysées. Cependant, la France métropolitaine est un territoire éloigné des grandes limites de plaques tectoniques. De ce fait, les moteurs de la sismicité sont peu connus et les données sismiques peu nombreuses, rendant difficile l'analyse de la sismicité passée. Cette thèse propose d'adapter la méthode du générateur de séismes synthétiques au contexte de sismicité faible à modérée. Pour cela, un modèle de calcul de la récurrence des séismes est implémenté et une nouvelle approche de génération des séismes synthétiques dans le temps et dans l'espace est proposée.

Short abstract for general public

Predicting earthquakes isn't an easy task if not impossible. Despite predict these seismic events, scientific community prefers to calculate annual probabilities of exceedance of a given seismic hazard level. Yet, domains like insurance need a clear definition of the seismic events that produce these hazard levels. In order to calculate these probabilities and precisely characterize involved earthquakes, one can generate synthetic earthquakes over large period of time (several hundred thousand years) and observe their impact on a given site. This approach takes into account past seismicity and its uncertainties. However, mainland France is far from major plate boundaries. Thus, the origins of seismicity aren't well known and seismic data are sparse. These facts make difficult to analyse past seismicity. This thesis proposes to adapt the generator methodology for application in low-to-moderate seismic areas. A recurrence model for earthquakes is implemented while a new way to localize synthetic earthquakes in space and time is proposed.

Résumé

La France métropolitaine est éloignée des principales limites de plaques tectoniques. De ce fait, l'origine de la sismicité est plus difficile à comprendre que dans les régions actives localisées le long des limites de plaques. De plus, les données sismiques disponibles (séismes, accumulation des contraintes dans le sol, failles, mécanismes au foyer, etc.) y sont moins nombreuses. Ces observations sont deux des principaux freins rencontrés lorsque l'on souhaite caractériser l'aléa sismique sur un territoire à sismicité faible à modérée.

L'approche proposée dans cette thèse s'appuie sur un générateur stochastique produisant des séismes synthétiques en combinant des observations de terrain et des connaissances théoriques sur la sismicité du territoire d'étude. Ce générateur de séismes se divise en trois parties : (i) le tirage temporel des séismes principaux, suivi du (ii) tirage de leur localisation conditionné par la magnitude et enfin de (iii) la production de répliques associées aux séismes principaux.

Le tirage temporel est réalisé à l'aide d'une distribution fréquence-magnitude (DFM) calculée sur l'intégralité du territoire d'étude via la méthode non paramétrique des temps de latence. Ce calcul à l'échelle du territoire entier permet d'augmenter le nombre de données, de diminuer les périodes de retour et donc d'estimer les fréquences de chaque magnitude directement à partir des données. Une attention particulière a été portée à la méthode des temps de latence afin de pallier la faible quantité de données observées notamment pour les fortes magnitudes. Le tirage spatial des séismes principaux nécessite, quant à lui, une régionalisation et une densité spatiale caractérisant la sismicité. La régionalisation permet de limiter spatialement la magnitude maximale autorisée des séismes synthétiques. La localisation d'un séisme synthétique d'une magnitude donnée est tirée dans la densité spatiale uniquement au sein des régions autorisant cette magnitude. Les répliques sont ensuite générées autour des séismes principaux à partir de la loi de Bâth et de la distribution proportion-magnitude des répliques.

L'aléa associé à chaque séisme (principal et réplique) généré est calculé à l'aide de modèles de prédiction des mouvements du sol pondérés. La pondération est dépendante de la magnitude du séisme et de la distance séisme-site. Cette pondération s'effectue en utilisant la base de données européenne d'accélération du sol RESORCE. Enfin, l'observation en un site donné des aléas produits par les séismes synthétiques sur un million d'années permet d'estimer les probabilités annuelles de dépassement de seuils d'aléa donnés.

Abstract

French mainland seismicity is considered as low to moderate due to its remoteness from tectonic plate boundaries. A first consequence is that the origins of its seismicity are harder to understand than in active regions close to tectonic plate boundaries. Another consequence is the lack of available data (earthquakes recorded but also strain rates, active faults...). These two observations make difficult to estimate seismic hazard in low-to-moderate seismic areas.

The proposed approach is to generate synthetic earthquakes by combining observation and theoretical knowledge on the seismicity of the studied territory. This generator is based on a three-step scheme: (i) the temporal draw of main shocks, (ii) their spatial draw conditioned by magnitude and finally (iii) the generation of aftershocks they produce.

The temporal step needs a recurrence rate. Past seismicity of the whole studied area is analysed thanks to the non-parametric inter event time method in order to obtain this wished recurrence rate. Computing the recurrence rate at the whole territory scale allows to keep the maximum quantity of data, reduce the return periods and so estimate main shock frequencies directly from observed data for each magnitude. An implementation has been developed to overcome the accuracy fall of the inter event time method observed when data are sparse. The spatial step needs a regionalization and a spatial density representing seismicity. The regionalization allows maximum magnitude limitation in space: each region is characterized by an allowed maximal magnitude. Location of a synthetic earthquake with a given magnitude is drawn in the spatial density only within regions that allow this magnitude. Aftershocks are generated around main shocks thanks to the Bath law and the proportion - magnitude distribution of aftershocks.

The seismic hazard produced by each of the generated earthquakes (main shocks and aftershocks) is computed thanks to a set of weighted Ground-Motion Prediction Equations. The weights are obtained as function of magnitude and distance thanks to The European ground-motion database RESORCE. Finally, from direct observation of the seismic hazard produced by synthetic earthquakes over one million years, annual probabilities of exceedance can be calculated with ease.

Abréviations

L'ensemble des abréviations utilisées dans ce manuscrit sont listées ci-dessous en fonction du domaine dans lequel elles s'appliquent.

Catalogue

- FCAT : *French seismic CATalog* [Manchuel et al., 2017]
- SIHex : Sismicité Instrumentale de l'Hexagone [Cara et al., 2015]
- SHEEC : *SHARE European Earthquake Catalogue* [Stucchi et al., 2013]
- N : nombre de séismes dans un catalogue
- T : nombre d'années couvert par un catalogue
- N_{cb} : nombre de catalogues bruités (catalogues obtenus après échantillonnage des incertitudes sur la magnitude des séismes)

Aléa

- PGA : accélération maximale du sol (*Peak Ground Acceleration*)
- PGV : vitesse maximale du sol (*Peak Ground Velocity*)
- a_0 : seuil d'aléa

Méthodes

- PSHA : *Probabilistic Seismic Hazard Assessment*
- DSHA : *Deterministic Seismic Hazard Assessment*

Magnitudes

- M_w : magnitude de moment
- $ErrM_w$: incertitude liée à la magnitude
- M_c : magnitude de coupure, magnitude minimale de l'étude
- A_c : année de coupure liée à M_c
- M_{max} : magnitude maximale du catalogue

- M_{maj} : magnitude majorée, magnitude maximale de l'étude
- $M_{max\ Region}$: magnitude maximale autorisée dans une région
- M_{sp} : magnitude d'un séisme principal
- $M_{sp\ min}$: magnitude minimale des séismes principaux pouvant produire des répliques
- $M_{r\ max}$: magnitude maximale possible des répliques
- ΔM : différence entre M_{sp} et $M_{r\ max}$
- MIMP : *Magnitude of Inferred Main shock Proportion*
- M_x : magnitude de croisement

Moment sismique

- M_0 : moment sismique
- $M_{0\ sp}$: moment sismique libéré par le séisme principal
- $\sum M_{0\ r}$: cumul des moments sismiques libérés par les répliques
- $\sum M_{0\ p}$: cumul des moments sismiques libérés par les séismes précurseurs
- R : ratio entre le cumul des moments sismiques libérés par les séismes dépendants et par le séisme principal

Proportions

- DPM : distribution proportion - magnitude
- p_{sp} : proportion de séismes principaux
- p_r : proportion de répliques
- τ : temps de latence normalisés
- TL : temps de latence

Fréquences

- DFM : distribution fréquence - magnitude
- GR : Gutenberg - Richter [Gutenberg et Richter, 1944]
- a et b : coefficients de la loi GR
- $\lambda_{i\ M_w}$: fréquence annuelle de séismes principaux de magnitude M_w à l'année i
- $n_{i\ M_w}$: nombre des séismes principaux de magnitude M_w générés à l'année i
- PR : période de retour

Modèles de prédiction des mouvements du sol

- GMPE : *Ground-Motion Prediction Equations*
- σ : incertitude liée à la valeur médiane d'aléa calculée par un modèle de prédiction des mouvements du sol
- ϵ : fraction de σ qui est ajoutée ou soustraite à la valeur médiane de PGA
- ϵ_{max} : ϵ maximale de l'étude
- $PGA_{réf}$: PGA de référence
- PGA_g : PGA calculée par la GMPE g
- a_g : poids lié à PGA_g
- N_{GMPE} : nombre de GMPE utilisées
- Am14 : GMPE d'Ameri [2014]
- Ak14 : GMPE d'Akkar et al. [2014a]
- Bi14 : GMPE de Bindi et al. [2014]
- Bo14 : GMPE de Boore et al. [2014]
- C15 : GMPE de Cauzzi et al. [2015]
- DC15 : GMPE de Drouet et Cotton [2015]

Distances

- D : distance entre le séisme et le site d'étude (km)
- D_{JB} : distance Joyner - Boore (km), distance la plus proche entre le projeté du plan de rupture en surface et le site d'étude
- $D_{emprise}$: distance d'emprise spatiale (km), distance du séisme au-delà de laquelle l'aléa est considéré comme négligeable et n'est donc pas calculé

Répliques

- OU : Omori - Utsu [Omori, 1894, Utsu, 1957]
- ETAS : *Epidemic Type Aftershocks-Sequence* [Ogata, 1988]
- $P_r(M_w)$: probabilité de tirer une réplique de magnitude M_w

Analyses des résultats

- En : écart normalisé
- Inc : incertitude liée à la PGA à période de retour donnée [Giardini et al., 2004, Douglas et al., 2014]

Autres études

- G85 : algorithme de *declustering* de Grünthal [1985]
- M2018 : étude française de Martin et al. [2018]
- B2020 : étude française de Beauval et al. [2020]

Remerciements

Tout commença en ce jour de décembre 2017 où se rassemblèrent la partie entrepreneuriale, représentée par Pierre et David, la partie universitaire, regroupant François et Guillaume, et un jeune géolien fraîchement diplômé, moi. Cette réunion avait pour but de lancer un doctorat sur l'étude de l'aléa sismique français entre trois parties qui ne se connaissaient pas ou peu. La clé de voûte de cette rencontre fut Jean-Marc, bon ami de David, qui proposa à CCR l'équipe RING comme cadre pour cette thèse.

Tu es passé, à ma grande et agréable surprise, de mon directeur d'école à mon directeur de thèse. Merci à toi pour ton incitation à publier et à croire en ce que je faisais. Je te remercie aussi pour la confiance que tu m'as accordée en me cédant quelques TD de statistiques.

Après de longs échanges scientifiques et quelques mentions administratives je commençais seulement à me rendre compte de ce que signifiait faire une thèse lorsque Guillaume me demanda si je me sentais prêt et si j'avais les épaules pour réaliser cette thèse. J'avoue que ma vision de la thèse, qui jusque-là se résumait à "un CDD de trois ans", fut grandement perturbée. Ajouté à cela le fort caractère et le franc-parlé de François qui allait devenir un de mes encadrants et me voilà pas mal bousculé.

Je me dois maintenant de rétablir la vérité ! Merci Fifi pour ton engagement permanent dans cette thèse. Bien que l'étude de l'aléa sismique ne soit pas, au début de la thèse, notre domaine de prédilection, tu as su remettre en question beaucoup de mes travaux grâce à ton envie de toujours faire mieux. Je garde de ce passage dans l'équipe RING également énormément de souvenirs extra-travail qui font de toi et ta belle famille (bisous à la chaleureuse Julie et aux petits diables Arthur et Thomas) de très bons amis !

Merci également à Guillaume pour tes bons conseils et bonnes remarques distillés à chacun de mes séminaires. Tu fais un très bon chef d'équipe !

Bien que je fusse déstabilisé par cette nouvelle vision de la thèse, la présence quasi-paternelle de Pierre et David me redonna confiance et j'acceptai fermement cette nouvelle aventure.

Pierre, merci pour ce que tu m'as appris, que ce soit durant le PFE à CCR ou bien au cours de cette thèse. Tu as toujours tout fait pour me rendre la vie plus facile côté CCR et pour te rendre disponible malgré ton emploi du temps chargé. J'ai également passé d'excellents moments avec toi aux pauses à CCR ou encore durant les conférences (AFPS et Journées Cat Nat). Maintenant que je m'apprête à prendre ta place, la boucle est bouclée !

En janvier 2018, me voilà donc parti à la découverte de ce monde complexe qui regroupe l'équipe RING, GeoRessources, le consortium et l'ASGA. Après plus de trois ans passé dans cette équipe je me rends compte de la chance que j'ai eu de travailler dans une telle structure.

Voici donc venir le moment opportun de remercier l'ASGA. L'ASGA étant une association, il est plus juste de remercier celles qui la font tourner à merveille : Armelle et Maryse. Grâce à votre bonne gestion de l'association et du consortium j'ai pu travailler dans les meilleures conditions. Merci également d'être restées accessibles lors de mon séjour parisien.

Sans oublier un énorme merci à nos collaboratrices Sophie, Fatima, Nageoi et Marie qui nous ont délesté des charges administratives. C'est un vrai plus pour l'équipe !

Me voilà plongé dans le bain, dans cette équipe que je connaissais peu. En effet, je faisais figure de mouton noir, moi qui n'ai pas fait Géol Num. Mais cette équipe composée d'anciens de mes enseignants est rapidement devenue un groupe d'amis formidables. Votre sympathie et les pauses tarot y sont pour beaucoup !

Paul, travailler avec toi sur un sujet pour lequel je ne suis pas à l'aise fut un plaisir grâce à ton enthousiasme et ta pédagogie. La bise à Léo et Ulysse que j'ai presque vu naître en ce fameux dernier soir libre de mars 2020 !

Merci Pauline et tes avis bien tranchés qui font ton charme et ont animé nos pauses thé à refaire le monde. Une des plus grandes satisfactions de ma thèse fut que ma soutenance t'ait plu, toi qui as toujours apprécié la venue de cette nouvelle thématique dans l'équipe.

Christine, ou devrais-je dire Mamoune ? Merci pour ta jovialité lors des pauses café et également pour tes cours de programmation qui m'ont permis de réaliser cette thèse. La bise à Denis. La prochaine fois que je viens, je vous prends à la pétanque et aux fléchettes !

Mustapha et Christophe, merci pour votre soutien scientifique et technique.

Au cours des premiers mois, je vois défiler les soutenances des vieux thésards : Marion, Antoine et Gaby.

Vous futes pour moi des références en ce début de thèse. Du plus extravagant

(Antoine), à la plus réservée (Marion), vous m'avez montré qu'il n'existe pas un seul type de thésard et que cela n'a pas d'importance sur la qualité du travail réalisé. Mention spéciale à Gaby qui fut mon co-bureau pendant quelques mois.

Vinrent ensuite les cours donnés aux élèves de l'école de géologie. Que ce soit les stages terrain, les TD de statistique, de géostatistique et de programmation, l'encadrement de Master ou encore un passage éclair à Telecom, ces expériences m'ont permis de découvrir l'enseignement que j'apprécie.

Merci Julien de m'avoir offert l'opportunité de t'accompagner deux fois en stage terrain que j'ai adoré en tant qu'étudiant et encadrant : les Baronnies ! Quelles magnifiques expériences et quels bons moments passés avec toi et Yann.

Merci à Christine, Pierre et Nico pour m'avoir accompagné dans mes premiers TD de programmation.

Encore merci à Jean-Marc (stats) et Guillaume (géostat) de m'avoir proposé des TD.

Merci également à Benoît de chapeauter si magistralement le module TSAS qui ne m'a pas réussi en tant qu'étudiant. Petite pensée aussi à ce fameux concours de tarot !

Célia, merci pour ton travail, ta sympathie et ta persévérance. Il en fallait pour maîtriser Dash ! J'ai appris beaucoup à t'encadrer.

Enfin, merci aux étudiants de l'ENSG pour leur indulgence durant mes TD. Mention spéciale au groupe 4 des Baronnies 2018 : Laura, Léna, Auriane et Pauline, et leur grande maîtrise des projections stéréographiques !

Cette première année de thèse à Nancy fut un concentré de travail mais également de soirées endiablées, que ce soit au bar, chez les uns ou les autres ou encore au mythique Foyer.

Merci à Nico pour ton soutien infailible aux Docks, au Délirium et consorts. Tu resteras pour moi un très bon ami avec qui j'ai refait le monde maintes fois dans ces antres en fréquente compagnie de Melchior et Paul. La bise à Claire et Guizmo.

Merci à tous les thésards de GeoRessources qui nous ont accompagné dans ces folles soirées au bar ou chez l'habitant : Martin, Laurie, Claire, Lucas, Pierrick et Yann (big up Steval !).

Paul, un exemple d'engagement dans l'équipe et dans tout ce que tu fais. On a partagé de précieux moments ensemble : sports (foot, hand, pmu), soirées (bar, appart et Foyer), jeux, vins, tarouges, belidres et j'en passe... Merci de m'avoir ouvert tes portes que ce soit pour boire des canons ou bien m'héberger. Bisous à Claire et son F en géostat ! Pour la petite anecdote, je n'avais pas encore signé de contrat que tu m'embarquais déjà comme cadre A d'une soirée Foyer !

En parlant du Foyer, merci à Judith et Gilles d'avoir maintenu cette institution malgré les nombreuses péripéties auxquelles vous avez du faire face. Merci

également aux étudiants pour leur énergie, leur folie et pour nous avoir mis à contribution lors de ces soirées.

La fin de ma première année rime avec la fin de mon premier séjour nancéien et s'accompagne de la fin du doctorat de Margaux, Pierre, Modeste et Julien.

J'ai pu voir à travers vous le déroulement de la dernière année de thèse. Malgré vos emplois du temps chargé vous gardiez toujours le temps pour discuter avec nous les plus jeunes. Je repense notamment aux discussions sur la géologie de la Guadeloupe avec Margaux et de la sismicité de la vallée du Var avec Pierre. Mention spéciale pour Margaux, ma co-bureau pendant plus d'un an. Encore désolé pour mes performances musicales dans le bureau. Je pense que tu me pardonnes au vu des superbes moments passés ensemble (jeux, sports, bars...).

La deuxième année rime avec Paris et CCR. Au cours de cette année je retrouve une atmosphère bien connue que j'avais quittée en septembre 2017 : le deuxième étage du 157 boulevard Haussmann. L'équipe a peu changé et est toujours aussi agréable. Je suis très heureux de pouvoir continuer mon aventure post-doctorale en votre compagnie.

Merci à vous tous. David, pour ton suivi avisé de la thèse, Dorothee pour nos échanges d'expériences doctorales, les collègues de la R&D (Jean-Philippe, Thomas, Roxane, ZX), les co-bureaux pour vos discussions autour du foot (Eric), de la géostat (Luc) et de MATLAB et HPC (Fred), Margaux pour nos souvenirs de PFE, Nathalie pour ta bonne humeur, les bergamottes et ton aide précieuse sur ArcGIS, Elsa pour notre futur apéro à Port-Brûlé et le reste de l'étage pour votre accueil et votre gentillesse : Georges, Christian, Geoffrey, Chadi, Jérémy, Lise, Monique, Joëlle, Malika, Jelena, Daniel, Antoine, Thierry et Nicolas.

Cette année 2019 a également été marquée par une succession effrénée de conférences entre août et septembre : IAMG en Pennsylvanie, RING meeting à Nancy et AFPS à Strasbourg.

La conférence au Etats-Unis m'ammène à parler de mes deux braves et sympathiques compagnons de route durant ces trois années et des brouettes de thèse : Melchior et Yves.

Melchior tu es fêtard, bon vivant, bon cuisinier, amoureux de la nature, artisan et tu aimes accueillir. Bref, un épicurien et fier de l'être ! Merci pour ton accueil toujours chaleureux et les bons moments passés avec toi.

Yves, ta passion, que ce soit pour les cailloux, les jeux vidéo et les mangas, m'a toujours impressionné. D'ailleurs merci beaucoup pour tous les bons tuyaux en termes de jeux vidéo et mangas ainsi que les soirées jeux vidéo que tu organisais.

Notre point commun à tous les trois est notre passion pour le Japon (rappelons-nous de la Japan expo, de l'Anim'Est ou encore de la soirée umami, sans parler

des nombreux resto japonais). J'espère qu'on pourra partager une coupe de saké tous les trois là-bas !

Le Japon m'emmène naturellement à Jade. Bise japonaise à toi.

Enfin, comment ne pas remercier Julien qui me suit (ou serait-ce moi qui le suis ?) depuis ces belles années de prépa à Carnot ? On a fait un bon bout de chemin ensemble et on a soutenu notre thèse à une semaine d'intervalle ! Merci à toi.

Le retour à Nancy pour la dernière année de thèse (2020) a été marqué par plusieurs événements. La découverte de nouvelles têtes féminines : Capucine et Zoé, ainsi que l'arrivée d'un virus mondialement connu : le COVID. Cette période de confinement nous a permis de se souder les coudes et de partager de bons moments à distance : jeux de société, tarots...

Capucine, merci pour ta gentillesse, tes conseils en tricot, les soirées jeux et les bars passés en ta compagnie. Passe le bonjour à Roro !

Merci Zoé pour ta participation très active et constructive lors des séminaires, ton franc-parlé et la découverte d'un monde qui m'était jusque-là inconnu: la course hippique. Mes salutations à Morgan. Quand je reviens on se fait une partie de Catane !

Paul, tes différents passages au sein de l'équipe RING ont laissé une bonne image du lorrain fier de ses origines rurales que je partage avec toi. Merci, par Jupiter !

Le confinement et surtout le rejet de la première version de ce présent manuscrit ont été des moments difficiles à vivre. Néanmoins, je les ai pris avec beaucoup de philosophie notamment grâce à toi, Salomé.

Merci d'avoir été présente à mes côtés durant ces périodes qui auraient dues être difficiles et éprouvantes mais qui sont finalement passées paisiblement. Merci également de me supporter depuis 10 ans maintenant ! Je t'aime.

Arrive maintenant l'année 2021 et la fin de la thèse est proche. Je me retrouve à la place des anciens et doit montrer la voie aux nouveaux arrivants que sont Fabrice et Enrico.

Merci à vous deux d'être différents, de ne pas venir du système classique ENSG - Géol Num. J'avais initié ce mouvement trois ans auparavant et vous le confirmez désormais. Je pense que l'équipe a besoin de se diversifier de la sorte. Conseil de vieux con : détendez-vous ! Enrico, tes progrès en français m'ont épaté. Tu as été un co-bureau distrayant et direct, l'idéal pour penser à autre chose que la rédaction du manuscrit.

En ce jour du 25 juin 2021, ma thèse prend fin sur une soutenance me permettant de défendre mes travaux de thèse face à un jury d'experts bienveillants et pointilleux.

Merci à Eric, Fabrice, Françoise, Céline et Paola d'avoir accepté d'évaluer mes travaux et pour vos discussions constructives lors de la soutenance. J'aimerais remercier tout spécialement Fabrice qui, de par le rejet de la première version du manuscrit, m'a permis d'améliorer grandement la pertinence et la cohérence de mes travaux vis-à-vis de la communauté de l'étude de l'aléa sismique. Un grand merci également à Céline pour ses échanges constructifs entre la première et la dernière version.

La soutenance est finie, le pot de thèse débute et les vacances commencent. Tout est bien qui finit bien.

J'aimerais clôturer cette longue séance de remerciements par ma famille. Merci à Rozenn et Cloé pour leur soutien de près comme de loin. Enfin, merci à mes parents de m'avoir inculqué des belles valeurs telles que le partage et la joie de vivre. Ces valeurs m'ont permis de suivre mon p'tit bonhomme de chemin et d'"entrer enfin dans le monde du travail".

Table des matières

Résumé court - Short Abstract	i
Résumé	iii
Abstract	v
Abréviations	vii
Remerciements	xi
Introduction générale	1
1 État de l'art sur l'estimation de l'aléa sismique	13
1.1 Introduction	14
1.2 Catalogues de sismicité	15
1.3 Méthodes d'estimation de l'aléa sismique	18
1.4 Mise en oeuvre des méthodes probabilistes	25
1.5 Conclusion	38
2 Du catalogue sismique à la distribution fréquence - magnitude	39
2.1 Introduction	40
2.2 Distribution fréquence - magnitude de séismes principaux	41
2.3 Echantillonnage de l'incertitude autour de la distribution fréquence - magnitude	45
2.4 Discussions et Conclusion	46
3 Générateur de séismes en contexte de sismicité faible à modérée	49
3.1 Introduction	50
3.2 Générateur de séismes principaux	51
3.3 Générateur de répliques	55
3.4 Discussions et Conclusion	62
4 Des séismes générés à l'aléa sismique	65
4.1 Introduction	66
4.2 Calcul de l'aléa sismique	67
4.3 Pondération des GMPE	72
4.4 Discussions et Conclusion	75

5 Application : France métropolitaine	77
5.1 Introduction	78
5.2 Données françaises et paramétrisation	79
5.3 Proportions et fréquences de séismes principaux en France	92
5.4 Génération de catalogues synthétiques français	104
5.5 Calcul de l'aléa sismique français	111
5.6 Estimation de l'aléa sismique en France	118
5.7 Conclusion	132
 Conclusions générales	 135
 Liste des Figures	 144
 Liste des Tables	 145
 Liste des Algorithmes	 145
 Bibliographie	 147

Introduction générale

De tout temps, les tremblements de terre, les éruptions volcaniques ou encore les tsunamis ont fasciné l'Humanité. Cette fascination est due à leur caractère soudain et dévastateur. L'incompréhension et l'injustice ressenties par les Hommes les ont poussés à les interpréter comme des actions de divinités en colère. Les Grecs tremblaient devant Poséidon (Neptune pour les Romains), les Egyptiens voyaient en Geb/Seb le créateur de séismes tandis que les Japonais comparaient leur pays à un poisson-chat (Namazu) qui provoquerait des séismes à chacun de ses mouvements. Cette affiliation divine laisse peu à peu sa place à une approche plus descriptive des événements naturels autour de la science naturelle qui s'intéresse à la Terre : la géologie. Celle-ci n'est apparue que tardivement dans l'environnement scientifique européen. En effet, les trois principes de Nicolas Sténon datent du XVII^{ème} siècle tandis que le mot *géologie* n'entre dans l'Encyclopédie [Diderot et D'Alembert, 1751] qu'en 1751. En 1755, le dévastateur tremblement de terre de Lisbonne fit couler beaucoup d'encre. Rousseau [1756] estime que l'Homme a une part de responsabilité dans cette catastrophe : « *convenez, par exemple, que la nature n'avoit point rassemblé là vingt mille maisons de six à sept étages, et que, si les habitants de cette grande ville eussent été dispersés plus également et plus légèrement logés, le dégât eût été beaucoup moindre et peut-être nul* ». Sans les nommer, il donne ici une description de la *Vulnérabilité* et de l'*Exposition*, deux des trois piliers du risque, défini comme suit:

$$\text{Risque} = \text{Aléa} \times \text{Vulnérabilité} \times \text{Exposition}. \quad (1)$$

Le risque sismique L'*Aléa* est le phénomène naturel en question. En contexte de risque sismique, il correspond à la vitesse ou l'accélération maximale du sol (*Peak Ground Velocity or Acceleration*, PGV ou PGA) produite en surface par un séisme. L'*Exposition* représente la quantité d'enjeux impactés par l'*Aléa* produit par ce séisme. Ces enjeux sont les vies humaines (Figure 1a) et les marchés économiques (Figure 1b). La *Vulnérabilité* définit la capacité des enjeux à résister à l'*Aléa* qu'ils subissent.

Sur la période 1970 - 2019, quatre des six catastrophes naturelles ayant fait plus de 100 000 victimes sont des séismes (Figure 1a). La rareté des puissants séismes implique des coûts assurantiels (Figure 1b) ponctuellement forts mais peu fréquents. À l'inverse, l'ensemble des catastrophes naturelles climatiques (inondations, cyclones, etc.), bien plus fréquentes, provoque des pertes assurées (Figure 1b) plus élevées. Ainsi, contrairement au nombre de victimes qu'ils provoquent, les séismes ne jouent pas les premiers rôles dans les pertes assurées causées par l'ensemble des autres catastrophes naturelles.

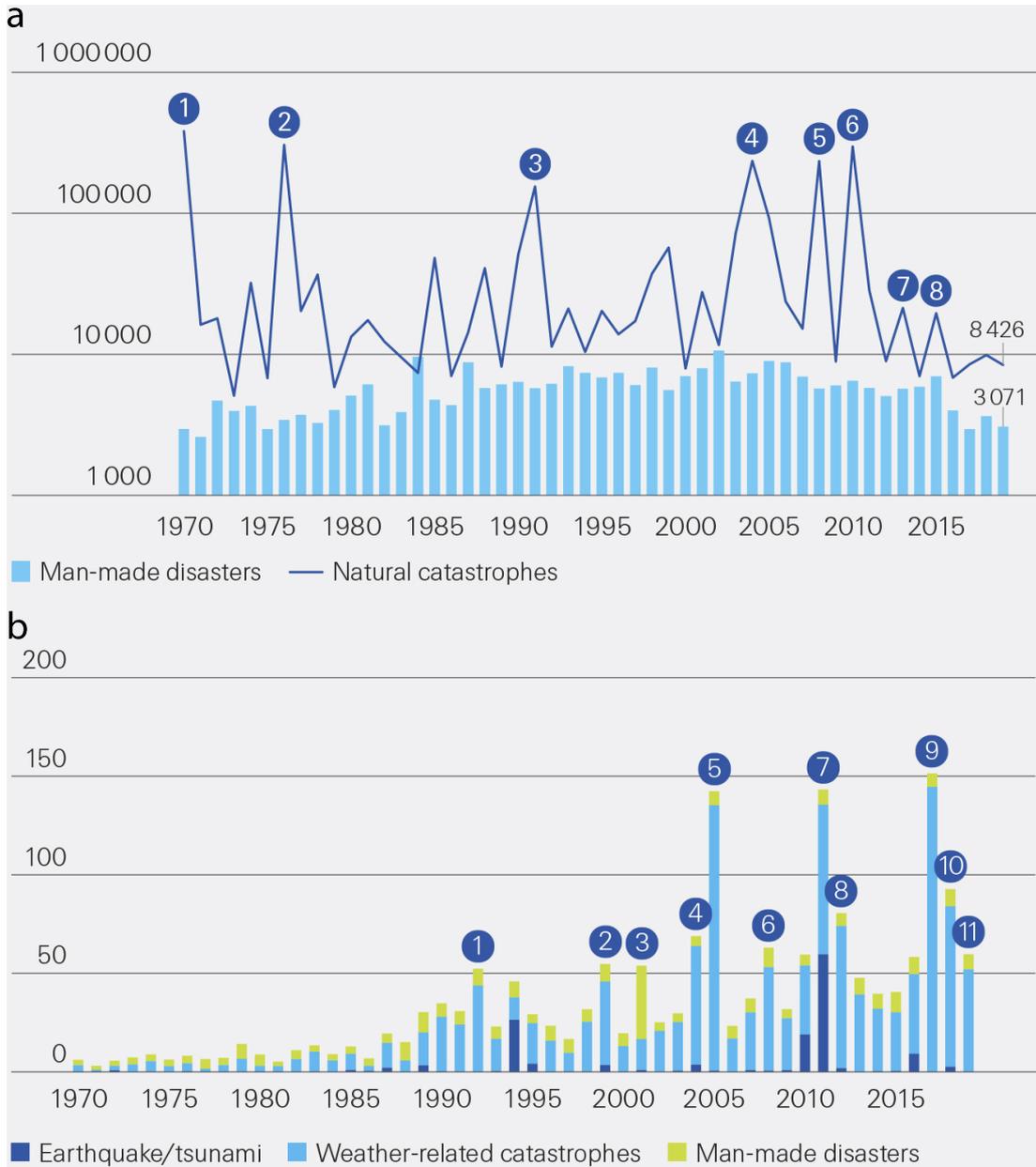


Figure 1: IMPACT DES SÉISMES SUR LA SOCIÉTÉ DE 1970 À 2019 (d'après Bevere et al. [2020]).

a : Bilan humain provoqué par des catastrophes naturelles et anthropiques. Sur les six catastrophes naturelles dépassant les 100 000 victimes, quatre sont des tremblements de terre et/ou tsunamis (1 - Pérou 1970, 2 - Chine 1976, 4 - Sumatra 2004, 6 - Haïti 2010).
 b : Pertes (Md\$) assurées dues aux catastrophes naturelles et anthropiques. Sur cette période, les séismes de Northridge (1994), de Tohoku (2011) et des provinces de Canterbury et de Christchurch (2010 et 2011) ont provoqué ponctuellement de larges pertes.

Ce constat est également fait en France métropolitaine, où l'activité sismique est faible à modérée. Peu de séismes ont été reconnus comme Catastrophe Naturelle depuis l'instauration du régime d'indemnisation assurantiel en 1982. Les coûts assurantiels causés par les six séismes principaux survenus depuis 1982 (Figure 2) s'élèvent à un peu moins de 500 M€, ce qui représente environ 1% de la sinistralité totale. Tout comme à l'échelle planétaire (Figure 1b), la sinistralité sismique française n'a que peu de poids comparée à l'inondation (50% de la sinistralité totale) ou encore la sécheresse (42%)¹. Pour autant, de puissants séismes ont eu lieu sur le territoire français dans le passé. La Caisse Centrale de Réassurance (CCR) et le Bureau de Recherches Géologiques et Minières (BRGM) ont estimé les pertes assurantielles que provoqueraient quelques-uns de ces séismes historiques s'ils survenaient de nos jours :

- Arette (1967) : 120 - 140 M€ [Rey et Tinard, 2017],
- Lambesc (1909) : 1,8 - 2,5 Mds€ [Rey et Tinard, 2015],
- Marais vendéen (1799) : 1,5 - 2,3 Mds€ [Rey et Tinard, 2016].

En plus des séismes historiques actualisés, des scénarios sismiques plausibles peuvent également servir d'étude pour l'estimation des pertes assurantielles. C'est le cas du séisme de magnitude 6,3 au large de Nice proposé par Mouroux et al. [2004]. Les pertes assurantielles estimées par CCR pour un tel scénario sont de 10 à 13 Mds€ dont 3 à 4,5 Mds€ seulement sur la commune de Nice [Rey et Tinard, 2015]. Il est donc nécessaire de prendre en compte le risque sismique même dans des régions caractérisées par une sismicité faible à modérée comme la France métropolitaine.

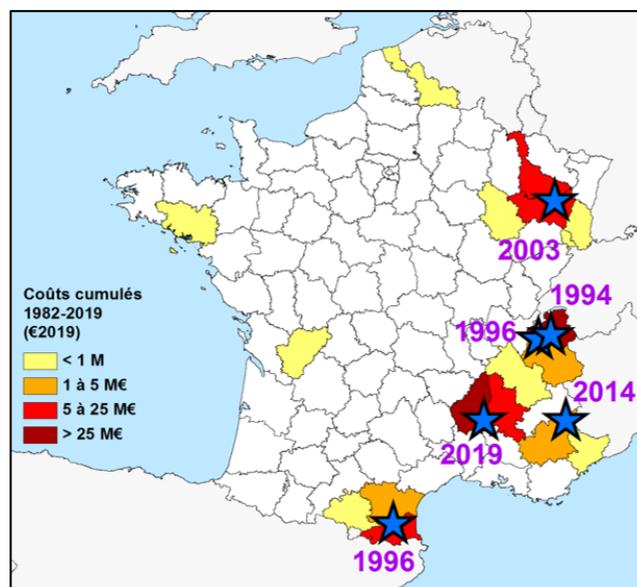


Figure 2: SINISTRALITÉ SISMIQUE EN FRANCE MÉTROPOLITAINE DE 1982 À 2019 (CCR).

Les étoiles bleues correspondent aux emplacements des six plus gros séismes ayant été reconnus Catastrophes Naturelles en France métropolitaine. Chaque département est représenté par les coûts cumulés associés aux pertes dues aux séismes.

¹<https://catastrophes-naturelles.ccr.fr/perils-couverts>

L'origine de la sismicité Comme le rappelle l'Équation 1, la première étape de la définition du *Risque* est une description de l'*Aléa*. La compréhension de l'aléa sismique passe par une compréhension des événements sismiques : les séismes. Ces séismes sont divisés en deux catégories : les séismes indépendants et les séismes dépendants. Les séismes indépendants comprennent la sismicité de fond et les séismes principaux. Les séismes dépendants ou corrélés sont les séismes produits par un séisme principal. Ils sont appelés précurseurs ou répliques suivant s'ils ont eu lieu avant ou après le séisme principal auquel ils sont liés.

Les séismes indépendants sont produits par l'accumulation de contraintes dans le sol. Lorsque ces contraintes dépassent localement la résistance frictionnelle de la croûte terrestre (Figure 3a), alors cette croûte rompt sous la forme d'un plan de rupture. Les plans de ruptures sont localisés le long des fragilités de la croûte, à savoir les failles. Cette rupture fait chuter l'état de contraintes (Figure 3a) et libère de l'énergie sismique qui, en se propageant dans la croûte sous la forme d'ondes sismiques, produira en surface l'aléa sismique. Les séismes dépendants sont, quant à eux, produits par la perturbation locale des contraintes lors du relâchement de l'énergie provoquée par un séisme principal.

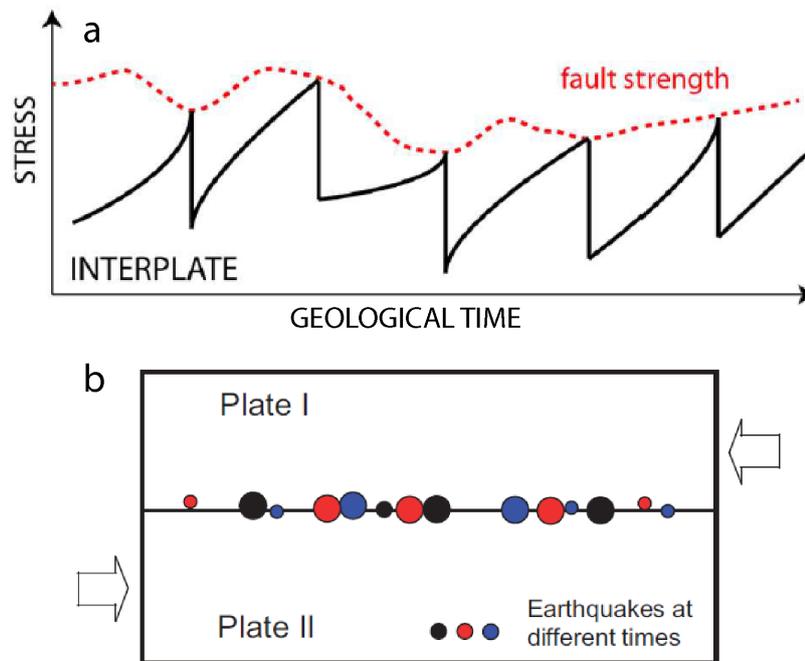


Figure 3: OCCURRENCE SPATIO-TEMPORELLE DES SÉISMES INTERPLAQUES [Calais et al., 2016, Stein et al., 2015].

a : L'occurrence des séismes majeurs en frontière de plaques est principalement régie par l'accumulation des contraintes tectoniques et la résistance frictionnelle de la faille (*fault strength*) [Kanamori et Brodsky, 2004].

b : Quelle que soit la période observée, les séismes majeurs des régions actives sont localisés le long des limites de plaques [Stein et al., 2015].

Les moteurs de l'accumulation des contraintes sont divers et variés. Dans les régions actives, localisées le long des limites de plaques tectoniques, le principal moteur est justement la tectonique des plaques. Ainsi, quelle que soit la période observée, les séismes majeurs de ces régions ont lieu sur les principales failles localisées le long des

limites de plaques tectoniques (Figure 3b). De plus, l'occurrence des séismes suit un cycle sismique régulier qui alterne entre périodes d'accumulation et de relâchement des contraintes tectoniques (Figure 3a). Ce cycle est attesté notamment grâce à la concordance entre les taux de glissement (mm/an) calculés à partir des séismes passés et des données géodésiques actuelles [par ex., Tong et al., 2014, pour la faille de San Andreas].

L'analyse de la sismicité Deux approches sont utilisées pour étudier la sismicité. La première, physique, s'attache à comprendre, expliquer et reproduire la sismicité dans son intégralité : de la rupture des roches à la propagation des ondes sismiques associées. Les modèles mécaniques et géophysiques développés dans cette approche enrichissent notre compréhension sur ce phénomène complexe. La deuxième approche, statistique, est basée sur l'observation de la sismicité passée. Cette observation permet de décrire la sismicité à travers des lois empiriques. L'approche statistique accepte que certains aspects du processus physique soient hors de portée, du moins à des fins pratiques, et doivent être remplacés par un processus aléatoire [Vere-Jones, 2010]. Les deux lois empiriques les plus réputées en sismicité sont celles de Gutenberg et Richter [1944] (GR) :

$$\log_{10}(N(\geq M_w)) = a - bM_w, \quad (2)$$

et d'Omori - Utsu (OU) [Omori, 1894, Utsu, 1957] :

$$n(t) = K(t + c)^{-p}. \quad (3)$$

Ces deux lois décrivent la décroissance du nombre cumulé $N(\geq M_w)$ de séismes en fonction de leur magnitude M_w (Équation 2) et la décroissance temporelle du nombre $n(t)$ des séismes dépendants (Équation 3).

Les deux approches sont complémentaires et ce sont les avancées sur ces deux fronts qui permettent non pas de prédire les séismes [par ex., Geller, 1997, Kagan, 2013] mais de mieux appréhender l'aléa sismique.

Dans la deuxième moitié du XX^{ème} siècle, l'approche statistique d'étude de la sismicité devient de plus en plus mature. Deux méthodes principales de description de l'aléa sismique voient le jour : le *Probabilistic Seismic Hazard Assessment* (PSHA, Cornell [1968]) et le *Deterministic Seismic Hazard Assessment* (DSHA, par ex., Reiter [1990], Kramer [1996]). Ces deux méthodes tentent de décrire l'aléa sismique de manières bien distinctes. Tandis que la première prend en compte l'ensemble des séismes possibles afin de calculer une probabilité annuelle de dépasser un seuil d'aléa, la deuxième se base sur un séisme de référence, considéré comme extrême. Une méthode probabiliste plus compréhensible que le PSHA classique a vu le jour à la fin du XX^{ème} siècle : la méthode des catalogues synthétiques. Cette méthode permet de garder l'avantage de l'approche probabiliste tout en conservant l'aspect événement propre au DSHA.

Problématique liée au contexte intraplaque

Le but de cette thèse est de produire un modèle d'estimation de l'aléa sismique français qui puisse s'intégrer au mieux à un modèle d'étude du risque sismique produit par CCR. Ce modèle de risque sismique vise à estimer l'impact financier des séismes dans un contexte assurantiel dans lequel CCR joue un rôle central via le régime d'indemnisation des catastrophes naturelles propre à la France.

D'après Bommer [2002], la façon d'estimer l'aléa ne doit pas être choisie uniquement en fonction des objectifs et des sorties requises par le risque auquel il appartient. Elle doit également être adaptée aux caractéristiques de la région d'étude et à la quantité et qualité des données disponibles. Or, cette thèse s'intéresse à l'estimation de l'aléa sismique en France métropolitaine. Ce territoire est considéré en grande partie comme une région stable, i.e., éloignée des limites de plaques tectoniques. Deux différences fondamentales pour l'analyse de l'aléa sismique sont observées entre la sismicité intraplaque, qui nous intéresse, et la sismicité des régions actives, présentée brièvement en amont.

La première différence est que le comportement spatio-temporel des séismes majeurs présenté Figure 3 est probablement différent en contexte intraplaque. Tout d'abord, les régions stables étant éloignées des limites de plaques, la tectonique des plaques ne constitue plus le moteur principal de l'accumulation des contraintes. Cet effacement de la contribution de la tectonique des plaques laisse place à d'autres processus complexes qui s'expriment également en contexte de forte sismicité mais sont occultés par la tectonique des plaques. Ces processus divers et variés comprennent les contrastes de densité et de topographie, les déchirures de slab, l'érosion ou encore l'ajustement isostatique [par ex., Mazzotti et al., 2020]. L'ensemble de ces contributions aboutissent néanmoins à une faible accumulation des contraintes. En effet, la France métropolitaine est caractérisée par des taux de déplacement verticaux et horizontaux ne dépassant pas localement les 0,5 mm/an [Nocquet, 2012, Nguyen et al., 2016] sauf dans les Alpes centrales (vertical : 1-2,5 mm/an, horizontal : 0,5-1,5 mm/an, Mazzotti et al. [2020]). À l'échelle d'une faille individuelle, l'accumulation des contraintes est donc considérée comme quasi-nulle (Figure 4a). Dans les régions stables, le cycle sismique (Figure 3a) n'est donc peut-être pas l'hypothèse la plus plausible pour expliquer l'occurrence des séismes majeurs [Calais et al., 2016]. Cette occurrence est causée, à l'échelle d'une faille individuelle, par d'autres facteurs comme la diminution de la résistance frictionnelle de la faille et/ou des variations temporaires de contraintes (Figure 4a Calais et al. [2016]).

De plus, l'accumulation des contraintes, quel que soit le moteur, ne se fait pas le long d'un réseau linéaire de failles parallèles aux limites de plaques tectoniques, comme observé en contexte actif (Figure 3b). Le réseau de failles en région stable est plus également réparti sur le territoire. Sur la base d'observation de 2 000 ans de sismicité chinoise, Liu et al. [2011] proposent que la localisation des séismes majeurs intraplaques varie dans le temps (Figure 4b), passant, à l'échelle d'une faille, de longues périodes de quiescence/dormance à une courte période d'activité [par ex., Liu et al., 2011, Stein et al., 2015]. Le récent séisme du Teil (2019) appuie cette hypothèse. Ce séisme a eu lieu sur la faille de la Rouvière qui date de l'Oligocène (23-34 Ma) et qui n'a pas montré de signe de déplacements depuis au moins plusieurs dizaines de milliers d'années [Ritz et al., 2020]. Ainsi, la sismicité intraplaque sur le long terme est probablement répartie

de manière plus homogène que ce que laisse penser la sismicité actuelle [par ex., Calais et al., 2016].

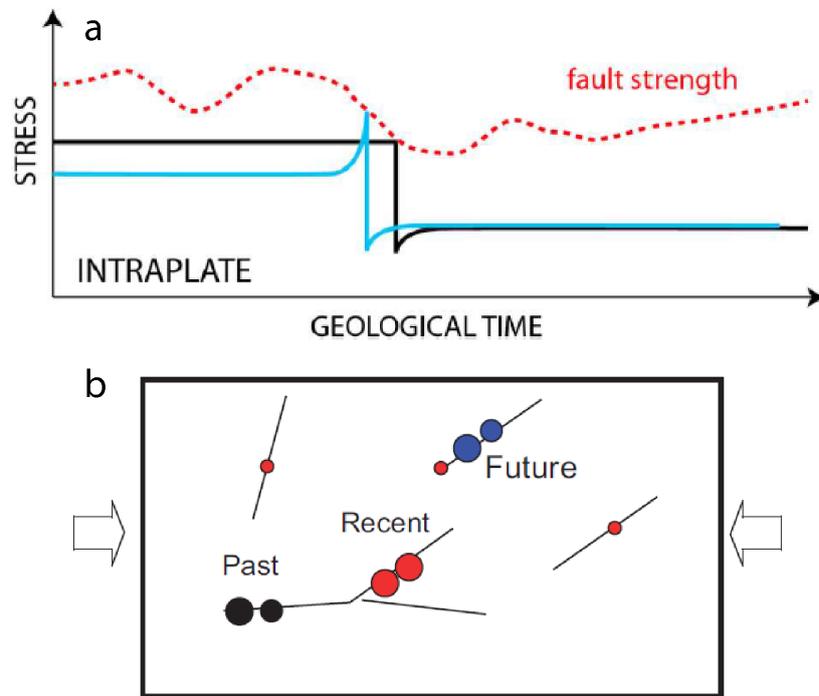


Figure 4: OCCURRENCE SPATIO-TEMPORELLE DES SÉISMES INTRAPLAQUES [Calais et al., 2016, Stein et al., 2015].

a : En contexte intraplaque, l'accumulation des contraintes tectoniques le long d'une faille individuelle est très faible (noir). Dans ce cas, l'occurrence des séismes majeurs est régie par une diminution de la résistance frictionnelle de la faille (rouge) et/ou une perturbation temporaire des contraintes (bleu) [Calais et al., 2016].

b : La localisation des séismes majeurs intraplaques semble changer au cours du temps au sein d'un réseau de failles interconnectées [Stein et al., 2015].

La deuxième différence est que le territoire français dispose de moins de données (séismes, accélérations, mécanismes au foyer, failles actives, etc.) du fait de sa sismicité limitée comparativement aux régions actives. Les lois empiriques et les modèles classiquement utilisés sont donc souvent formulés à l'aide de données provenant de régions plus actives (Figure 5). Ce constat induit deux conséquences principales. La première est que ces lois et modèles ne sont peut-être pas totalement adaptés au territoire français étant donné que le comportement spatio-temporel des séismes majeurs semble différer entre une région active (Figure 3) et une région stable (Figure 4). De plus, ces lois et modèles sont pour la plupart paramétriques. La définition la plus objective possible des paramètres peut s'avérer compliquée à réaliser étant donné le peu de données disponibles. Ainsi, des paramètres calibrés dans les régions actives sont généralement aussi appliqués aux territoires stables sans que cela soit totalement vérifiable. Ces constats rendent encore plus difficile l'estimation de l'aléa sismique en région stable [Stein et al., 2015].

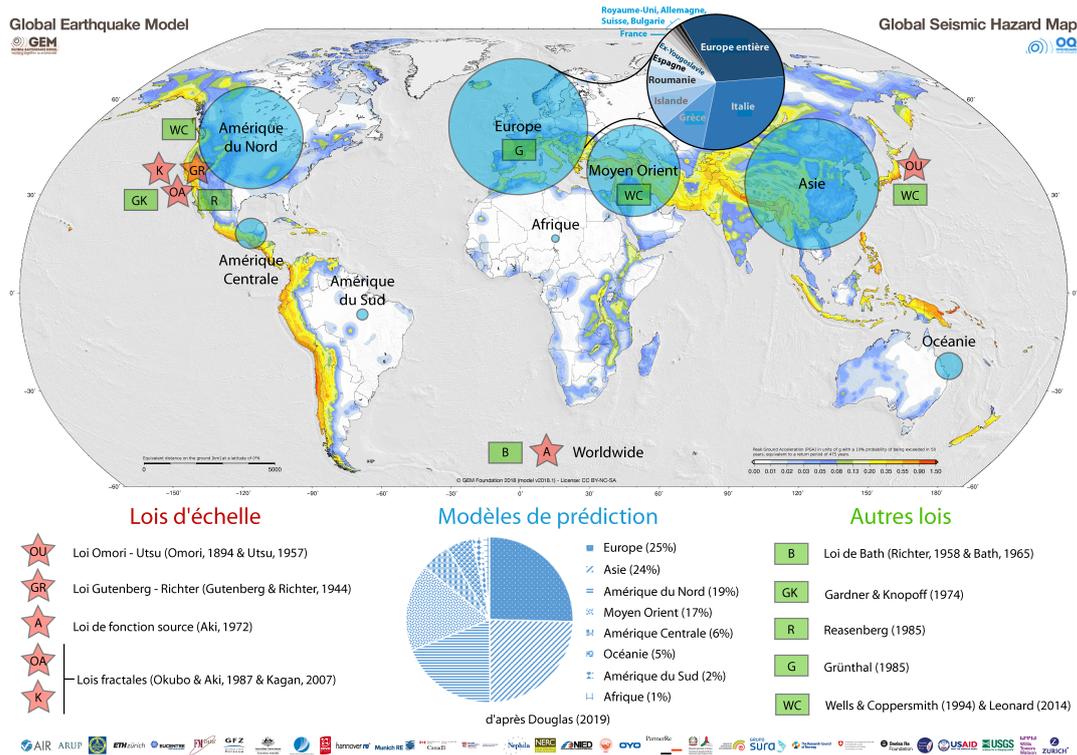


Figure 5: RÉPARTITION MONDIALE DES LOIS ET MODÈLES DÉCRIVANT LA SISMICITÉ. La carte de fond représente l'aléa sismique proposé par le *Global Earthquake Model* (GEM).

La plupart des lois et modèles ont été établis à partir de données provenant de Californie ou du Japon. Les lois et modèles ayant été formulés à partir de données françaises sont rares en raison de la faible sismicité enregistrée sur ce territoire.

La problématique qui se pose alors est : comment décrire l'aléa dans le but d'estimer le risque sismique en France métropolitaine ?

Cette problématique se structure autour de deux axes principaux :

- Comment décrire la sismicité lorsque peu de données sont disponibles ?
- Quelles méthodologies d'estimation de l'aléa sismique s'incorporent bien à l'étude du risque sismique ?

Contributions de cette thèse

Notations: ^(a) *article à comité de lecture*, ^(c) *article de conférence*, ^(o) *présentation orale*, ^(d) *présentation à distance*, ^(p) *poster*.

Exhaustivité et dépendance des données sismiques

Cette thèse a été l'occasion de me questionner sur l'importance des données sismiques et notamment de leur exhaustivité. Le début de ce doctorat a donc été principalement axé sur cette thématique ainsi que sur l'analyse de la dépendance des séismes entre eux. Deux présentations, deux articles de conférence ainsi qu'un poster ont mis en exergue ces travaux préliminaires :

- ^(p,c) C. Louarn, F. Bonneau, et C. Gouache (2019). A web application to analyse seismic catalogue. **RING Meeting 2019**, 10-13 septembre 2019, Nancy, France, ASGA.
- ^(o) C. Gouache, F. Bonneau, et P. Tinard (2019). Regional and Exhaustive French Seismic Catalogue. **20th Annual Conference IAMG 2019**, 12-15 août 2019, State College, USA, S0204.
- ^(o,c) C. Gouache, F. Bonneau, et P. Tinard (2018). Representative seismic zone-dependent catalogue for France. **RING Meeting 2018**, 18-21 septembre 2018, Nancy, France, ASGA.

Pondération de modèles de prédiction des mouvements du sol à partir de données d'accélération

Un workflow a été élaboré durant cette thèse afin de pondérer un ensemble de modèles de prédiction en fonction de la magnitude du séisme et de la distance séisme - site à l'aide de la base de données européennes d'accélération du sol RESORCE. Ainsi, les poids des modèles de prédiction sont obtenus de manière objective et sont dépendant de deux paramètres principaux qui gouvernent le calcul de l'aléa : la magnitude et la distance. Cette approche a été mise en avant lors de présentations orale et poster ainsi que dans un article de conférence :

- ^(d,c) C. Gouache, P. Cupillard, et P. Tinard (2020). Enriching the RESORCE database by synthetic waveforms to weight Ground-Motion Prediction Equations for seismic hazard assessment. **RING Meeting 2020**, 08-11 septembre 2020, ASGA.
- ^(p) C. Gouache, P. Cupillard, et P. Tinard (2019). Calibration de lois d'atténuation par simulation de propagation d'ondes sismiques. **10ème Colloque national AFPS 2019**, 24-26 septembre 2019, Strasbourg, France, ST-26-P-H-5.

Estimation non paramétrique de la distribution fréquence- magnitude de séismes principaux

Au cours de cette thèse, j'ai implémenté une méthode non paramétrique préexistante (temps de latence) afin de pouvoir l'appliquer lorsque les données sont peu nombreuses. Cette implémentation permet de stabiliser les résultats des proportions de séismes principaux pour les magnitudes les plus fortes. Ainsi, l'estimation de manière non paramétrique de la distribution fréquence- magnitude des séismes principaux est rendue possible pour les régions à sismicité faible à modérée. Ce travail a fait l'objet d'un article à comité de lecture, d'un article de conférence et d'une présentation orale :

- ^(a) C. Gouache, F. Bonneau, P. Tinard, et J.-M. Montel (2020). Estimation of main shock frequency – magnitude distributions by adapting the inter event time method for low-to-moderate seismicity areas. Application to French mainland. **Journal of Seismology**, 2021, doi: <https://doi.org/10.1007/s10950-021-10001-8>.
- ^(a,c) C. Gouache, F. Bonneau, et P. Tinard (2019). Stochastic estimation of French annual main shock frequencies. **RING Meeting 2019**, 10-13 septembre 2019, Nancy, France, ASGA.

Générateur spatio-temporel de séismes

Un générateur de séismes a été mis place dans le but de pallier le manque de données sismiques dû aux faibles périodes d'observation comparées aux périodes de retour de séismes dévastateurs. Cette approche permet de générer de la sismicité fictive sur un territoire donné et sur une période de temps souhaitée. Ce générateur se base sur la sismicité passée tout en ne la recopiant pas à l'identique. L'application de modèles de prédiction sur les séismes générés permet d'estimer l'aléa sismique sur des périodes de retour élevées. Ce générateur a été présenté lors d'une conférence internationale et de deux conférences nationales et fait l'objet d'un article en cours de rédaction :

- ^(d) C. Gouache, et P. Tinard (2020). L'estimation de l'aléa sismique en France. **11ème Journée CCR CAT Live 2020**, 15-19 juin 2020.
- ^(d) C. Gouache, F. Bonneau, P. Tinard, et J.-M. Montel (2020). Stochastic generator of earthquakes in French mainland. **EGU General Assembly 2020**, 04-08 mai 2020, EGU2020-5150. doi: [10.5194/egusphere-egu2020-5150](https://doi.org/10.5194/egusphere-egu2020-5150).
- ^(o) C. Gouache, P. Tinard, et J.-M. Montel (2019). Génération stochastique de catalogues sismiques pour le territoire français. **10ème Colloque national AFPS 2019**, 24-26 septembre 2019, Strasbourg, France, ST-25-O-S2-10:40.

Contexte de la thèse

Cadre scientifique et administratif

Cette thèse a été réalisée à l'Université de Lorraine au sein du laboratoire de recherche GeoRessources², dont les objectifs scientifiques sont centrés sur les ressources géologiques, minérales comme énergétiques. Ce laboratoire s'intéresse aux problématiques liées à l'ensemble du cycle de vie de ces ressources (exploration, exploitation, traitement, valorisation et gestion des résidus). Le laboratoire GeoRessources est divisé en plusieurs axes et en plusieurs équipes de recherche. Cette thèse s'intègre dans le projet de recherche de l'équipe RING³ (*Research for Integrative Numerical Geology* – Recherche en Géologie Numérique Intégrative), au sein de l'axe GéoModèles. La recherche de l'équipe RING, dirigée par le Pr. Guillaume Caumon, porte sur le développement de méthodes innovantes en géomodélisation intégrant (i) la modélisation stochastique structurale et stratigraphique, (ii) la simulation stochastique d'objets sédimentaires et diagénétiques, (iii) le maillage adaptatif et la gestion des échelles, et (iv) la simulation de processus physiques.

L'ASGA⁴ (Association Scientifique pour la Géologie et ses Applications) est une association loi de 1901 dont l'objectif est le soutien à l'enseignement et à la recherche en géosciences. Dans ce cadre, cette association gère depuis trente ans un consortium international d'entreprises (majoritairement pétrolières) et d'universités : le consortium RING-Gocad. L'ASGA a été mon employeur durant toute la durée du doctorat.

Cette thèse a été financée par CCR⁵ (Caisse Centrale de Réassurance) via ce consortium. CCR est un réassureur public qui propose depuis 1982 une couverture illimitée aux assureurs avec la garantie de l'Etat des principaux périls suivants : les inondations, la sécheresse géotechnique, les tremblements de terre, les submersions marines et les cyclones dans les territoires d'Outre-Mer. CCR développe des modèles permettant d'estimer l'impact, notamment financier, des aléas naturels en partenariat avec les principaux organismes scientifiques : Météo-France, l'INRAE et l'IFSTTAR pour les risques hydrologiques, le BRGM et l'ENSG pour les risques géologiques. Cette thèse fait suite à un projet de fin d'études réalisé au sein de CCR afin de créer un modèle probabiliste d'aléa sismique pour les territoires français.

La première année de thèse (2018) s'est déroulée à Nancy au sein de l'équipe RING afin d'être aiguillé au mieux dans ma stratégie de recherche et notamment la bibliographie mais également afin de m'imprégner de l'ambiance, de l'atmosphère d'un laboratoire de recherche. La deuxième année (2019) s'est, quant à elle, passée dans les locaux de CCR à Paris, au sein de l'équipe R&D. Cela a été l'opportunité d'appliquer les idées élaborées l'année précédente et de comparer avec les nombreuses données de CCR. Enfin, en 2020, dernière année de thèse, mon retour à Nancy dans l'équipe RING m'a permis de prendre du recul sur les applications faites l'année précédente et de mieux gérer la rédaction d'articles et du manuscrit de thèse.

²<http://georessources.univ-lorraine.fr/>

³<http://www.ring-team.org/>

⁴<http://asgageol.univ-lorraine.fr/>

⁵<https://www.ccr.fr/>

Dans le cadre de cette thèse, j'ai présenté mes travaux lors de deux conférences internationales [Gouache et al., 2019a, 2020b], un colloque national [Gouache et al., 2019d,c], trois éditions de conférences annuelles (RING Meetings) organisées par l'équipe RING et destinées aux membres sponsors du consortium RING-Gocad [Gouache et al., 2018, 2019b, Louarn et al., 2019, Gouache et al., 2020a], ainsi qu'un rendez-vous annuel des assureurs français [Gouache et Tinard, 2020]. Lors de ces conférences, mes travaux ont été présentés sous différentes formes : présentations orales, virtuelles (en contexte de la COVID-19) et poster, et sessions de démonstration et de tutoriel des développements numériques associés à cette thèse. Enfin, un article sur l'estimation de la fréquence-magnitude de séismes principaux en France métropolitaine a été publié dans *Journal of Seismology* [Gouache et al., 2021].

Cadre académique et enseignement

Cette thèse a été l'occasion d'assurer différents cours au sein de l'Université de Lorraine. Tout d'abord et principalement, à l'École Nationale Supérieure de Géologie (ENSG) où j'ai dispensé 122 h de cours variés, réparties entre deux stages de terrain dans les Baronnies (69 h) et des TD de statistiques et géostatistiques (33 h), et programmation informatique (10 h). J'ai également eu l'opportunité d'enseigner 10 h de TD d'initiation aux géosciences à Télécom Nancy au travers d'un projet collaboratif sur l'utilisation des outils de Big Data dans l'analyse de la sismicité. La Figure 6 résume l'entièreté des enseignements dispensés tout au long de la thèse.

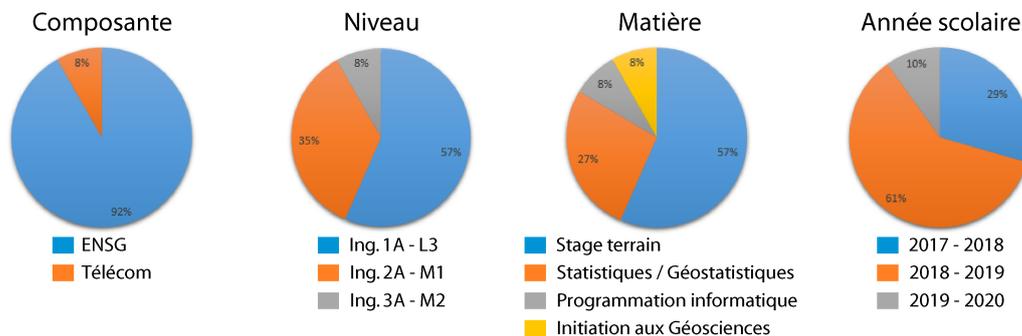


Figure 6: RÉSUMÉ DES ENSEIGNEMENTS DISPENSÉS AU COURS DE LA THÈSE.

L'équipe RING, qui accueille le master Géologie Numérique, m'a également mis à contribution afin d'encadrer le projet de recherche de master de Célia Louarn au cours de l'année scolaire 2018-2019. Son travail consistait à développer une application web en Python permettant d'analyser un catalogue sismique donné en entrée, via différentes lois de sismicité.

Chapitre 1

État de l'art sur l'estimation de l'aléa sismique

Sommaire

1.1	Introduction	14
1.2	Catalogues de sismicité	15
1.2.1	Catalogue instrumental	15
1.2.2	Catalogue historique	15
1.2.3	Catalogue paléosismique	16
1.2.4	Représentativité des catalogues sismiques	16
1.3	Méthodes d'estimation de l'aléa sismique	18
1.3.1	Méthode déterministe	18
1.3.2	Méthode probabiliste intégrative	19
1.3.3	Méthode probabiliste des catalogues synthétiques	21
1.4	Mise en oeuvre des méthodes probabilistes	25
1.4.1	Modèle source	25
1.4.1.1	Le zonage sismotectonique	25
1.4.1.2	La sismicité diffuse	27
1.4.1.3	Les failles actives	28
1.4.2	Modèle de récurrence	28
1.4.2.1	Loi Gutenberg-Richter	28
1.4.2.2	Limites à l'application de la loi GR	30
1.4.2.3	Méthode des temps de latence	32
1.4.3	Génération des catalogues synthétiques	34
1.4.4	Modèle de prédiction des mouvements du sol	35
1.5	Conclusion	38

Résumé Dans cet état de l'art sont présentées les différentes données utilisées lors de l'estimation de l'aléa sismique. Les méthodes d'estimation les plus couramment utilisées sont également citées et expliquées. Ces dernières regroupent la méthode déterministe (*Deterministic Seismic Hazard Assessment*, DSHA) et la méthode probabiliste (*Probabilistic Seismic Hazard Assessment*, PSHA). La méthode probabiliste peut être intégrative ou bien s'appuyer sur des catalogues stochastiques de sismicité générés sur de très longues périodes temporelles. C'est la méthode des catalogues synthétiques. Nous discutons également les avantages et les limites de chacune de ces approches vus sous le prisme du contexte de sismicité faible à modérée et de l'analyse du risque.

1.1 Introduction

L'aléa sismique est le mouvement du sol produit par un séisme. Ce mouvement est classiquement exprimé sous la forme d'accélération et/ou de vitesse maximale du sol (*Peak Ground Acceleration or Velocity*, PGA ou PGV). Les mouvements du sol enregistrés par les accéléromètres sont regroupés dans un catalogue d'accélération et/ou de vitesses. L'intensité macrosismique classe, quant à elle, les conséquences que produisent ces mouvements du sol sur la société (bâti et population). Cette métrique est plus adaptée au risque qu'à l'aléa sismique. Dans cette thèse, nous nous intéressons uniquement aux mouvements du sol et plus particulièrement à la PGA.

L'aléa sismique est un phénomène complexe. Il dépend des caractéristiques intrinsèques au séisme l'ayant engendré (magnitude, dimension de la rupture, type de rupture, etc.), du chemin parcouru par les ondes sismiques (distance, rhéologies traversées, vitesses des ondes, etc.) et du site ayant reçu l'aléa (type de sol, topographie, etc.). L'utilisation des seuls catalogues d'accélération est trop limitée pour décrire ce phénomène complexe. Les études de l'aléa sismique ont donc pris l'habitude de le décrire à travers deux axes : l'étude de la source (le séisme) d'une part et l'étude de la propagation de l'énergie sismique de la source à la surface d'autre part. Dans cette thèse, la notion de scénario sismique comprend la source ainsi que l'aléa qu'elle produit en surface.

Dans ce Chapitre, je présente les trois principales méthodes utilisées depuis la deuxième moitié du XX^{ème} siècle pour étudier d'une part la sismicité et d'autre part la production en surface de l'aléa sismique par la sismicité. Elles comprennent la méthode déterministe, et les méthodes probabilistes intégrative et des catalogues synthétiques, toutes trois présentées Section 1.3. Ces méthodes ont besoin de données afin d'estimer l'aléa sismique. Ces données sont regroupés dans différents catalogues sismiques définis Section 1.2. Enfin, la Section 1.4 présente les modèles requis pour le bon fonctionnement des méthodes d'estimation de l'aléa sismique. Par exemple, le modèle de propagation de l'énergie sismique de la source à la surface en est un.

1.2 Catalogues de sismicité

Un séisme est classiquement représenté par un temps d'occurrence, une localisation dans l'espace et une magnitude. Les séismes sont répertoriés dans un catalogue, appelé catalogue sismique. Ces catalogues sismiques se répartissent en trois catégories suivant la manière de repérer un séisme : instrumental, historique et paléosismique.

Cette Section offre une description brève des trois types de catalogue sismique. Je discute des informations disponibles à travers ces catalogues et les incertitudes qu'ils contiennent. Enfin, je mentionne Section 1.2.4 la complémentarité de ces catalogues ainsi que de la nécessité de traiter les données qu'ils fournissent.

1.2.1 Catalogue instrumental

Comme son nom l'indique, les séismes que contient le catalogue instrumental ont été enregistrés à l'aide d'un réseau d'instruments de mesure. Ces appareils sont appelés sismographes et sont capables de détecter un séisme à partir des mouvements du sol qu'il produit en surface.

Ce catalogue est précis car élaboré à partir d'instruments de mesure. Les incertitudes concernant la magnitude, la localisation ainsi que le temps d'occurrence des séismes qu'il contient sont donc faibles ou assumées. En revanche, il couvre moins de 100 ans de sismicité du fait de l'invention et l'utilisation récentes de tels appareils. Ainsi, ce type de catalogue propose une représentation conséquente et précise des séismes survenus ces dernières décennies.

1.2.2 Catalogue historique

Le catalogue historique regroupe tous les séismes ayant été ressentis et immortalisés à l'écrit par une population. L'étude des textes témoignant de l'évènement sismique et de la description des dégâts sur le bâti par différentes personnes à différents endroits permet d'estimer la localisation et l'intensité macrosismique du séisme. La magnitude peut ensuite être déterminée à l'aide de méthodologies permettant de convertir les intensités macrosismiques en magnitude [par ex, Traversa et al., 2017, Baumont et al., 2018, pour la sismicité historique française]. Si ces récits sont datés, le séisme peut être associé à un temps d'occurrence défini à la minute, à l'heure, au jour, au mois ou à l'année près suivant la précision des dates indiquées. De par leur méthode d'acquisition, les séismes historiques sont associés à des incertitudes plus fortes que celles des séismes instrumentaux.

La qualité et la quantité des informations fournies par le catalogue historique varient grandement d'une région du monde à l'autre. En effet, la profondeur temporelle d'un tel catalogue dépend de la présence de population sachant écrire sur le territoire étudié et de la bonne conservation des écrits au cours des péripéties de l'Histoire. Dans les régions du monde où la présence de l'écriture est ancienne (par ex., Chine, Japon, Méditerranée), ces catalogues historiques peuvent atteindre plusieurs centaines d'années voire plus d'un millénaire. Par exemple, le catalogue historique européen proposé par Stucchi et al. [2013] (SHEEC) couvre une période s'étalant du XI^{ème} au XIX^{ème} siècles.

1.2.3 Catalogue paléosismique

Le catalogue paléosismique rassemble les séismes déduits à partir d'indices de déplacements plus ou moins anciens (jusqu'à plusieurs millions d'années) le long des failles. Ces indices peuvent être, par exemple, des couches géologiques décalées, des escarpements topographiques ou des modifications de cours de rivières. Ces déplacements doivent être assez marquants pour être repérés de nos jours et ne reflètent donc que des séismes majeurs. Malgré les très fortes incertitudes en temps et en magnitude associées à ces catalogues, ils représentent une liste de séismes majeurs passés très utile pour l'étude de l'aléa sismique.

En France métropolitaine, ces indices ont pendant un temps été regroupés au sein du projet NEOPAL [Bertrand et al., 2007], couvrant une période de deux millions d'années.

1.2.4 Représentativité des catalogues sismiques

Représentativité temporelle des catalogues Une des caractéristiques de la sismicité est que la distribution du nombre de séismes en fonction de la magnitude suit une loi de puissance de coefficient proche de 1 : la loi Gutenberg et Richter [1944] (GR, Équation 2). Ainsi, plus la magnitude augmente, plus la période de retour du séisme augmente. Les catalogues possédant une plus grande période d'observation répertorient donc plus de séismes de fortes magnitudes. De cette loi découle la complémentarité observée entre les catalogues sismiques : le nombre de séismes enregistrés augmente drastiquement dans la période instrumentale tandis que les séismes les plus forts, i.e., les plus intéressants pour une étude d'aléa, sont concentrés dans les catalogues historique et paléosismique. Néanmoins, on l'a vu, plus la période d'observation d'un catalogue augmente, plus les incertitudes affectant la magnitude, la localisation et le temps d'occurrence des séismes sont grandes.

Classiquement, la loi GR est donc estimée à partir des magnitudes faibles à modérées fournies par le catalogue instrumental précis. Cette loi d'échelle est ensuite prolongée afin d'estimer la fréquence des séismes aux plus fortes magnitudes. Quant à eux, les catalogues historiques et paléosismiques sont notamment utiles à la définition de la magnitude maximale de la région étudiée ou encore à l'estimation des périodes de retour des séismes aux magnitudes les plus élevées.

Exhaustivité des catalogues Pour des raisons diverses et à des degrés différents, les trois types de catalogue sismique énumérés ci-dessus souffrent du même défaut : la non-exhaustivité des données. L'exhaustivité ou complétude sismique est définie, pour une période d'observation donnée, à partir d'une magnitude de coupure M_c . Tous les séismes du catalogue ayant une magnitude supérieure ou égale à M_c et ayant eu lieu durant la période d'observation donnée sont considérés comme exhaustifs. Il existe plusieurs méthodes de recherche de complétude afin de pallier la non-exhaustivité des catalogues sismiques. Ces dernières sont généralement basées sur la loi GR. Mignan et Woessner [2012] proposent une revue de telles méthodes.

La non-exhaustivité du catalogue instrumental est expliquée par la précision des instruments de mesure mais également par la densité spatiale d'instrumentalisation du territoire étudié. Ces deux aspects évoluent bien évidemment au gré du temps.

Ainsi, la magnitude de coupure d'un catalogue instrumental n'est pas la même de nos jours que celle d'il y a 30 ans. En ce qui concerne les séismes historiques, ils ont dû être ressentis par la population avant d'être portés à l'écrit. Pour cela, la magnitude des séismes doit être forte et/ou la distance à ces séismes doit être faible. À titre d'exemple, un séisme de magnitude 4 ne peut vraisemblablement pas être senti par une population vivant à plus de 100 km. Ainsi, la magnitude de coupure est plus élevée pour le catalogue historique que pour le catalogue instrumental. Bien évidemment, cette magnitude évolue dans le temps au gré de la couverture démographique du territoire étudié. En ce qui concerne le catalogue paléosismique, il est envisageable de définir une magnitude minimale au-dessus de laquelle tous les séismes sont détectables à l'aide d'indices de déplacements. Encore faut-il pour les trouver lancer des programmes de recherche. Or, ces programmes peuvent être longs et coûteux à mettre en place et ne couvrent pas l'entièreté du territoire. L'exhaustivité du catalogue paléosismique est donc bien plus complexe à définir.

Corrélations sismiques Les séismes peuvent être divisés en deux catégories : séisme indépendant et séisme dépendant. La première catégorie concerne la sismicité de fond et les séismes principaux. La deuxième catégorie regroupe les séismes engendrés par les séismes principaux : les précurseurs et les répliques qui ont eu lieu respectivement avant et après le séisme principal. Dans cette thèse, les séismes indépendants sont appelés séismes principaux.

Le comportement de ces deux types de séisme est bien distinct. En effet, il est classiquement admis que l'occurrence temporelle des séismes principaux est représentée à l'aide d'une simple loi de Poisson tandis que la dépendance des séismes est un phénomène bien plus complexe à modéliser (par ex., loi OU présentée en Équation 3). Ainsi, même si les répliques et précurseurs sont finalement pris en compte dans l'analyse de la sismicité, la communauté scientifique a pris l'habitude de séparer les séismes dépendants des séismes principaux.

Pour se faire, il est nécessaire de ségréguer les séismes principaux des séismes dépendants dans le catalogue sismique avant de procéder à son analyse. La méthode la plus usitée pour réaliser cette tâche est l'algorithme de *declustering*. Introduits par Gardner et Knopoff [1974], les algorithmes de *declustering* proposent de différencier les séismes principaux des séismes dépendants à l'aide de fenêtres spatio-temporelles [par ex., Grünthal, 1985, Uhrhammer, 1986, Burkhard et Grünthal, 2009]. D'autres types d'algorithmes définissent les relations spatio-temporelles des séismes à l'aide de paramètres [Reasenber, 1985] telle que la longueur du plan de rupture. Enfin, des algorithmes probabilistes plus complexes reposent sur l'inversion d'un modèle ponctuel marqué [par ex., Zhuang et al., 2002] ou encore sur une approche non paramétrique [par ex., Marsan et Lengline, 2008]. Le catalogue sismique débarrassé des séismes dépendants est appelé catalogue résiduel.

1.3 Méthodes d'estimation de l'aléa sismique

1.3.1 Méthode déterministe

La méthode déterministe [par ex., Reiter, 1990, Kramer, 1996] consiste à définir (magnitude, localisation, géométrie de faille) un séisme de référence jugé extrême. La visualisation de ce séisme concret aide à la compréhension de comment l'aléa a été estimé et de quel type de séisme il conviendra de se protéger. Cet aspect de la méthode déterministe est un avantage pour la communication de l'aléa sismique calculé auprès des ingénieurs, assureurs et décideurs qui ne sont pas forcément experts. Cette méthode est appelée DSHA pour *Deteministic Seismic Hazard Assessment* et est également référencée comme méthode basée évènement.

La définition du séisme de référence s'appuie sur les connaissances géologiques et sismologiques du site d'étude ainsi que sur la sismicité passée, généralement à partir des catalogues historique et/ou paléosismique (Section 1.2). À partir de ces catalogues, sont définis les séismes les plus pénalisants (en termes d'impact au site) susceptibles de se produire dans la région d'étude sur une période d'environ 1 000 ans, soit à peu près la période d'observation du catalogue historique dans les régions bien documentées comme la France métropolitaine. Ces séismes sont appelés Séismes Maximaux Historiquement Vraisemblables (SMHV). Si la présence et la connaissance de failles actives autour du site le permettent, ces SMHV peuvent être définis comme les séismes de magnitudes maximales pouvant être envisagés sur ces failles. Afin de tenir compte des incertitudes inhérentes à la détermination des caractéristiques des SMHV (magnitude, localisation, etc.), une marge de sécurité leur est appliquée pour aboutir au Séisme Majoré de Sécurité. Tous les choix concernant la marge de sécurité ou encore la définition du terme "pénalisant" sont régis par des règles. Pour la sûreté nucléaire française, ces règles sont énoncées dans la Règle Fondamentale de Sûreté [ASN, 2001]¹.

À partir des informations de la source et du site, l'application du modèle de prédiction des mouvements du sol (Section 1.4.4) permet de calculer l'aléa sismique que produit ce scénario extrême sur le site d'étude. Le DSHA étant déterministe, le calcul de l'aléa n'intègre pas l'incertitude aléatoire (σ) associée au modèle de prédiction. Initialement, seule la valeur médiane de l'aléa était utilisée pour calculer l'aléa. Une pratique [Krinitzsky, 2002] permettant de définir un scénario sismique (séisme et aléa qu'il produit) plus pénalisant consiste à considérer le 84^{ème} centile (médiane+ σ) de l'aléa calculé.

Suivant les règles établies, cette méthode a tendance à surestimer l'aléa du site d'étude. Cela a pour conséquence de surdimensionner les ouvrages. C'est pour cette raison que cette méthode est utilisée principalement pour dimensionner des ouvrages critiques tels que les barrages, les ponts ou encore les centrales nucléaires. Sur de tels ouvrages, où les enjeux sont forts, le surdimensionnement ne pose pas de problème et est même source de confiance pour les décideurs et la population. Néanmoins, afin d'éviter des coûts de mise aux normes faramineux, il est inenvisageable d'appliquer un tel surdimensionnement sur l'entièreté du bâti du territoire étudié.

¹<https://www.asn.fr/Reglementer/Regles-fondamentales-de-surete>

1.3.2 Méthode probabiliste intégrative

L'idée principale de la méthode probabiliste intégrative [Cornell, 1968] est de reproduire la variété des sources sismiques qui participent à la production de l'aléa sur le site d'étude. L'ensemble des contributions de ces sources sismiques sont intégrées en une probabilité $P(a > a_0|t)$ de dépasser un seuil d'aléa a_0 sur une période t d'observation donnée. Pour se faire, l'approche probabiliste se base sur l'hypothèse que l'occurrence temporelle des séismes principaux suit une loi de Poisson (Section 1.2.4) :

$$P(a > a_0|t) = 1 - e^{-\gamma(a > a_0)t}, \quad (1.1)$$

où $\gamma(a > a_0)$ est la fréquence annuelle de dépassement de l'aléa a_0 . Cette fréquence annuelle de dépassement est l'inverse de la période de retour associée à l'aléa a_0 . Par exemple, la période de retour 475 ans ($\gamma(a > a_0) = 1/475 = 0,0021 \text{ an}^{-1}$) correspond à une probabilité de dépassement de 10% sur une période d'observation de $t = 50$ ans (Équation 1.1). Cette période de retour est classiquement utilisée dans l'analyse de l'aléa sismique. L'objectif final est d'estimer $P(a > a_0|t)$ en fonction de chaque seuil d'aléa a_0 afin d'obtenir la courbe probabilité-aléa visible Figure 1.1d. Cette méthode est appelée PSHA pour *Probabilistic Seismic Hazard Assessment*.

Pour parvenir à la courbe probabilité-aléa, il convient tout d'abord de distinguer deux types d'incertitudes. Les premières sont inhérentes à un phénomène aléatoire et ne peuvent pas être réduites par l'acquisition de nouvelles données ou connaissances. Ces incertitudes peuvent en revanche être quantifiées sous la forme de distributions à l'aide d'observation dudit phénomène. Elles sont appelées incertitudes aléatoires ou naturelles. Elles comprennent par exemple la localisation des séismes, la géométrie des plans de rupture (azimut, pendage, dimensions, etc.), l'incertitude des modèles de prédiction (σ , Section 1.4.4), etc. Les deuxièmes résultent du manque de connaissances concernant les modèles et les paramètres utilisés pour estimer l'aléa. Elles sont appelées incertitudes épistémiques. Les choix du type de source (Section 1.4.1), du modèle de prédiction des mouvements du sol ou encore de la magnitude maximale autorisée produisent de telles incertitudes.

Les incertitudes aléatoires sont directement intégrées dans le calcul de la courbe probabilité-aléa (Équation 1.2). L'utilisation des mêmes modèles et données d'entrée (et leurs incertitudes aléatoires) aboutit toujours à la même courbe probabilité-aléa (Figure 1.1d). En revanche, les incertitudes épistémiques sont propagées à travers un arbre logique [Kulkarni et al., 1984, Coppersmith et Youngs, 1986]. Sur un même niveau de ramification, chaque branche de l'arbre propose une donnée d'entrée, un paramètre ou un modèle différent des autres branches. Chaque branche est associée à un poids reflétant la confiance attribuée à ce modèle, ce paramètre ou cette donnée. Au bout de chaque extrémité de l'arbre est calculée une courbe probabilité-aléa à l'aide d'une combinaison spécifique de modèles, paramètres et données. Enfin, une courbe moyenne peut être estimée en effectuant la moyenne pondérée de toutes les courbes obtenues en extrémité de l'arbre. Des statistiques comme les centiles peuvent également être calculées, permettant de définir la confiance attribuée à cette courbe moyenne.

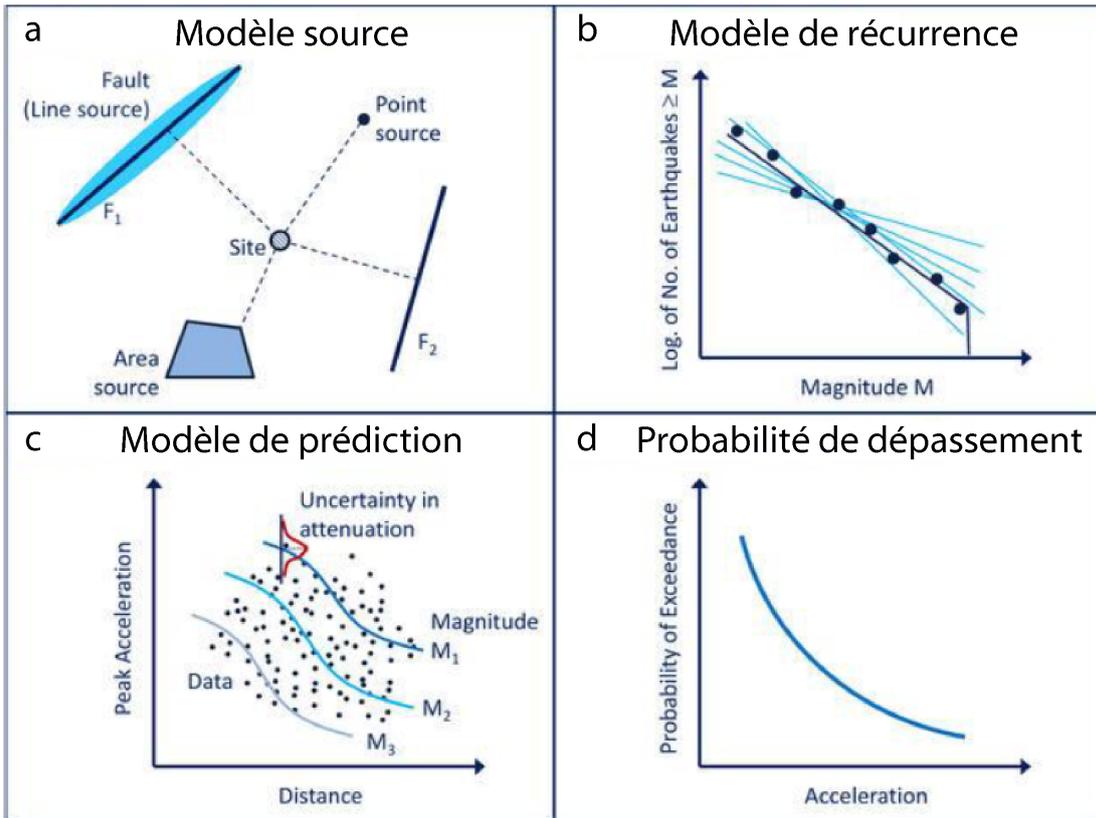


Figure 1.1: LES QUATRE ÉTAPES PRINCIPALES DU PSHA [Cornell, 1968] [d'après Chartier, 2019].

a : Le modèle source permet de définir la géométrie des sources sismiques (faille, zone homogène, sismicité diffuse) et la distribution spatiale des séismes au sein de ces sources.
 b : Le modèle de récurrence calcule un ensemble de distributions fréquence- magnitude afin de définir l'occurrence des séismes dans le temps en fonction de leur magnitude.
 c : Le modèle de prédiction des mouvements du sol estime, à l'aide de son incertitude aléatoire σ , la distribution des aléas sismiques que peut produire un séisme de magnitude M_w sur un site situé à une distance D .
 d : Enfin, l'ensemble des contributions sont intégrées au sein des Équations 1.2 et 1.1 afin de calculer la courbe probabilité- aléa.

Pour calculer une courbe probabilité- aléa (Équation 1.1), il est nécessaire d'estimer la fréquence annuelle $\gamma(a > a_0)$ de dépassement d'un seuil d'aléa a_0 comme suit :

$$\gamma(a > a_0) = \sum_i \left[\nu_i \int_{M_w} \int_D P(D)P(M_w)P(a > a_0|M_w, D) \right], \quad (1.2)$$

où la somme est réalisée sur l'ensemble des sources sismiques considérées. ν_i est le nombre annuel de séismes de magnitude supérieure ou égale à la magnitude minimale de l'étude dans la source sismique i . $P(D)$ est la probabilité qu'un séisme de la source i soit localisé à une distance D du site. $P(M_w)$ est la probabilité qu'un séisme de la source i ait une magnitude M_w . Enfin, $P(a > a_0|M_w, D)$ est la probabilité que l'aléa produit par un séisme de magnitude M_w et situé à une distance D du site dépasse le seuil d'aléa a_0 .

Pour calculer chaque terme de l'Équation 1.2, le PSHA s'articule autour de quatre étapes illustrées Figure 1.1. Chacune de ces étapes repose sur un modèle bien précis. L'objectif de ces modèles est présenté ci-dessous tandis qu'une description détaillée de leur fonctionnement est proposée Section 1.4.

En premier lieu, la géométrie de la source sismique et la localisation des séismes susceptibles d'impacter le site doivent être définies via un modèle source. Ce modèle correspond à la Figure 1.1a. Une fois le type de source choisi (faille, zone homogène, sismicité diffuse), la distribution spatiale des séismes au sein de cette source est estimée à partir des données observées et de leurs incertitudes aléatoires. Cette distribution permet de calculer la probabilité $P(D)$ (Équation 1.2). Ensuite, la fréquence des séismes en fonction de leur magnitude est à établir au sein de la source. Cela est réalisé à l'aide d'un modèle de récurrence, illustré Figure 1.1b. Toujours à partir des données observées et de leurs incertitudes aléatoires, une distribution fréquence-magnitude est estimée. De cette distribution découle le nombre annuel ν_i de séismes de magnitude supérieure ou égale à la magnitude minimale de l'étude au sein de la source sismique. Cette distribution permet également de définir la probabilité $P(M_w)$ (Équation 1.2). Par la suite, la distribution des aléas a qu'un séisme de magnitude M_w peut produire à une distance D du site est échantillonnée à l'aide du modèle de prédiction des mouvements du sol et de son incertitude aléatoire σ (Figure 1.1c). Au sein de cette distribution, la proportion d'aléas a supérieurs au seuil a_0 donne la probabilité $P(a > a_0 | M_w, D)$ de l'Équation 1.2. La dernière étape (Figure 1.1d) consiste à intégrer l'ensemble des contributions des sources sismiques afin de calculer la fréquence annuelle de dépassement $\gamma(a > a_0)$ (Équation 1.2) puis la probabilité $P(a > a_0 | t)$ de dépasser un seuil a_0 d'aléa sur une période de temps donnée (Équation 1.1).

La prise en compte du temps et des incertitudes aléatoires rend cette méthode plus robuste et réaliste que la méthode déterministe présentée Section 1.3.1. Elle permet de considérer non pas un seul cas dimensionnant mais l'ensemble des scénarios sismiques possibles. En revanche, de par son aspect intégratif, le PSHA perd le côté pédagogique du DSHA, à savoir d'associer un aléa sismique à un séisme concret. En effet, le PSHA ne calcule pas la probabilité de dépassement à l'aide de sources identifiables mais via l'ensemble des sources possibles. De ce fait, la méthode probabiliste est souvent qualifiée de "boîte noire", n'aidant pas forcément les décideurs dans leurs choix.

Une alternative à cette approche intégrative a été proposée afin de rendre les résultats probabilistes plus compréhensibles pour un public non-expert. Cette méthode est appelée méthode des catalogues synthétiques.

1.3.3 Méthode probabiliste des catalogues synthétiques

La méthode Monte Carlo ou des catalogues synthétiques [Musson, 1998, 1999] est également une méthode probabiliste. Dans la suite de ce manuscrit, afin de distinguer la méthode des catalogues synthétiques de la méthode intégrative de Cornell [1968], le terme PSHA sera exclusivement réservé à cette dernière. Le terme "catalogues synthétiques" est préféré au terme "Monte Carlo" car ce dernier fait écho à un processus plus général de propagation des incertitudes. La méthode des catalogues synthétiques consiste non pas à intégrer l'ensemble des séismes possibles mais à en simuler stochas-

tiquement sur une période temporelle donnée. Les catalogues synthétiques regroupant ces séismes synthétiques forment alors une base de données représentative de la sismicité du territoire étudié. Pour générer ces catalogues synthétiques, les mêmes données d'entrée et modèles (Section 1.4) que le PSHA peuvent être utilisés. Si c'est le cas, la méthode des catalogues synthétiques aboutit exactement [par ex., Musson, 1998, 2012] à la même courbe probabilité-aléa moyenne obtenue via le PSHA (Figure 1.1d). Seule la manière de l'obtenir change.

L'observation directe des effets des séismes synthétiques sur de très longues périodes temporelles permet d'obtenir la probabilité de dépassement (Figure 1.2). La procédure est simple. Dans l'exemple Figure 1.2, dix millions d'années sont simulées. Pour chaque année simulée, l'aléa maximal produit par l'ensemble des séismes synthétiques sur le site est répertorié dans une liste. Cette liste représente donc un catalogue d'accélération (PGA) maximales annuelles. Contrairement aux catalogues d'accélération établis sur des observations réelles qui ne couvrent au mieux que quelques décennies, celui-ci s'étend sur dix millions d'années. Cette liste d'aléas est ensuite triée par ordre décroissant. L'aléa qui a une probabilité annuelle d'être dépassé égale à $1/10\,000$ correspond au 1001^{ème} aléa de la liste d'aléas maximums annuels. Autrement dit, cet aléa a été dépassé au moins 1 000 fois sur les dix millions années d'observation, soit une probabilité annuelle d'être dépassé égale à 10^{-4} .

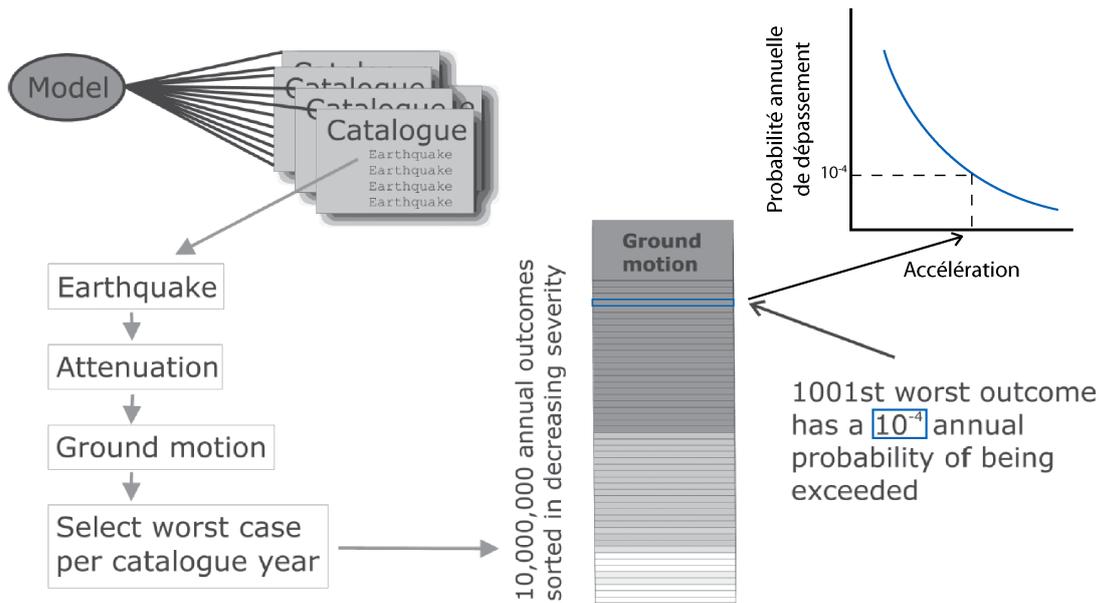


Figure 1.2: CALCUL DE LA PROBABILITÉ ANNUELLE DE DÉPASSEMENT VIA LA MÉTHODE DES CATALOGUES SYNTHÉTIQUES [d'après Musson, 2012].

Dans cet exemple, un ensemble de N catalogues synthétiques de T années sont générés. Le nombre total d'années simulées est de dix millions ($N \times T$). Les dix millions d'aléas maximums annuels sont triés par ordre décroissant. L'aléa qui a une probabilité annuelle d'être dépassé égale à 10^{-4} correspond au 1001^{ème} aléa de la liste. L'aléa trouvé est enfin associé à la probabilité 10^{-4} dans la courbe probabilité annuelle-aléa. Cette courbe est entièrement construite en répétant l'opération pour différentes probabilités.

Les avantages d'une telle méthode sont nombreux. Tout d'abord et tout comme la méthode PSHA (Section 1.3.2), elle est probabiliste de part la prise en compte des incertitudes aléatoires. Cela la rend plus réaliste que la méthode déterministe (Section 1.3.1).

Ensuite, elle est plus facile à comprendre grâce à son aspect concret. Elle ne nécessite pas de formulations mathématiques complexes comme l'intégration des sources sismiques (Équation 1.2). Il n'est pas difficile de comprendre que la probabilité annuelle de dépassement d'un aléa donné ayant été dépassé dix fois sur 10 000 années simulées est égale à 10^{-3} . De plus, les catalogues synthétiques peuvent servir de données d'entrée pour des modèles d'estimation d'autres aléas naturels tels que les glissements de terrain, la liquéfaction ou encore les tsunamis [par ex., Goda et De Risi, 2017].

Enfin, l'aspect concret couplé à l'approche probabiliste permet de quantifier la contribution des séismes sur l'aléa sismique notamment en termes de magnitude et de distance [Bommer et al., 2002]. Par exemple, supposons qu'en un site donné la PGA ayant une probabilité annuelle de dépassement de 10^{-3} soit égale à 2 m/s^2 . Il suffit d'observer l'ensemble des séismes ayant produit une PGA supérieure ou égale à 2 m/s^2 et d'en analyser les distances au site et les magnitudes. Ces contributions sont des informations importantes pour l'analyse du risque sismique. Il est à noter que le calcul des contributions peut également être réalisé en désagrégant la courbe probabilité-aléa du PSHA [par ex., McGuire, 1995, Romeo et Prestininzi, 2000]. Cependant, cela nécessite l'ajout d'un modèle et de ses incertitudes, alors que ce calcul est naturel dans la méthode des catalogues synthétiques.

En revanche, le plus grand inconvénient de cette méthode est la nécessité de produire les catalogues synthétiques sur de très longues périodes temporelles. En prenant N le nombre de catalogues synthétiques et T leur durée en années, Musson [2012] a montré que pour reproduire à l'identique la courbe probabilité-aléa du PSHA jusqu'à la probabilité 10^{-7} , il faut que le nombre d'années totales simulées ($N \times T$) soit au moins égale à un milliard. Musson [1999] attribue à Rosenhauer [1983] la première mention de la méthode des catalogues synthétiques mais constate qu'elle est peu utilisée et associe ce fait à la lourdeur des calculs.

Néanmoins, depuis le début du XXI^{ème} siècle, la puissance de calcul a augmenté et a donc permis une plus large utilisation et des adaptations diverses et variées de cette méthode [par ex., Ebel et Kafka, 1999, Beauval et al., 2006, Han et Choi, 2008, Bommer et al., 2015, Bourne et al., 2015, Shahbazi et Mansouri, 2019, Milner et al., 2021]. De plus, ce nombre d'années simulées dépend principalement de la probabilité minimale ou de la période de retour maximale que l'utilisateur souhaite analyser. En effet, Assatourians et Atkinson [2013] conseillent d'utiliser des catalogues synthétiques d'une profondeur temporelle totale ($N \times T$) au moins 1 000 fois supérieure à la période de retour (inverse de la probabilité) analysée. Ainsi, pour la période de retour 475 ans, $N \times T$ doit être au moins égal à 475 000 ans et non un milliard d'années.

Enfin, certains auteurs [par ex., Assatourians et Atkinson, 2013] ne pensent pas qu'il soit primordial de distinguer les incertitudes aléatoires et épistémiques. D'une part, parce que la distinction entre ces deux types d'incertitude n'est pas forcément

aisée. D'autre part, parce que, d'un point de vue pratique, l'élaboration d'un arbre logique n'est pas innée [Bommer et al., 2005]. De plus, dans certains cas, le nombre de branches peut être important (plus de cent), augmentant considérablement les temps de calcul.

Assatourians et Atkinson [2013] proposent une alternative à l'utilisation des arbres logiques. Ils produisent un catalogue synthétique d'une très longue durée (par ex., dix millions d'années Figure 1.3) en tirant suivant des poids tous les paramètres, données et modèles nécessaires à la génération d'un séisme synthétique (localisation, profondeur, récurrence, magnitude maximale, etc.). Il en est de même pour le calcul de l'aléa à partir de chaque séisme synthétique : pour ce même catalogue synthétique, l'ensemble des modèles de prédiction des mouvements du sol sont utilisés à travers une combinaison linéaire dont les coefficients sont leur poids respectif. Ainsi, les aléas calculés témoignent des deux types d'incertitudes. La courbe moyenne (Figure 1.3) résulte directement de l'observation des aléas maximaux annuels, comme présenté Figure 1.2. Ces auteurs affirment que la courbe moyenne ainsi obtenue est identique à celle obtenue par l'arbre logique. Les centiles associés à l'aléa d'une période de retour donnée sont en revanche obtenus comme suit. La liste des aléas maximaux annuels est divisée en sous-listes de taille égale à la période de retour souhaitée (Figure 1.3). L'aléa maximal obtenu dans chacune des sous-listes représente une réalisation de l'aléa associé à ladite période de retour. Le tri de ces représentations donne les centiles de l'aléa recherché (Figure 1.3). À l'inverse de la courbe moyenne, les courbes centiles peuvent différer des courbes centiles obtenues via l'arbre logique [Assatourians et Atkinson, 2013].

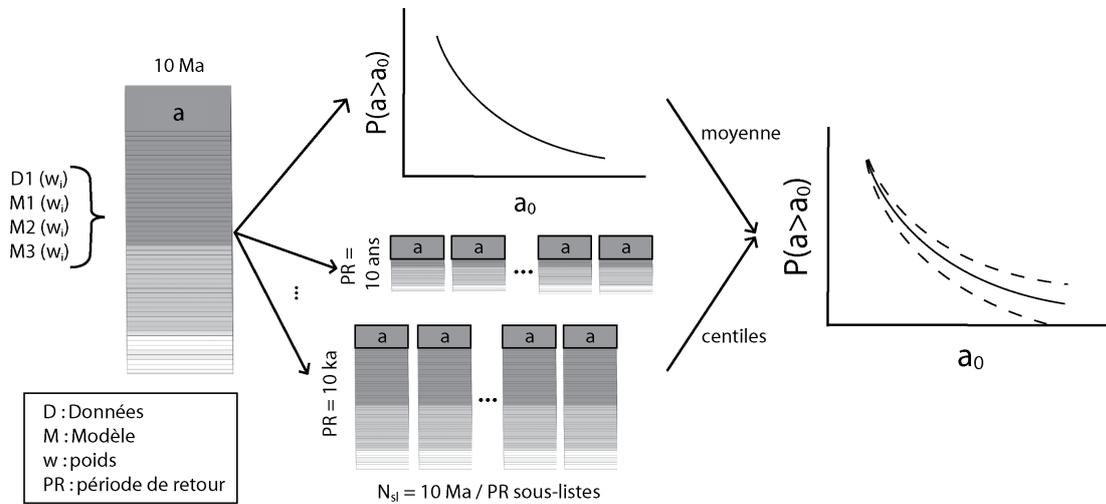


Figure 1.3: VARIABILITÉ DES COURBES PROBABILITÉ - ALÉA.

Toutes les données (D), tous les modèles (M) et tous les paramètres nécessaires à l'élaboration de la liste des aléas maximaux annuels sont sélectionnés avec leur poids respectif. À partir de la liste des aléas maximaux annuels triée dans l'ordre décroissant, la courbe moyenne probabilité-aléa est obtenue (Figure 1.2). En divisant cette liste en N_{sl} sous-listes de taille égale à la période de retour (PR), il est possible de calculer les centiles de l'aléa associé à cette période de retour. En répétant l'opération pour chaque PR souhaitée, les courbes centiles de probabilité-aléa sont associées à la courbe moyenne.

1.4 Mise en oeuvre des méthodes probabilistes

1.4.1 Modèle source

Le modèle source établit tout d'abord la géométrie des sources sismiques : zones homogènes (Section 1.4.1.1), sismicité diffuse (Section 1.4.1.2) ou failles actives (Section 1.4.1.3). Ensuite, au sein de chaque source sismique, la distribution de la localisation des séismes est définie. Cette distribution donne la probabilité $P(D)$ (Équation 1.2) d'obtenir un séisme à une distance D du site d'étude.

1.4.1.1 Le zonage sismotectonique

Initialement [Cornell, 1968], le modèle source est défini par un ensemble de zones sismotectoniques au sein desquelles les régimes géodynamiques ont des comportements tectoniques comparables et une sismicité uniforme. En d'autres termes, la distribution spatio-temporelle de la sismicité ainsi que les paramètres de faille sont homogènes au sein d'une même zone. L'ensemble de ces zones est appelé zonage sismotectonique.

Les zones sismotectoniques sont définies en combinant des données sismologiques, géologiques et géophysiques [par ex., Martin et al., 2002, Baize et al., 2013, Le Dortz et al., 2019, pour une application à la France métropolitaine]. Les données sismologiques concerne les failles actives ainsi que les données présentées Section 1.2 : sismicités instrumentale, historique et paléosismique. Ces données représentent la sismicité récente, i.e., inférieure à deux millions d'années et sont qualifiées de dynamiques. Les données géologiques et géophysiques regroupent par exemple les profondeurs de bassins sédimentaires, l'âge et la lithologie du sous-sol ou encore la profondeur du Moho. Ces données sont caractérisées par une plus grande échelle temporelle et sont donc qualifiées de statiques. La prise en compte de l'ensemble de ces données permet au zonage sismotectonique de décrire au mieux le processus de production des séismes dans la région d'étude.

L'absence de consensus dans la manière de créer un zonage sismotectonique ainsi que la variété des données utilisées et les hypothèses et choix considérés font qu'il en existe une grande quantité. Par exemple, la France métropolitaine compte les zonages de GEOTER [Martin et al., 2002, Le Dortz et al., 2019], de EDF et de IRSN [Baize et al., 2013] (Figure 1.4). Le zonage GEOTER 2002 (Figure 1.4b) a servi de base pour l'élaboration de la carte d'aléa sismique réglementaire en vigueur en France à ce jour. Dans certains cas (par ex., IRSN), cette variabilité est renforcée par des incertitudes associées à l'existence de certaines frontières ou à leur position. Ainsi, à partir du seul zonage proposé Figure 1.4d, une multitude de zonages sont réalisables suivant la prise en compte de ces incertitudes.

Des études en France métropolitaine ont montré que le choix du zonage sismotectonique a un impact sur l'estimation de l'aléa. Par exemple, Beauval et al. [2020] montrent que les écarts observés entre les résultats d'aléa peuvent atteindre 30% suivant le choix du zonage. De leur côté, Drouet et al. [2020] utilisent trois zonages sismotectoniques différents et constatent, pour un même site, des différences entre les distributions fréquence-magnitude obtenues à partir de ces trois zonages. Ces différences sont visibles Figure 1.5a et peuvent atteindre un ordre de grandeur.

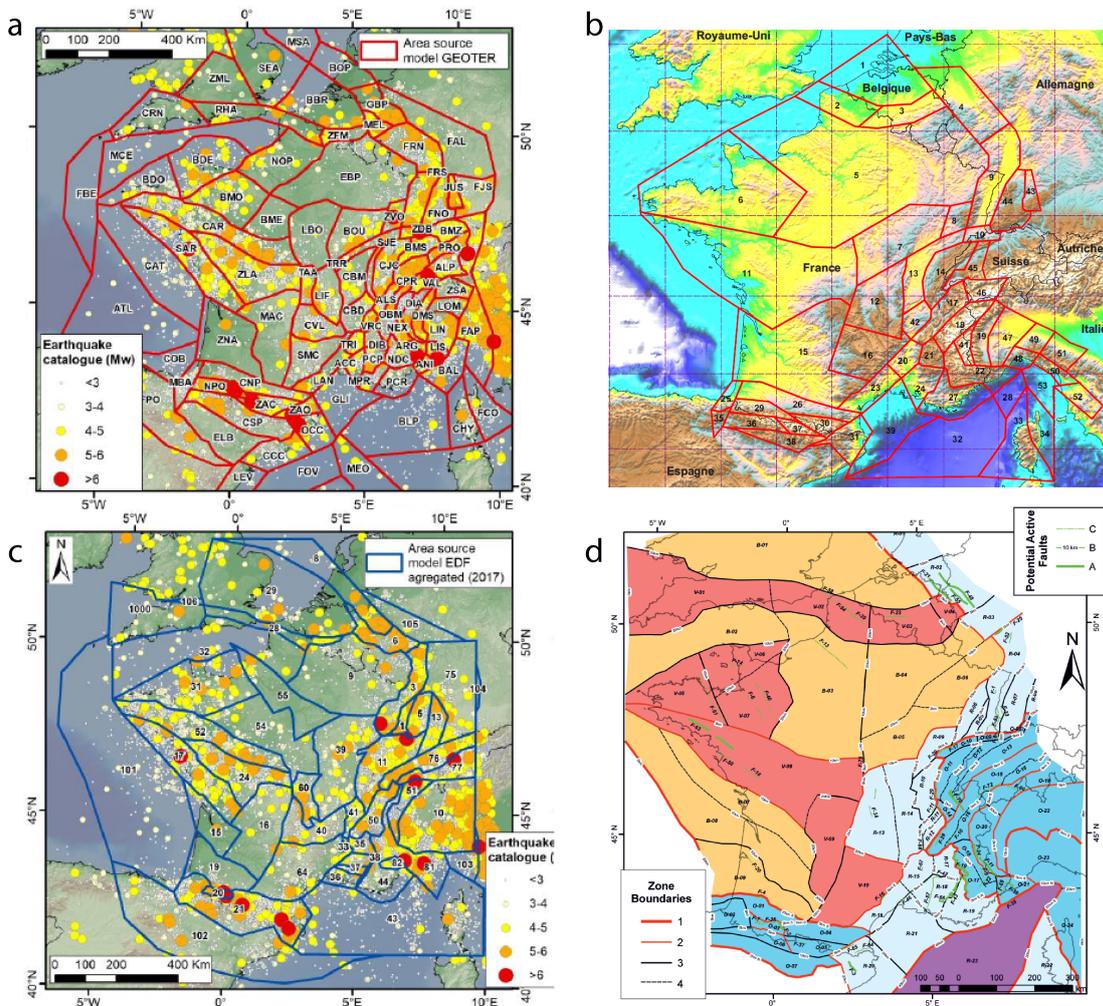


Figure 1.4: EXEMPLES DE ZONAGES SISMOTECTONIQUES FRANÇAIS [a et c d'après Drouet et al., 2020].

a et b : Zonage GEOTER de 2019 [Le Dortz et al., 2019] et 2002 [Martin et al., 2002].

c : Zonage EDF.

d : Zonage IRSN [Baize et al., 2013].

1.4.1.2 La sismicité diffuse

Dans les années 90, un modèle est proposé [par ex., Jacob et al., 1994, Woo, 1996, Frankel et al., 1996] dans le but de s'affranchir de la difficulté et la subjectivité liées à la création du zonage sismotectonique. L'idée est d'analyser la sismicité de manière continue en appliquant un modèle de lissage sur la sismicité observée. Cette approche traduit une variation progressive d'une sismicité à une autre contrairement au zonage catégorique (Figure 1.5b).

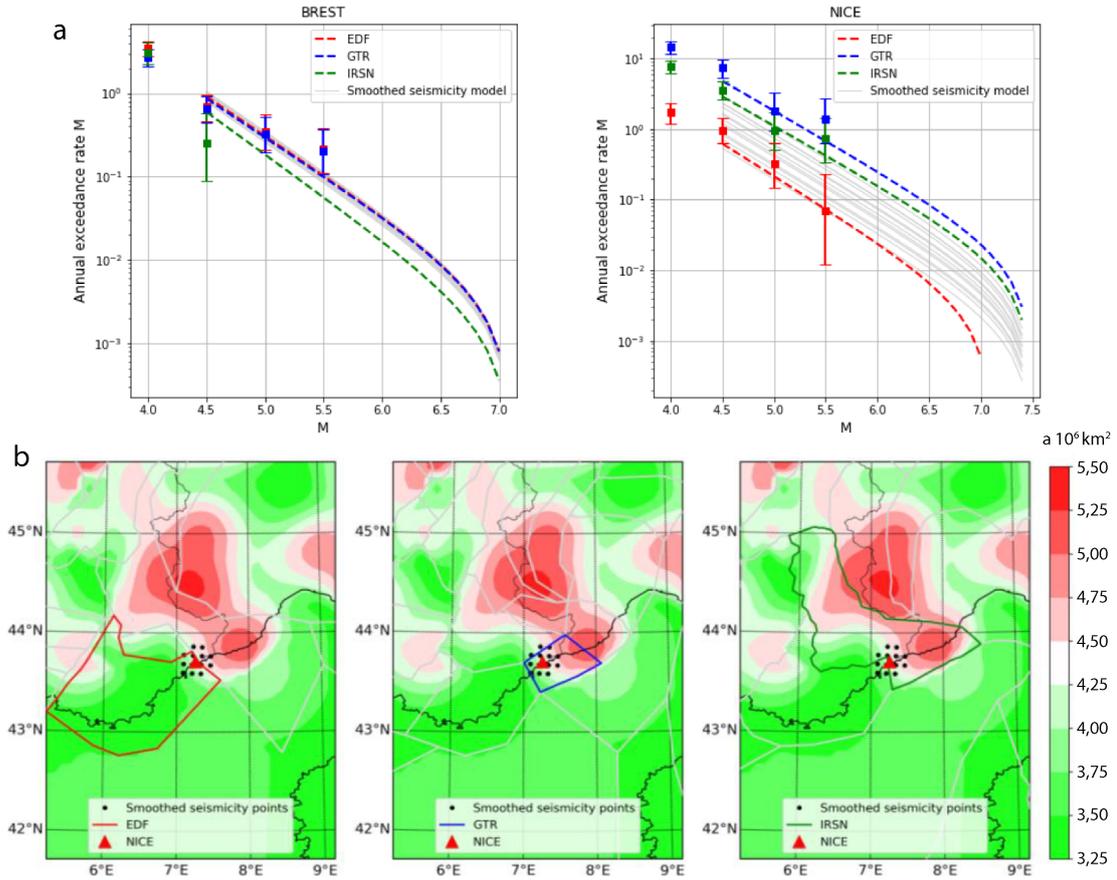


Figure 1.5: COMPARAISON DES MODÈLES SOURCE ZONAGE ET LISSAGE [d'après Drouet et al., 2020].

Les zonages utilisés sont ceux de EDF (rouge, Figure 1.4a), de GEOTER (bleu, Figure 1.4c) et de l'IRSN (vert, Figure 1.4d).

a : Comparaison des distributions fréquence- magnitude dans les villes de Brest (gauche) et Nice (droite) obtenues à l'aide des zonages. Les courbes grises représentent les DFM obtenues à partir des points de calcul de la sismicité diffuse. Ces points sont localisés dans un rayon de 20 km autour de la ville (b).

b : Localisation de la ville de Nice vis-à-vis des trois zonages utilisés et de la sismicité diffuse calculée. Les points noirs sont les points de calcul de la sismicité diffuse.

Pour réaliser le lissage de la sismicité observée, l'aire d'étude est divisée en une grille à maille régulière (par ex., 10×10 km). En chaque maille, une valeur est obtenue en moyennant les densités de probabilité spatiales d'une métrique calculées en chaque point de données. Cette métrique est généralement l'activité sismique (a de la loi GR, Figure 1.5b).

L'application du lissage de la sismicité au sein de la méthode des catalogues synthétiques est très répandue comme par exemple dans les Pyrénées [Beauval et al., 2006], en Corée du Sud [Han et Choi, 2008] ou encore en Iran [Shahbazi et Mansouri, 2019].

1.4.1.3 Les failles actives

Ce modèle source se base uniquement sur les failles actives majeures connues et pour lesquelles de nombreuses études et données sont disponibles. Cela donne plus de poids aux régions où se trouvent ces failles et permet de mieux détailler la sismicité le long de celles-ci. Ces failles actives sont généralement de grandes failles associées aux plus forts séismes passés connus. De ce fait, les failles sont généralement utilisées uniquement pour représenter la forte sismicité (par ex., $M_w \geq 6,5$ pour Woessner et al. [2015] et $M_w \geq 7$ pour Bommer et al. [2002] et Crowley et Bommer [2006]).

Récemment, une méthode des catalogues synthétiques utilisant uniquement ce modèle source a été appliquée dans le Sud de la Californie [Milner et al., 2021]. Les auteurs ont généré plus de 700 000 ans de sismicité le long des nombreuses failles actives connues dans cette région. Cependant, dans les régions stables, le nombre de failles actives bien documentées reste faible malgré les efforts soutenus concernant la découverte et la description des ces dernières [par ex., Feuillet et al., 2001, 2002, Terrier et al., 2002, 2018, Cushing et al., 2008, Guyonnet-Benaize, 2011, Guyonnet-Benaize et al., 2015, Jomard et al., 2017, pour la France métropolitaine].

Le modèle source "Faille" est donc rarement utilisé dans les régions stables. Lorsque c'est le cas, son application ne se fait pas sur l'entièreté du territoire et surtout se fait en complément d'un ou des deux autres modèles source. Ainsi, Martin et al. [2018] ont pris en compte les failles actives dans l'analyse de l'aléa sismique seulement au niveau d'un réseau de dix failles en Provence. Estimer l'aléa sismique dans ce type de région n'est donc pas chose aisée [Stein et al., 2015].

1.4.2 Modèle de récurrence

Le modèle de récurrence définit la distribution fréquence-magnitude observée dans chaque source sismique. De cette distribution sont extraits le nombre annuel ν (Équation 1.2) de séismes à magnitude supérieure ou égale à la magnitude minimale de l'étude et la probabilité $P(M_w)$ (Équation 1.2) d'obtenir un séisme de magnitude M_w dans la source sismique.

1.4.2.1 Loi Gutenberg-Richter

La récurrence des séismes est classiquement définie à l'aide de la loi Gutenberg-Richter (GR, Gutenberg et Richter [1944], Équation 2). Lorsqu'un catalogue sismique de bonne qualité contient assez de séismes, une régression de la distribution du nombre annuel de séismes de magnitude supérieure ou égale à M_c sur ces données permet de calculer les coefficients a et b . Cette distribution fréquence-magnitude (DFM) s'obtient facilement à l'aide de l'Équation 2 tel que :

$$\forall M_w \in [M_c, M_{maj}], F_{GR}(\geq M_w) = \nu e^{-\beta(M_w - M_c)}, \quad (1.3)$$

où $\beta = b \ln(10)$ et ν (Équation 1.2), le nombre annuel de séismes de magnitude supérieure ou égale à M_c , est égal à 10^a .

En analyse de l'aléa sismique, cette DFM est tronquée [Cornell et Vanmarcke, 1969] par la magnitude maximale M_{maj} :

$$\forall M_w \in [M_c, M_{maj}], \quad F_{GR\text{tronc}}(\geq M_w) = \nu \times \frac{e^{-\beta(M_w - M_c)} - e^{-\beta(M_{maj} - M_c)}}{1 - e^{-\beta(M_{maj} - M_c)}}. \quad (1.4)$$

Cette pratique trouve sa justification dans le fait que les plans de ruptures sont limités par l'étendue de la zone sismogène. La dimension du plan de rupture étant directement liée à la magnitude du séisme [par ex., Wells et Coppersmith, 1994, Leonard, 2014], la magnitude des séismes est également bornée par une valeur maximale M_{maj} . Or, la loi GR est une relation affine entre la magnitude et le \log_{10} du nombre de séismes (Équation 2). Il est donc nécessaire de tronquer cette droite afin qu'elle ne se prolonge pas à des magnitudes infinies. En appliquant l'Équation 1.4, la DFM tend donc vers 0 lorsque la magnitude tend vers M_{maj} . La Figure 1.6 illustre un exemple d'application de la loi GR tronquée à la magnitude $M_{maj} = 7$ sur les données sismiques françaises.

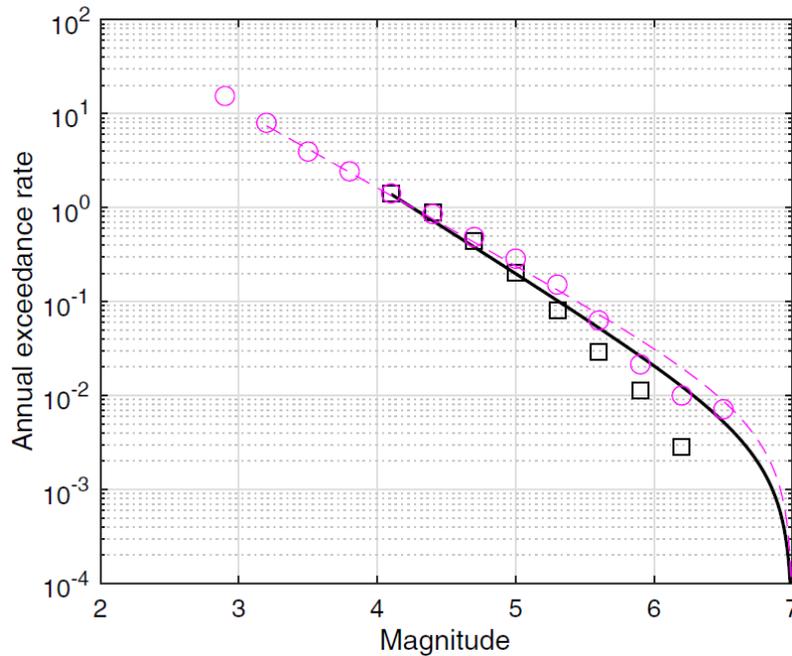


Figure 1.6: LOI GR TRONQUÉE APPLIQUÉE À LA SISMICITÉ FRANÇAISE [Beauval et al., 2020].

Application de la loi GR tronquée (Équation 1.4) sur les données sismiques françaises provenant des catalogues FCAT (rose, Manchuel et al. [2017]) et SHEEC (noir, Stucchi et al. [2013]). Les marqueurs représentent les taux annuels de sismicité observés dans les catalogues tandis que les courbes sont les modèles GR obtenus à partir de ces observations. Les coefficients a et b de la loi GR sont calculés entre les magnitudes 3,2 et 5,6 pour le catalogue FCAT ($b = 0,82$) et entre les magnitudes 4,1 et 5,3 pour le catalogue SHEEC ($b = 0,93$). La magnitude maximale M_{maj} est égale à 7.

1.4.2.2 Limites à l'application de la loi GR

Pour appliquer la loi GR, il est nécessaire que les données observées soient exhaustives. Comme mentionné Section 1.2, l'exhaustivité peut être atteinte en bornant par le bas la gamme de magnitudes via une magnitude de coupure M_c . La régression de la loi GR sur les données se fait donc à partir de M_c . De plus, l'Équation 1.4 stipule que la fréquence des séismes plonge asymptotiquement au niveau de la magnitude maximale M_{maj} (Figure 1.6). Il est donc nécessaire de définir une magnitude de coupure, maximale cette fois-ci, bornant par le haut la gamme de magnitudes utilisée pour la régression de la loi GR. Dans l'exemple Figure 1.6 [Beauval et al., 2020], pour la France métropolitaine, les gammes de magnitudes utilisées pour appliquer cette régression sont 3,2-5,6 pour le catalogue FCAT et 4,1-5,3 pour le catalogue SHEEC. Dans une étude intégrant le monde entier (Figure 1.7 Sachs et al. [2012]), cette gamme de magnitudes s'étend de 5,25 à 7,5. Les coefficients a et b obtenus par régression dans cette gamme de magnitudes sont utilisés Équation 1.4 afin de prolonger la DFM pour les plus fortes magnitudes (Figure 1.6).

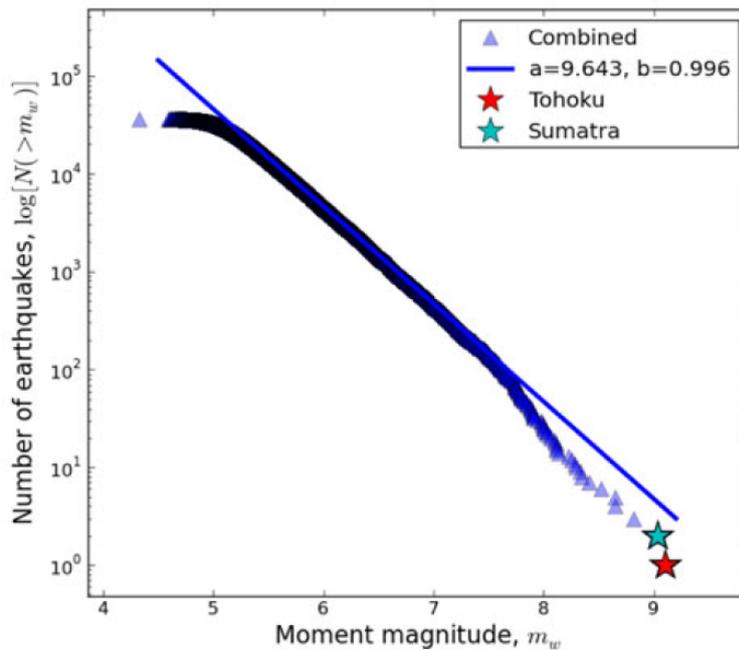


Figure 1.7: LOI GR APPLIQUÉE À L'ÉCHELLE MONDIALE [Sachs et al., 2012].

Loi GR appliquée sur le catalogue *Global Central Moment Tensor* (<https://www.globalcmt.org/>) du 01/01/1977 au 30/06/2010. Une régression de la loi GR entre les magnitudes 5,25 et 7,5 donne un b proche de 1. En revanche, au-delà de M_w 7,5, la distribution du nombre de séismes suit une deuxième loi de puissance avec une pente plus élevée.

Ce prolongement fait sens parce que la loi GR est observée dans une grande partie de la planète et pour une large gamme de magnitudes. Néanmoins, dans certains cas, deux lois de puissance sont observées au lieu de la seule loi GR. La troncation exponentielle (Équation 1.4) ne parvient pas à décrire ce comportement. Dans l'exemple français Figure 1.6 [Beauval et al., 2020], cette observation peut être faite au-delà de la magnitude 5,3 pour le catalogue SHEEC (noir). Pour le même territoire d'étude, le catalogue FCAT

(rose) aboutit, quant à lui, à une seule loi de puissance, malgré des fréquences annuelles instables au-delà de $M_w 5,6$. La méthode d'estimation des magnitudes étant différente d'un catalogue à l'autre, il est normal d'observer ces comportements différents. Il est cependant difficile de définir si ce constat traduit une sous-représentation des fortes magnitudes dans le catalogue SHEEC ou bien une sur-représentation de celles-ci dans le catalogue FCAT. Dans le cas d'une sous-représentation des fortes magnitudes dans le catalogue SHEEC, la seule loi GR est observée. Cependant, si le catalogue FCAT sur-représente les fortes magnitudes, alors deux lois de puissance sont observées. Dans l'exemple mondial Figure 1.7 [Sachs et al., 2012], l'observation de deux lois de puissance ne peut pas être attribuée à une sous-représentation des données de fortes magnitudes étant donné que ces dernières sont bien échantillonnées sur une courte période temporelle (par ex., vingt séismes $M_w \geq 8$ en 33 ans). Cette dualité de la DFM est principalement observée sur la sismicité mondiale [Sornette et al., 1996, Sachs et al., 2012, Corral et González, 2019]. Dans le cas où l'existence de deux lois de puissance est avérée, prolonger la loi GR aux fortes magnitudes a pour conséquence de grandement surestimer la fréquence des séismes à fortes magnitudes. Par exemple, la Figure 1.7 illustre une surestimation du nombre de séismes $M_w \geq 8$: vingt observés contre cinquante estimés par la loi GR. Cela pose la question suivante : est-il juste de prolonger la loi GR pour les fortes magnitudes à l'aide de coefficients estimés à l'aide de magnitudes plus petites ?

Comme mentionné Section 1.2.4, en analyse de l'aléa sismique, l'application de la loi GR se fait uniquement sur les séismes principaux. Néanmoins, la loi GR ne permet pas intrinsèquement de distinguer les séismes principaux des séismes dépendants. Elle est donc classiquement associée à un algorithme de *declustering* qui permet de séparer les séismes principaux des séismes corrélés (Section 1.2.4). La dépendance des séismes est un phénomène complexe que la communauté scientifique étudie encore activement. De ce fait, il n'existe pas de consensus concernant l'élaboration d'un algorithme de *declustering*. Des études [par ex., Luen et Stark, 2012, van Stiphout et al., 2012] ont montré, pour une application en Californie, que les résultats de différents algorithmes sont divers et variés. Il est par exemple reconnu que les algorithmes de Reasenber [1985] et de Grünthal [1985] sont ceux qui enlèvent respectivement le moins et le plus de séismes dépendants dans un catalogue. La Figure 1.8 illustre parfaitement la différence qu'il peut exister entre les catalogues résiduels de plusieurs algorithmes de *declustering*. Dans l'exemple Figure 1.8b, le nombre de séismes contenus dans les catalogues résiduels peut varier du simple au triple. De plus, l'algorithme de *declustering* de Reasenber [1985] a été développé à l'aide de faibles magnitudes M_c de coupure. La faible sismicité est donc le "moteur" de cet algorithme pour assembler un *cluster*. Christophersen et Smith [2008] estiment que cette caractéristique empêche l'algorithme de trouver les dépendances sismiques lorsque M_c est grande puisque les petits séismes ne sont plus observables. Enfin, van Stiphout et al. [2012] ont statistiquement démontré, dans une application aux données californiennes, que les catalogues résiduels obtenus par certains algorithmes de *declustering* [par ex., Reasenber, 1985] ne respectent pas l'hypothèse poissonienne.

Toutes ces spécificités font qu'il est difficile de choisir le ou les algorithmes adaptés à une étude d'aléa donnée. L'impact du choix de l'algorithme de *declustering* sur

l'estimation de l'aléa sismique est analysé dans de nombreuses études. Certaines, appliquées à la Nouvelle-Zélande [Christoffersen et al., 2011], l'Arabie Saoudite [Sokolov et al., 2016], la Turquie [Azak et al., 2017], le Kamchatka et le Caucase [Galina et al., 2019], ont montré que cet impact peut être élevé (jusqu'à 50%) et que ces algorithmes doivent être utilisés avec précaution. D'autres études, appliquées en France métropolitaine [Drouet et al., 2020, Beauval et al., 2020], aboutissent à une conclusion opposée. En effet, ils démontrent que cet impact est faible (variation d'environ 1% de b) lorsque la magnitude de coupure est peu élevée (3-3,5).

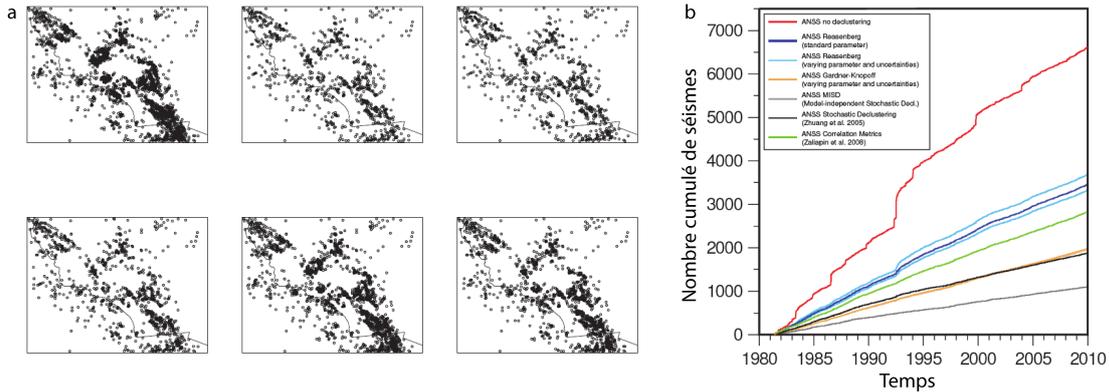


Figure 1.8: COMPARAISON DE DIFFÉRENTS ALGORITHMES DE DECLUSTERING.

a : Impact spatial des algorithmes de *declustering* sur des données californiennes [Luen et Stark, 2012]. Chaque case montre la distribution spatiale des catalogues résiduels résultant du *declustering* du catalogue initial (en haut à gauche) par cinq algorithmes différents.

b : Impact temporel des algorithmes de *declustering* sur des données californiennes [van Stiphout et al., 2012]. Chaque courbe illustre le nombre cumulé de séismes dans les catalogues résiduels résultant du *declustering* du catalogue initial (rouge) par cinq algorithmes différents.

Enfin, certains algorithmes nécessitent une paramétrisation [par ex., Reasenber, 1985, Zhuang et al., 2002]. La sélection des paramètres ne peut rarement se faire à l'aide du peu de données disponibles en région stable, augmentant la part de subjectivité.

1.4.2.3 Méthode des temps de latence

Le couple algorithme de *declustering*-loi GR est classiquement utilisé dans l'analyse de l'aléa sismique. Cependant, il existe des approches alternatives permettant aussi de déterminer la distribution fréquence-magnitude (DFM) des séismes principaux. La méthode des temps de latence en est une. Cette méthode est non paramétrique et est initialement proposée par Corral [2004] et Molchan [2004]. Ces auteurs estiment la proportion de séismes principaux à partir des temps de latence τ_i normalisés (Équation 1.5). Le temps de latence correspond à la durée entre deux séismes successifs. La normalisation se fait via le temps de latence moyen :

$$\forall i \in [1, N - 1], \tau_i = (t_{i+1} - t_i) / (t_{total} / N), \quad (1.5)$$

avec N le nombre de séismes, t_{total} le temps total entre le premier et le dernier séisme et t_{i+1} et t_i les temps d'occurrence des $(i+1)^{ème}$ et $i^{ème}$ séismes du catalogue étudié.

Corral [2004] démontre que la distribution $f(\tau)$ des temps de latence normalisés τ suit une loi gamma :

$$f(\tau) = C \tau^{\gamma-1} e^{-\tau\gamma}. \quad (1.6)$$

où C est une constante et γ le ratio entre la moyenne $\bar{\tau}$ et la variance σ_{τ}^2 des temps de latence normalisés. Les temps de latence étant normalisés, leur moyenne converge vers 1. Ainsi γ est égal à l'inverse de la variance des temps de latence normalisés. D'un côté, l'Équation 1.6 converge vers une loi de Poisson lorsque γ tend vers 1. On l'a vu précédemment, cette loi est classiquement utilisée pour décrire le comportement temporel des séismes principaux. De l'autre côté, la distribution des temps de latence converge vers une loi inverse, analogue à la loi OU (Équation 3), lorsque γ tend vers 0. Or, la loi OU décrit le comportement temporel des séismes dépendants. Ainsi, Molchan [2004] démontre que la proportion p_{sp} de séismes principaux est égale à γ :

$$p_{sp} = \frac{\bar{\tau}}{\sigma_{\tau}^2} \text{ tel que } \begin{cases} p_{sp} \rightarrow 1 & \text{séismes principaux seulement} \\ p_{sp} \rightarrow 0 & \text{séismes dépendants seulement} \end{cases}. \quad (1.7)$$

Hainzl et al. [2006] proposent d'estimer la DFM de séismes principaux pour chaque magnitude M_w à l'aide de la méthode des temps de latence. Pour se faire, seuls les séismes de magnitude supérieure ou égale à M_w sont gardés dans le catalogue initial. À partir de ces séismes conservés, la proportion $p_{sp}(\geq M_w)$ (Équation 1.7) de séismes principaux est calculée pour chaque magnitude. La fréquence annuelle $F_{sp}(\geq M_w)$ de séismes principaux pour chaque magnitude peut donc s'obtenir à partir du nombre $Nb(\geq M_w)$ de séismes dont la magnitude est supérieure ou égale à M_w , du nombre d'années T du catalogue et de $p_{sp}(\geq M_w)$:

$$F_{sp}(\geq M_w) = p_{sp}(\geq M_w) \frac{Nb(\geq M_w)}{T}. \quad (1.8)$$

En calculant $F_{sp}(\geq M_w)$ pour chaque magnitude M_w il est possible d'obtenir la DFM de séismes principaux, autrement dit le même résultat que le couple algorithme de *declustering* - loi GR.

Dans les régions à sismicité faible à modérée où peu de données sont disponibles, l'absence de paramétrisation de cette méthode est un avantage non négligeable. De plus, Beauval et al. [2006] et Hainzl et al. [2006] ont simulé, respectivement dans les Pyrénées et en Californie, une sismicité synthétique à l'aide d'un modèle de processus ponctuel marqué. La proportion de séismes principaux de la sismicité synthétique est donnée en entrée de leur modèle et est donc connue. Ces auteurs comparent les proportions de séismes principaux calculées sur la sismicité simulée à l'aide de la méthode des temps de latence et de l'algorithme de *declustering* de Reasenberg [1985]. Dans les deux applications, ils ont montré que la méthode des temps de latence aboutit à la proportion de séismes principaux la plus proche de celle donnée en entrée du modèle.

En revanche, cette méthode ne distingue pas les répliques des séismes principaux; seule une proportion de séismes principaux est calculée. Elle ne permet donc pas de

définir quels séismes sont dépendants et lesquels sont indépendants. De plus, la DFM est calculée pour chaque pas de magnitude ainsi les fréquences des magnitudes les plus élevées ne sont pas extrapolées à partir d'observations faites aux magnitudes plus faibles. Si le catalogue utilisé pour calculer ces DFM est considéré comme complet en ce qui concerne les plus fortes magnitudes alors cet aspect de la méthode des temps de latence est un avantage. Au contraire, si peu de confiance est attribuée au catalogue en ce qui concerne cette exhaustivité, la méthode des temps de latence n'est pas applicable.

1.4.3 Génération des catalogues synthétiques

La génération des catalogues synthétiques est réalisée lorsque le modèle source et le modèle de récurrence sont choisis et appliqués. Initialement [Musson, 1998, 1999], cette approche est appliquée avec un modèle source de type "Zone". Dans ce cas, un catalogue fictif est construit comme suit :

- Le nombre de séismes principaux par an et par zone est tiré suivant une loi de Poisson de paramètre égal à l'activité sismique ν . Cette dernière est obtenue à l'aide du modèle de récurrence (Équation 1.3).
- Les épicentres des séismes produits sont placés de manière uniforme au sein de la zone.
- Leurs paramètres tels que la profondeur ou les informations sur le plan de rupture (azimut, pendage, régime) sont tirés aléatoirement dans des intervalles de valeurs définis au préalable.
- La magnitude des séismes simulés est déterminée suivant la distribution fréquence-magnitude des séismes principaux calculée dans la zone à l'aide du modèle de récurrence.

Classiquement, les catalogues synthétiques ne produisent que des séismes principaux. Néanmoins, les répliques peuvent avoir de l'importance sur l'aléa sismique et encore plus sur le risque sismique. En effet, Faenza et al. [2007] ont montré, pour une application dans les Pyrénées, que ne pas prendre en compte les répliques entraîne une sous-estimation de l'aléa pouvant aller jusqu'à 10%. De plus, cette sous-estimation est encore plus accentuée lorsque l'on s'intéresse au risque sismique puisque les répliques impactent des structures déjà fragilisées par le séisme principal. Ces répliques peuvent donc être encore plus désastreuses car elles peuvent détruire un bâtiment qui était seulement fragilisé par le séisme principal. La prise en compte des répliques est donc nécessaire notamment dans le contexte de cette thèse.

Le modèle de génération de sismicité complète (séismes principaux et répliques) le plus répandu est le modèle ETAS (*Epidemic Type Aftershocks-Sequence*). Ce modèle a été introduit par Kagan et Knopoff [1981] puis démocratisé par Ogata [1988]. Tout comme la procédure présentée ci-dessus, le but est de créer un catalogue synthétique de sismicité mais en produisant également les séismes dépendants à l'aide de différents modèles. L'idée est de produire des séquences sismiques en cascade. Chaque séisme principal simulé par une loi de Poisson de paramètre ν peut produire un ensemble

de répliques. Chacune de ces répliques peut également produire sa propre séquence sismique et ainsi de suite. Pour chaque génération de répliques, la magnitude est tirée dans la même distribution fréquence-magnitude que celle des séismes principaux. Pour chaque séisme (principal ou répliques), un nombre $n(t)$ de répliques sont produites dans le temps à l'aide de la loi Omori-Utsu [Omori, 1894, Utsu, 1957] :

$$n(t) = K(t + c)^{-p}, \quad (1.9)$$

où K , c et p sont des constantes. c représente le délai de détection des répliques après un fort séisme. Il permet à l'Équation 1.9 de converger lorsque $t = 0$. p est le coefficient de la décroissance temporelle du nombre de répliques : plus p est élevé, plus la production de répliques au cours du temps est faible. Sa gamme de valeurs s'étend de 0,6 à 2,5 avec une médiane égale à 1,1 [Utsu et al., 1995]. K est la constante de productivité exprimée par Helmstetter et Sornette [2003b] [d'après Utsu, 1970] comme suit :

$$K = k \times 10^{\alpha(M_{sp} - M_c)}, \quad (1.10)$$

avec M_c la magnitude de coupure, M_{sp} la magnitude du séisme principal et k et α des constantes.

Enfin, la localisation des répliques se fait généralement de manière isotropique autour du séisme principal bien que certains modèles permettent une distribution spatiale anisotropique liée à la géométrie du plan de rupture [par ex., Hainzl et al., 2010].

Le principal problème à l'application du modèle ETAS est la nécessité de quantifier au moins cinq paramètres : ν (Équation 1.3), c et p (Équation 1.9) et k et α (Équation 1.10). Afin de définir ces paramètres de manière objective, le mieux est de les calibrer sur des séquences sismiques passées et bien enregistrées. Ces séquences sont malheureusement peu nombreuses en région stable. Ainsi, pour une application dans les Pyrénées, Beauval et al. [2006] sont contraints de définir deux de ces paramètres à l'aide des données en fixant les trois autres. La définition de ces trois paramètres s'appuie sur des travaux réalisés dans des régions plus actives : Californie pour α [Helmstetter, 2003], Japon pour p [Ogata, 1998] et Italie pour c [Console et Murru, 2001]. Ces régions ne partagent pas forcément les mêmes propriétés que les régions stables ainsi, la pertinence d'utilisation de ces trois paramètres n'est pas totalement vérifiable.

1.4.4 Modèle de prédiction des mouvements du sol

La dernière étape du PSHA avant de produire les courbes probabilité-aléa consiste à calculer cet aléa. À ce stade, seuls des séismes définis par un ensemble de paramètres tels que leur localisation, leur magnitude et leur temps d'occurrence sont disponibles. Il existe des formulations mathématiques qui traduisent cet ensemble de paramètres en une valeur d'aléa. Ces formulations sont appelées modèles de prédiction des mouvements du sol ou encore GMPE pour *Ground Motion Prediction Equations*. Ils permettent d'estimer la probabilité $P(a > a_0 | M_w, D)$ (Équation 1.2) qu'un séisme de magnitude M_w situé à une distance D du site d'étude produise un aléa supérieur au seuil a_0 , en échantillonnant l'incertitude aléatoire σ .

Les GMPE sont des équations empiriques qui visent à reproduire l'aléa sismique Y (PGA ou PGV) sans chercher à simuler avec exactitude les phénomènes physiques. Elles requièrent, *a minima*, la magnitude M_w du séisme et la distance D séisme - site. D'autres paramètres tels que le type de sol sur site, le mécanisme au foyer ou la profondeur du séisme sont pris en compte dans certaines GMPE. Un exemple de GMPE est donné par Drouet et Cotton [2015] comme suit :

$$\ln(Y) = b_1 + b_2(M_w - 8) + b_3(M_w - 8)^2 + (b_4 + b_5 M_w) \ln\left(\sqrt{D^2 + b_6^2}\right) + b_7 D + \epsilon\sigma. \quad (1.11)$$

Les coefficients b_i sont obtenus par régression de l'Équation 1.11 sur des données réelles d'accélération ou de vitesse comme l'illustre la Figure 1.9. Cette Figure montre également la grande dispersion qu'il existe autour des données de PGA ou PGV. Ainsi, $\ln(Y)$ ne donne qu'une représentation médiane du véritable aléa tandis que σ correspond à l'écart-type de cette valeur médiane. ϵ est la fraction de σ que l'on ajoute ou soustrait à $\ln(Y)$. Cette fraction suit une distribution normale centrée réduite. Autrement dit, $\epsilon\sigma$ représente l'erreur aléatoire à distribution normale de moyenne zéro et d'écart-type σ . La probabilité $P(a > a_0 | M_w, D, \sigma)$ (Équation 1.2) qu'un séisme donné (M_w, D) produise au site un aléa supérieur au seuil a_0 est estimée à l'aide de $\epsilon\sigma$.

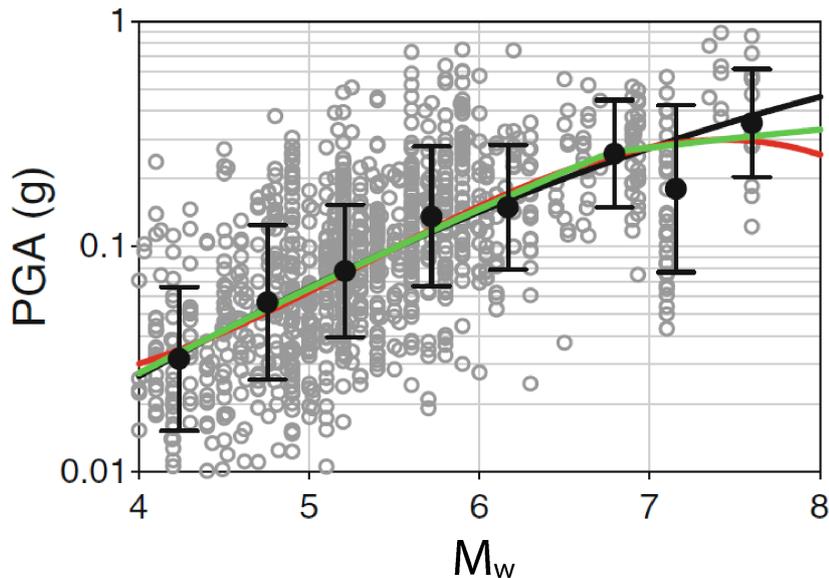


Figure 1.9: DISTRIBUTION PGA - MAGNITUDE ET MODÈLES ASSOCIÉS [Akkar et al., 2014a].

Application de différentes formes fonctionnelles (noir : quadratique continue, rouge : cubique, vert : quadratique discontinue) sur des PGA enregistrées (cercles gris). Ces données sont enregistrées au rocher à une distance de 10 km et sont produites par des séismes à régime décrochant.

Peu de modèles de prédiction des mouvements du sol sont formulés à l'aide de données de régions à sismicité faible à modérée. En effet, comme ces régions possèdent peu de données d'accélération ou de vitesse, la plupart des GMPE sont établies à partir de données de régions sismiquement actives. L'utilisation de GMPE formulées dans

ces territoires à fortes sismicité est fréquente pour une application aux régions stables. Le choix de ces GMPE externes est basé sur deux hypothèses. La première est qu'il n'y a pas de variabilité spatiale des mouvements du sol, à distance et magnitude fixes, entre toutes les régions crustales du monde. La deuxième est qu'il n'y a pas d'effet d'échelle de magnitude, i.e., que la forme fonctionnelle des GMPE est indépendante de la magnitude. Ces deux hypothèses sont débattues dans la littérature [voir Beauval et al., 2012, pour une liste non exhaustive]. L'effet d'échelle de magnitude a par exemple été mis en évidence [par ex., Cotton et al., 2008]. Cette observation a nécessité une modification des formes fonctionnelles des modèles de prédiction au milieu des années 2000. Cette modification correspond au passage d'une relation linéaire (courbe noire Figure 1.9) à des relations plus complexes (courbes rouge et verte Figure 1.9) entre l'aléa et la magnitude [par ex., Zhao et al., 2006, Bindi et al., 2009]. Ce point est important car dans les régions actives, l'estimation de l'aléa est fréquemment étudié à partir de $M_w \geq 5$. Certains modèles de prédiction des mouvements du sol établis dans ces régions actives ne sont applicables qu'à partir de cette magnitude. Néanmoins, dans les régions stables, l'estimation de l'aléa sismique est réalisé à partir de magnitudes plus faibles, qui jouent un rôle sur l'aléa final [par ex., Beauval et al., 2008, Nievas et al., 2020]. De plus, couplées à d'autres paramètres tels que la distance (par ex., Le Teil, 2019) ou le type de sol (par ex., Mexico, 1985), ces magnitudes peuvent accroître le risque sismique de manière significative. Il est donc nécessaire, surtout dans les régions stables, de ne pas prendre en compte que les magnitudes modérées à fortes (classiquement $M_w \geq 5$). Les formes fonctionnelles des GMPE doivent donc être adaptées en conséquence. Ce constat tend également à penser que la pondération d'un ensemble de GMPE ne peut *a priori* pas se faire indépendamment de la magnitude.

La sélection du modèle de prédiction des mouvements du sol est importante dans l'analyse de l'aléa sismique. Certains auteurs [par ex., Cotton et al., 2006] donnent des conseils pour choisir parmi les nombreuses GMPE existantes (Douglas [2019] en recense plus de 400) celles qui sont les plus adaptées au territoire d'étude et aux objectifs de l'étude envisagée. Ces GMPE doivent être définies par des intervalles de validité de magnitude et de distance cohérents avec les gammes de magnitudes et distances utilisées dans l'étude de l'aléa sismique. Enfin, ces GMPE doivent être formulées, si possible, avec des données de la région d'étude ou au moins de régions à contexte sismique similaire.

1.5 Conclusion

L'aléa sismique produit en un site donné peut être estimé en modélisant tout d'abord la sismicité (localisation, régime, magnitude) puis en traduisant la propagation et l'atténuation des ondes sismiques du séisme jusqu'au site. Dans cette optique, un scénario sismique est considéré comme la combinaison de la survenance d'un séisme et de l'aléa qu'il produit en surface. L'estimation de l'aléa peut être réalisée de façon déterministe en considérant un seul des scénarios sismiques possibles, qualifié de dimensionnant. L'approche probabiliste permet, quant à elle, d'échantillonner la variabilité des scénarios sismiques possibles afin de calculer une probabilité annuelle de dépasser un seuil d'aléa au site. Au sein de la méthode probabiliste, l'option des catalogues synthétiques se base non pas sur l'intégration de l'ensemble des scénarios sismiques mais sur la création de scénarios synthétiques au sein de catalogues afin de calculer cette probabilité. Cela permet d'être plus pédagogique, compréhensible et surtout d'illustrer la contribution de chaque séisme sur l'aléa produit au site. Pour cet aspect et la prise en compte des incertitudes, la méthode probabiliste des catalogues synthétiques est adaptée à une étude du risque sismique. C'est cette dernière qui est utilisée dans cette thèse dans le but d'établir une analyse de l'aléa qui s'insère dans un modèle de risque sismique.

La méthode des catalogues synthétiques nécessite un modèle source et un modèle de récurrence afin de créer un catalogue synthétique de très longue durée ou un ensemble de sous-catalogues de durée plus courte. Par son absence de paramètres et le fait qu'elle calcule la fréquence annuelle pour chaque pas de magnitude sans extrapolation, la méthode des temps de latence paraît être une alternative pertinente à l'algorithme de *declustering* et à la loi Gutenberg-Richter classiquement utilisés pour le modèle de récurrence. Malgré ses avantages, la méthode des temps de latence n'a, *a priori*, jamais été utilisée comme modèle de récurrence dans une étude probabiliste de l'aléa sismique. Cette thèse est l'occasion de l'utiliser comme modèle de récurrence au sein d'une méthode des catalogues synthétiques. Le Chapitre 2 y est consacré.

En ce qui concerne le modèle source, trois approches ont été présentées Section 1.4.1 : zone homogène, sismicité diffuse (ou lissage) et failles actives. Dans le Chapitre 3, je propose une approche hybride permettant de prendre en compte la variabilité spatiale de la sismicité en contexte intraplaque (modèle source) à l'aide d'un couplage entre les approches par zone et par lissage. Une méthode de production de répliques, moins paramétrique que le modèle ETAS, est également présentée dans ce Chapitre.

Enfin, un ensemble de modèles de prédiction des mouvements du sol pondérés est ensuite utilisé pour estimer l'aléa sismique produit par chaque séisme synthétique. Malgré l'existence de plusieurs analyses comparatives, les poids des GMPE sont souvent déterminés à dire d'experts. Dans cette thèse, une pondération réalisée à l'aide de données d'accélération et en fonction de la magnitude du séisme et de la distance séisme-site est proposée. Le Chapitre 4 détaille comment cette pondération est effectuée et comment l'aléa est calculé.

L'ensemble des contributions apportées dans cette thèse est appliqué à la France métropolitaine Chapitre 5, en lien avec l'estimation du risque sismique qui a motivé ces travaux de recherche.

Chapitre 2

Du catalogue sismique à la distribution fréquence - magnitude

Sommaire

2.1	Introduction	40
2.2	Distribution fréquence- magnitude de séismes principaux	41
2.2.1	Problématique	41
2.2.2	Implémentation MIMP (<i>Magnitude of Inferred Main shock Proportion</i>)	42
2.3	Echantillonnage de l'incertitude autour de la distribution fréquence - magnitude	45
2.3.1	Propagation des incertitudes sur la magnitude	45
2.3.2	Magnitude maximale observée et magnitude majorée	45
2.4	Discussions et Conclusion	46

Résumé Un modèle de récurrence basé sur la méthode des temps de latence est détaillé dans ce Chapitre. Ce modèle est divisé en deux parties : le calcul de la distribution proportion - magnitude (DPM) de séismes principaux d'une part et l'estimation de la distribution fréquence - magnitude (DFM) de séismes principaux d'autre part. La méthode proposée ici vise à adapter la méthode des temps de latence aux régions à sismicité faible à modéré. L'idée est de stabiliser les proportions de séismes principaux pour les fortes magnitudes qui sont en général peu représentées. L'implémentation MIMP, pour *Magnitude of Inferred Main shock Proportion*, permet de corriger la DPM et d'estimer une DFM non biaisée. Enfin, un processus de Monte Carlo est appliqué afin de propager les incertitudes sur la magnitude et d'obtenir des DPM et DFM stochastiques.

2.1 Introduction

L'analyse temporelle de la sismicité passe par la définition de la distribution fréquence - magnitude (DFM) de séismes indépendants ; appelés principaux dans cette thèse. Cette DFM permet d'estimer la probabilité d'occurrence d'un séisme principal en fonction de sa magnitude. Que ce soit pour la méthode PSHA intégrative (Cornell [1968] Équation 1.2) ou bien pour la méthode PSHA des catalogues synthétiques [par ex., Musson, 1999], cette DFM constitue une des données d'entrée permettant de calculer la courbe probabilité annuelle - aléa.

Dans cette thèse, cette DFM est calculée en deux étapes à l'aide de la méthode des temps de latence. Tout d'abord, la proportion p_{sp} de séismes principaux est calculée en fonction de la magnitude à l'aide de la variance σ_τ^2 des temps de latence entre deux séismes successifs [Molchan, 2004] :

$$p_{sp} = \frac{1}{\sigma_\tau^2} \begin{cases} p_{sp} \rightarrow 1 & \text{séismes principaux seulement} \\ p_{sp} \rightarrow 0 & \text{séismes dépendants seulement} \end{cases} . \quad (2.1)$$

Ensuite, la fréquence annuelle des séismes principaux est obtenue pour chaque pas de magnitude en multipliant le nombre annuel de séismes observés dans le catalogue par la proportion de séismes principaux calculée précédemment [Hainzl et al., 2006] :

$$F_{sp}(\geq M_w) = p_{sp}(\geq M_w) \times \frac{Nb(\geq M_w)}{T} , \quad (2.2)$$

avec $Nb(\geq M_w)$ le nombre de séismes à magnitude supérieure ou égale à M_w et T le nombre d'années du catalogue étudié.

La Section 2.2 présente l'implémentation MIMP visant à pallier le manque de fiabilité de la méthode des temps de latence lorsque peu de données sont disponibles. Un processus de Monte Carlo est appliqué Section 2.3 afin de propager l'incertitude sur les magnitudes dans le calcul de la DFM. Un ensemble stochastique de DFM est donc obtenu. Cet ensemble représente la variabilité des magnitudes du catalogue étudié.

2.2 Distribution fréquence - magnitude de séismes principaux

2.2.1 Problématique

D'après la loi de Bâth [Richter, 1958, Bâth, 1965], la magnitude du séisme principal est supérieure à celles des répliques qu'il provoque. En partant de ce postulat, la population de séismes ayant les magnitudes les plus fortes est susceptible d'abriter une proportion de séismes principaux plus élevée que celle de séismes ayant des magnitudes plus faibles. De ce fait, plus la magnitude augmente, plus la probabilité qu'un séisme soit un séisme principal augmente. En appliquant cette hypothèse, une représentation graphique théorique de la distribution proportion - magnitude de séismes principaux attendue est proposée Figure 2.1.

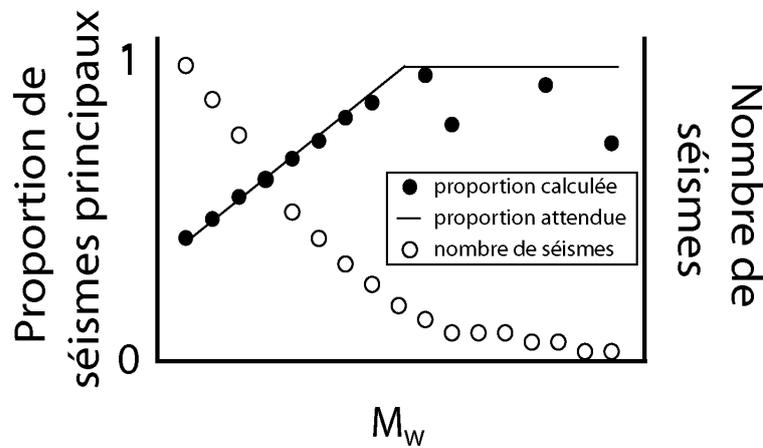


Figure 2.1: DISTRIBUTIONS DE PROPORTIONS DE SÉISMES PRINCIPAUX CALCULÉE ET ATTENDUE.

Distributions synthétiques de proportions de séismes principaux attendue (courbe) et calculée via la méthode des temps de latence (cercles pleins Équation 2.1).

Le nombre de données (cercles vides) disponibles pour calculer de la proportion de séismes principaux diminue lorsque la magnitude augmente.

Un exemple de distribution proportion - magnitude de séismes principaux calculée à l'aide de la méthode des temps de latence est illustré Figure 2.1. Alors que la loi de Bâth [Richter, 1958, Bâth, 1965] suggère que la proportion de séismes principaux tende vers 1 à partir d'une certaine magnitude (courbe), certaines valeurs calculées (cercles pleins) sont anormalement faibles (<1) voire impossibles (>1). Les valeurs supérieures à 1 ne sont pas représentées Figure 2.1. Ainsi, à partir d'une certaine magnitude, l'évaluation des proportions de séismes principaux est fortement impactée par la faible quantité de séismes contenue dans le catalogue (cercles vides). En effet, la méthode des temps de latence se base sur le temps entre deux séismes successifs et nécessite donc au moins la présence de deux séismes. Le calcul de la variance de ces temps de latence implique au moins deux valeurs de temps de latence et donc au moins trois séismes. Il va sans dire que le calcul d'une variance sur deux mesures n'a que peu de sens et peut aboutir à des résultats incohérents avec la loi de Bâth.

Ainsi, les proportions associées aux fortes magnitudes sont sensées être élevées, i.e., proches de 1 (courbe Figure 2.1), mais sont en réalité calculées difficilement et de manière incertaine à l'aide de la méthode des temps de latence (cercles pleins Figure 2.1).

2.2.2 Implémentation MIMP (*Magnitude of Inferred Main shock Proportion*)

L'idée de cette Section est de détecter la magnitude à partir de laquelle les données ne permettent pas à la méthode des temps de latence d'estimer une proportion de séismes principaux en accord avec la loi de Båth (courbe Figure 2.1). On identifie alors une magnitude à partir de laquelle toutes les proportions associées doivent être égales à 1. Cette magnitude s'appelle *Magnitude of Inferred Main shock Proportion* (MIMP). L'implémentation proposée dans cette Section se base sur deux hypothèses. La première est que le catalogue est supposé complet pour toutes les magnitudes qu'il contient. La deuxième est que nous admettons qu'à partir d'une certaine magnitude tous les séismes sont des séismes principaux. Les instabilités observées Figure 2.1 pour les fortes magnitudes sont attribuées au peu de données disponibles.

L'implémentation MIMP s'articule autour de trois axes composés de cinq étapes représentées Figure 2.2.

Calcul des proportions de séismes principaux Tout d'abord, une proportion de séismes principaux est calculée via la méthode des temps de latence pour chaque magnitude M_w comprise entre la magnitude de coupure M_c et la magnitude maximale M_{max} du catalogue. Pour les magnitudes où ce calcul n'est pas possible à cause de la faible quantité de séismes enregistrés, et sous l'hypothèse que le catalogue est complet et représentatif, nous considérons que tous ces séismes sont des séismes principaux. Dans ce cas, la proportion de séismes principaux est fixée à 1.

La Figure 2.2a et l'Algorithme 1 représentent cette étape.

Algorithme 1 : CALCUL DES PROPORTIONS DE SÉISMES PRINCIPAUX

Données : Catalogue

Résultat : Proportions de séismes principaux

```
1 pour chaque magnitude  $M_w$  allant de  $M_c$  à  $M_{max}$  faire
2   | si le nombre de séismes  $\geq M_w$  est supérieur ou égal à 3 alors
3   |   | Calculer la proportion de séismes principaux à magnitude supérieure ou
4   |   | égale à  $M_w$  (Équation 2.1) ;
5   | sinon
6   |   | La proportion de séismes principaux à magnitude supérieure ou égale à
7   |   |  $M_w$  est fixée à 1 ;
8   | fin
9 fin
```

Recherche de MIMP Dans un second temps, si cela n'a pas été fait lors de l'étape précédente, la proportion de séismes principaux associées à la plus forte magnitude M_{max} de l'étude est fixée à 1 (Figure 2.2b). Cette étape découle directement de la loi de Bâth [Richter, 1958, Bâth, 1965] : comme la magnitude d'un séisme principal est plus élevée que celles de ses répliques, nous considérons qu'aucune réplique ne peut être de magnitude M_{max} .

Ensuite, nous introduisons un seuil de proportion appelé seuil MIMP (Figure 2.2c). Cette valeur seuil correspond à la proportion de séismes principaux à partir de laquelle nous attribuons moins de confiance aux résultats de la méthode des temps de latence. L'idée est qu'à partir du seuil MIMP ces résultats doivent se stabiliser entre cette valeur seuil et la valeur maximale 1. L'instabilité de la méthode des temps de latence est attendue aux fortes magnitudes, caractérisées par une faible quantité de données. Ainsi, le seuil MIMP doit être élevé (par ex., 0,9).

La recherche de la magnitude MIMP passe en premier lieu par l'identification de la magnitude seuil M_{seuil} (Figure 2.2d). Cette dernière correspond à la plus petite magnitude associée à une proportion de séismes principaux atteignant ou dépassant le seuil MIMP. À partir de cette magnitude, le catalogue est sensé contenir pratiquement que des séismes principaux. Si M_{seuil} est différente de M_{max} alors MIMP correspond à la plus petite magnitude se trouvant entre M_{seuil} et M_{max} à être associée à une proportion de séismes principaux inférieure au seuil MIMP. Dans le cas où la distribution initialement calculée est stable (par ex., courbe Figure 2.1), aucune MIMP n'est trouvée.

L'Algorithme 2 décrit ces trois étapes.

Algorithme 2 : RECHERCHE DE MIMP

Données : Proportions de séismes principaux (Algorithme 1),
Seuil MIMP

Résultat : MIMP

- 1 La proportion de séismes principaux associée à M_{max} est fixée à 1 ;
 - 2 Recherche de M_{seuil} , la plus petite magnitude à être associée à une proportion de séismes principaux supérieure ou égale au seuil MIMP ;
 - 3 **si** M_{seuil} est inférieure à M_{max} **alors**
 - 4 Recherche de MIMP, la plus petite magnitude supérieure à M_{seuil} ,
inférieure à M_{max} et dont la proportion de séismes principaux associées est inférieure au seuil MIMP ;
 - 5 **fin**
-

Correction des proportions de séismes principaux Enfin, dans le cas où MIMP a été trouvée, toutes les proportions associées à une magnitude se trouvant dans l'intervalle $[MIMP, M_{max}]$ sont fixées à 1 (Figure 2.2e et Algorithme 3). Si la première distribution est déjà stable, aucune MIMP n'est trouvée et donc aucune correction n'est apportée. La distribution finale des proportions de séismes principaux en fonction de la magnitude correspond à celle calculée par la méthode des temps de latence dans la première étape qui a ensuite été corrigée dans cette étape.

Algorithme 3 : CORRECTION DES PROPORTIONS DE SÉISMES PRINCIPAUX

Données : Proportions de séismes principaux (Algorithme 1),
MIMP (Algorithme 2)

Résultat : Distribution de proportions de séismes principaux

- 1 si *MIMP existe* alors
 - 2 | Les proportions de séismes principaux associées aux magnitudes comprises
| entre MIMP et M_{max} sont fixées à 1 ;
 - 3 fin
-

La distribution proportion-magnitude de séismes principaux ajustée à l'aide de l'implémentation MIMP (Figure 2.2f) se rapproche de la distribution théorique attendue à l'image de la courbe visible Figure 2.1.

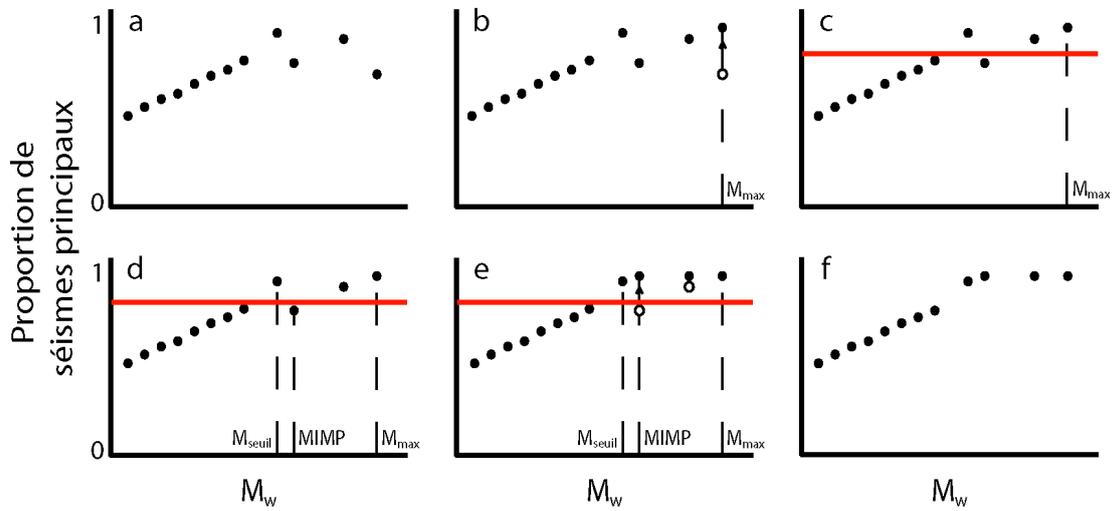


Figure 2.2: REPRÉSENTATION GRAPHIQUE DE L'IMPLÉMENTATION MIMP.

a : Les proportions de séismes principaux sont calculées à partir de la méthode des temps de latence via l'Équation 2.1 ou bien fixées à 1 suivant le nombre de données disponibles. Dans cet exemple, ce dernier cas n'est pas observé ; aucune proportion n'est arbitrairement fixée à 1.

b : Fixation à 1 de la proportion associée à la magnitude maximale M_{max} .

c : Introduction du seuil MIMP (ligne rouge).

d : Recherche de M_{seuil} et MIMP.

e : Fixation à 1 de toutes les proportions dont la magnitude associée est supérieure ou égale à MIMP.

f : Obtention de la distribution proportion - magnitude de séismes principaux corrigée.

Une fois les proportions de séismes principaux corrigées pour chaque pas de magnitude, la distribution fréquence - magnitude est obtenue à l'aide de l'Équation 2.2.

2.3 Echantillonnage de l'incertitude autour de la distribution fréquence - magnitude

2.3.1 Propagation des incertitudes sur la magnitude

Parmi toutes les incertitudes associées aux séismes, les incertitudes liées à leur magnitude ($ErrM_w$) sont les plus critiques dans l'estimation de l'aléa sismique. En effet, la magnitude est à la base des méthodes d'estimation de l'aléa et une augmentation de 1 de la magnitude de moment équivaut à une énergie environ 32 fois plus élevée [Hanks et Kanamori, 1979]. Or, dans certains catalogues (historiques et paléosismiques) ces incertitudes peuvent dépasser 1 [par ex., Manchuel et al., 2017, Traversa et al., 2017, pour le catalogue historique français]. Dans cette thèse, nous nous limitons à l'étude de la seule incertitude $ErrM_w$ sur la magnitude.

Pour prendre en compte cette incertitude ($ErrM_w$), nous proposons de produire un ensemble de N_{cb} catalogues obtenus en perturbant la magnitude du catalogue initial. Ces catalogues sont appelés catalogues bruités et sont représentatifs de la sismicité de la région. L'application conjointe de la méthode des temps de latence sur l'ensemble de ces catalogues propose une vision complète (optimiste et pessimiste) de la DPM et de la DFM des séismes principaux. Les fréquences obtenues sont donc probabilistes : pour chaque pas de magnitude, une distribution de fréquences est obtenue. Ainsi, chaque pas de magnitude est associé à une densité de probabilité de fréquences de séismes principaux. Ces densités de probabilité sont estimées à l'aide des N_{cb} résultats équiprobables obtenus. Étant donné que la méthode des temps de latence est sensible au nombre de données disponibles, ces densités de probabilité sont calculées en pondérant les N_{cb} résultats à l'aide du nombre de données utilisées pour calculer ces résultats. Cela permet de donner plus de poids et donc de confiance aux résultats obtenus avec le plus de données.

Pour produire ces N_{cb} catalogues bruités, nous échantillons l'incertitude $ErrM_w$ à l'aide d'un processus de Monte Carlo. La magnitude de chaque séisme est tirée suivant une loi et en fonction de son incertitude. Trouver la loi qui gouverne la distribution de l'incertitude sur la magnitude n'est pas une tâche aisée. Pour autant, la loi normale est la norme pour tirer des valeurs de variable aléatoire indépendante. C'est cette loi qui est couramment utilisée pour propager l'incertitude sur la magnitude des séismes. Nous proposons donc de tirer la magnitude dans une loi normale de moyenne M_w et d'écart-type $ErrM_w/2$ au sein d'un processus de Monte Carlo. La magnitude tirée est finalement arrondie à 0,1 près.

2.3.2 Magnitude maximale observée et magnitude majorée

Dans cette thèse, la magnitude maximale du territoire étudié correspond à la magnitude majorée définie comme suit :

$$M_{maj} = \max(M_w + ErrM_w), \quad (2.3)$$

avec M_w les magnitudes du catalogue sismique et $ErrM_w$ leur incertitude. Elle est donc constante et associée à aucune incertitude. En appliquant l'Équation 2.3 sur le

catalogue historique français FCAT [Manchuel et al., 2017], la magnitude majorée M_{maj} du territoire français vaut $6,7 + 0,6 = 7,3$.

C'est cette magnitude majorée qui est utilisée comme magnitude maximale de l'étude. En revanche, c'est bien la magnitude maximale observée dans chaque catalogue bruité qui est utilisée dans le calcul des temps de latence (M_{max} dans les Algorithmes 1, 2 et 3). La magnitude majorée est fixe durant toute l'étude, à l'inverse de la magnitude maximale M_{max} qui, à cause du bruitage des magnitudes, peut changer d'un catalogue bruité à l'autre. Il est évident que M_{max} doit toujours être inférieure ou égale à M_{maj} . Étant donné que la magnitude est tirée dans une loi normale de moyenne M_w et d'écart-type $ErrM_w/2$, 5% des magnitudes tirées sont en dehors de l'intervalle $M_w \pm ErrM_w$. La loi normale étant symétrique, il y a donc 2,5% de chance de tirer une magnitude supérieure à $M_w + ErrM_w$. Ainsi, dans certains cas particuliers la magnitude tirée est supérieure à la magnitude majorée. Si cela se produit, la magnitude est retirée jusqu'à ce qu'elle respecte la condition $M_{max} \leq M_{maj}$.

2.4 Discussions et Conclusion

La méthode des temps de latence sert à estimer la distribution fréquence - magnitude des séismes principaux à partir de leur distribution proportion - magnitude. Cette dernière est calculée à l'aide des temps de latence entre deux séismes successifs, normalisés par le temps de latence moyen. Comme pour toutes les méthodes statistiques, la méthode des temps de latence montre des difficultés à estimer des proportions non biaisées lorsque les données sont peu nombreuses, i.e., pour les fortes magnitudes.

L'implémentation MIMP, pour *Magnitude of Inferred Main shock Proportion*, est proposée dans ce Chapitre afin de pallier ce problème. Cette implémentation corrige la proportion de séismes principaux calculée à partir d'une valeur seuil de proportion, appelée seuil MIMP. À partir de ce seuil, la proportion est sensée se stabiliser entre cette valeur seuil et la valeur maximale 1. Si à partir d'une certaine magnitude MIMP des instabilités sont observées alors les proportions associées aux magnitudes supérieures ou égales à MIMP sont fixées à 1. Autrement dit, à partir de la magnitude MIMP tous les séismes sont considérés comme étant des séismes principaux. Le seul paramètre nécessaire à l'implémentation MIMP est le seuil MIMP.

En première approche, la magnitude maximale du territoire d'étude est estimée de manière déterministe comme étant la somme maximale des magnitudes du catalogue et de leur incertitude (Équation 2.3). Ce choix est critiquable pour deux raisons. La première est que la magnitude maximale observée dans un catalogue n'est pas forcément la magnitude maximale que l'on peut atteindre, notamment si la période de retour d'une telle magnitude est supérieure à la profondeur temporelle du catalogue. L'ajout de l'incertitude pour estimer la magnitude majorée est une mesure de précaution mais ne constitue pas une certitude d'avoir atteint la vraie magnitude maximale. La deuxième raison est que la magnitude maximale est estimée de façon déterministe, sans incertitude associée. Ainsi, la variabilité de la magnitude maximale n'est pas traitée dans cette thèse. Une solution pour prendre en compte cette variabilité est de définir

cette magnitude à l'aide d'analyses Bayésiennes comparatives avec des régions analogues [par ex., Ameri et al., 2015, Zentner, 2018], d'analyses du cumul du moment sismique [Makropoulos et Burton, 1983] ou encore d'analyses intégratives [Kijko, 2004]. Ces analyses peuvent définir une M_{maj} fixe associée à une incertitude ou alors définir M_{maj} par une densité de probabilité. Elles permettent donc de prendre en compte la variabilité de M_{maj} . Nuancions néanmoins ces limites par le fait que plus les périodes de retour étudiées sont élevées, plus l'impact du choix de la magnitude maximale sur l'aléa est fort. Or, les périodes de retour analysées dans cette thèse sont inférieures à 2 500 ans. À ce niveau-là, le choix de la magnitude maximale a *a priori* un moindre impact. Pour cette même raison, le choix de la magnitude maximale de la récente étude française de Drouet et al. [2020] s'est également porté sur une valeur constante.

La méthode des temps de latence constitue un modèle de récurrence alternatif au couple algorithme de *declustering*-loi GR classiquement utilisé. À l'inverse de l'algorithme de *declustering*, la méthode des temps de latence ne permet pas de distinguer les séismes principaux des répliques, elle calcule seulement une proportion de séismes principaux en fonction de la magnitude. De plus, elle estime la DFM dans son intégralité, i.e., pour chaque pas de magnitude. Aucune extrapolation de coefficients des faibles magnitudes aux fortes magnitudes n'est réalisée, contrairement à l'application classique de la loi GR. Cela a pour avantage de reproduire fidèlement ce que les données nous montrent. En revanche, pour que les résultats soient dignes de confiance, les données doivent l'être également. En effet, la profondeur temporelle de ces dernières ainsi que leur exhaustivité doivent être assez élevée afin d'échantillonner, à l'aide des incertitudes sur la magnitude, les plus fortes magnitudes du territoire étudié. Cette condition empêche son application lorsque les données ne sont pas assez représentatives, où l'application de la loi GR lui est préférée.

Chapitre 3

Générateur de séismes en contexte de sismicité faible à modérée

Sommaire

3.1	Introduction	50
3.2	Générateur de séismes principaux	51
3.2.1	Nombre de séismes principaux par an et par magnitude	51
3.2.2	Localisation des séismes principaux en fonction de la magnitude	52
3.2.3	Définition des autres paramètres du séisme principal	53
3.3	Générateur de répliques	55
3.3.1	Magnitude maximale des répliques	55
3.3.2	Tirage des répliques produites par un séisme principal	56
3.3.3	Définition des paramètres des répliques	59
3.4	Discussions et Conclusion	62

Résumé Dans ce Chapitre, je présente le fonctionnement du générateur de séismes adapté au manque de données sismiques lié au contexte des régions stables. Ce générateur nécessite une distribution fréquence - magnitude (DFM) de séismes principaux, une régionalisation du territoire étudié et une densité spatiale caractérisant sa sismicité. Les séismes principaux sont tout d'abord générés dans l'espace temps-magnitude grâce à la DFM, puis ils sont spatialement implantés à l'aide de la densité spatiale et de la régionalisation. Cette dernière limite la magnitude maximale dans l'espace tandis que la densité spatiale indique la répartition spatiale des séismes dans chaque région. Seules les régions où la magnitude maximale est supérieure ou égale à la magnitude du séisme à planter sont considérées. La variabilité spatiale de la sismicité est ainsi respectée grâce à la magnitude qui joue le rôle de liant entre le tirage temporel et le tirage spatial des séismes principaux. Enfin, une série de répliques est simulée autour de chaque séisme principal à l'aide de la loi de Bâth et de la distribution proportion - magnitude de répliques fournie en entrée du générateur.

3.1 Introduction

L'objectif de cette thèse est de produire des catalogues synthétiques sur de longues périodes temporelles afin d'échantillonner l'ensemble des sources sismiques possibles. Pour cela, un générateur de séismes synthétiques est utilisé. Ce générateur a besoin d'informations concernant la temporalité de la sismicité du territoire d'étude ainsi que d'informations décrivant la répartition spatiale des séismes. Le premier type d'informations est regroupé dans un modèle de récurrence (Section 1.4.2) tandis que le deuxième est décrit dans un modèle source (Section 1.4.1).

Dans le Chapitre 2, un modèle de récurrence utilisant la méthode des temps de latence a été présenté à cet effet. L'analyse d'un catalogue répertoriant les séismes du territoire d'étude permet de décrire le comportement temporel des séismes principaux dudit territoire à travers une distribution fréquence-magnitude (DFM). Néanmoins, ce modèle requiert une quantité importante de données afin d'aboutir à des résultats fiables. De plus, la méthode des temps de latence calcule la DFM pour chaque pas de magnitude, sans extrapolation. Les données sismiques doivent donc être complètes pour toutes les magnitudes étudiées. L'approche proposée dans ce Chapitre est d'appliquer ce modèle de récurrence à l'échelle du territoire entier dans le but de conserver un maximum de données pour estimer au mieux la DFM de séismes principaux. Cela implique que la distribution temporelle des séismes principaux est la même sur l'entièreté du territoire. La sismicité n'étant pas stationnaire, une méthodologie se basant sur une densité spatiale et une régionalisation est proposée Section 3.2. Pour cela, la densité spatiale est définie par région en fonction de la magnitude.

La Section 3.3 détaille, quant à elle, comment des répliques sont générées autour des séismes principaux avec le moins de paramètres possible. Cela est réalisé grâce à la loi de Bâth et de la proportion de répliques calculées en amont (Chapitre 2).

3.2 Générateur de séismes principaux

3.2.1 Nombre de séismes principaux par an et par magnitude

Comme pour la méthode classique des catalogues synthétiques [par ex., Musson, 1999, Ebel et Kafka, 1999], le nombre annuel de séismes principaux est tiré année après année dans une loi de Poisson de paramètre λ_{i, M_w} . Le nombre N_a d'années est défini par l'utilisateur.

Dans cette thèse, λ_{i, M_w} correspond à la fréquence annuelle de séismes principaux. Il est obtenu à partir des DFM de séismes principaux calculées Chapitre 2. Les fréquences annuelles étant dépendantes de la magnitude, λ_{i, M_w} l'est aussi. De plus, les DFM sont exprimées pour chaque pas de magnitude sous la forme d'une densité de probabilité obtenue grâce à la prise en compte des incertitudes sur la magnitude (Section 2.3). Ainsi, pour chaque pas de magnitude M_w et pour chaque année i , une fréquence annuelle $F_i(\geq M_w)$ peut être tirée dans la densité de probabilité associée à l'aide de la méthode de rejet. Le λ_{i, M_w} est ensuite calculé tel que :

$$\forall i \in [1, N_a], \forall M_w \in [M_c, M_{maj}], \quad \lambda_{i, M_w} = \begin{cases} F_i(\geq M_w) & \text{si } M_w = M_{maj} \\ F_i(\geq M_w) - F_i(\geq M_w + dM_w) & \text{sinon} \end{cases}, \quad (3.1)$$

avec dM_w le pas de magnitude, M_c la magnitude de coupure, i.e., la magnitude minimale de l'étude, et M_{maj} la magnitude maximale de l'étude. Les densités de probabilité ne donnent pas la fréquence annuelle des séismes principaux à magnitude égale à M_w mais celle des séismes principaux à magnitude supérieure ou égale à M_w (Équation 2.2). Afin d'obtenir le taux de séismes principaux de magnitude M_w , il est nécessaire de soustraire de celle associée à M_w celle associée à la magnitude supérieure ($M_w + dM_w$ Équation 3.1). Dans certains cas, $F_i(\geq M_w)$ est plus faible que $F_i(\geq M_w + dM_w)$, auquel cas λ_{i, M_w} est négatif. Dans ces cas de figure, $F_i(\geq M_w)$ est retirée dans sa densité de probabilité associée jusqu'à ce que λ_{i, M_w} soit positif.

Finalement, le nombre n_{i, M_w} de séismes principaux de magnitude M_w ayant lieu à l'année i est obtenu comme suit :

$$\forall i \in [1, N_a], \forall M_w \in [M_c, M_{maj}], n_{i, M_w} \sim \text{Poisson}(\lambda_{i, M_w}). \quad (3.2)$$

Au final, le résultat de cette étape est une liste de séismes principaux définis par leur année i d'occurrence et leur magnitude M_w .

La DFM utilisée pour calculer le nombre de séismes principaux par an représente la sismicité totale du territoire étudié. Ainsi, le tirage du nombre de séismes principaux est réalisé sur l'ensemble du territoire, sans prendre en compte les variabilités spatiales qu'il peut exister entre différentes régions dudit territoire.

3.2.2 Localisation des séismes principaux en fonction de la magnitude

Le but du tirage spatial est de localiser les séismes principaux déjà générés en temps et en magnitude sur l'entièreté du territoire. Cette localisation doit (i) prendre en compte la variabilité spatiale de la sismicité et (ii) assurer la cohérence de la localisation des séismes aux magnitudes les plus élevées. En effet, comme la définition temporelle des séismes principaux se fait à l'échelle du territoire entier, il est nécessaire de limiter la distribution spatiale des magnitudes extrêmes. À titre d'exemple, un séisme $M_w 6,8$ ne peut physiquement pas survenir n'importe où en France métropolitaine.

Une carte de densité spatiale caractérisant la sismicité est utilisée pour positionner les séismes principaux. C'est elle qui permet de respecter la première condition, à savoir de prendre en compte la variabilité spatiale de la sismicité. La régionalisation permet, quant à elle, de respecter la deuxième condition : positionner les séismes là où il peuvent physiquement se produire. Pour cela, chaque région se voit attribuer une magnitude maximale $M_{max Région}$. Elle peut être fixe, exprimée par un intervalle de valeur ou encore décrite à travers une densité de probabilité [par ex., Ameri et al., 2015, Zentner, 2018].

Le couplage entre la régionalisation et la carte de densité aboutit à un ensemble de régions définies par leur propre densité spatiale et par une magnitude maximale autorisée. La localisation d'un séisme de magnitude M_w généré à l'année i ne peut se faire que dans les régions éligibles (Figure 3.1), i.e., les régions où M_w est inférieure ou égale à $M_{max Région}$. Une fois l'ensemble des régions éligibles défini, le tirage spatial dudit séisme principal est réalisé dans les densités spatiales propres aux régions éligibles (Figure 3.1).

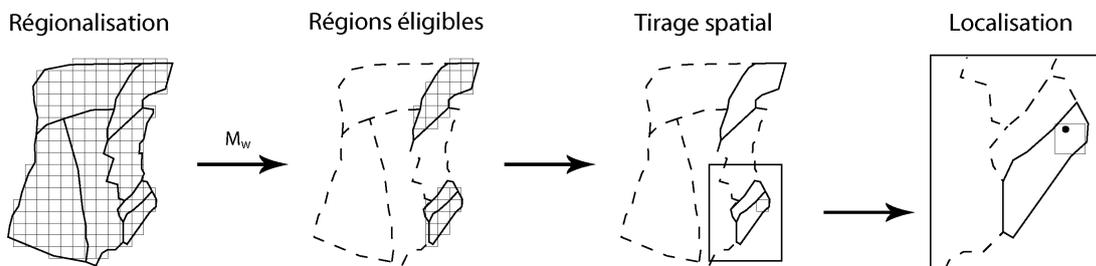


Figure 3.1: SCHÉMA GÉNÉRAL DU TIRAGE SPATIAL.

Une densité spatiale est définie sur une grille cartésienne. Chaque région est définie par sa densité spatiale et sa magnitude maximale. La magnitude M_w du séisme généré limite le nombre de régions éligibles. Le tirage dans la densité spatiale est ensuite réalisé uniquement dans les régions éligibles. Une fois une maille de la grille tirée, les coordonnées épicentrales du séisme (point noir) sont tirés de manière uniforme au sein de cette maille.

Dans cette thèse, la densité spatiale est calculée sur une grille cartésienne (Figure 3.1). Le tirage spatial réalisé ci-dessus consiste donc à tirer une maille de cette grille. Une fois une maille tirée, la localisation épicentrale du séisme est tirée de manière uniforme au sein de cette maille. La Figure 3.1 résume la procédure du tirage spatial.

Grâce à cette procédure, la distribution spatiale des séismes principaux générés est physiquement limitée par la magnitude et respecte la densité spatiale utilisée pour caractériser la sismicité.

3.2.3 Définition des autres paramètres du séisme principal

Une fois les séismes générés en temps, en magnitude et dans l'espace, d'autres caractéristiques leur sont associées. Ces dernières comprennent l'azimut, le pendage et le régime (normal, inverse, décrochant) de faille ainsi que la profondeur de l'hypocentre. Pour chaque région, l'ensemble de ces paramètres sont tirés dans des intervalles de densité de probabilité uniformes définis au préalable à l'aide des données disponibles. Ces données sont principalement des catalogues de mécanismes au foyer (profondeur, azimut, pendage, régime), des catalogues sismiques (profondeur) ou encore des cartes de tracés de failles (azimut). Les cartes de tracés de failles (actives ou non), issues de cartes géologiques, sont riches d'information car les tracés répertoriés sont généralement nombreux même au sein de régions peu actives. Cependant, ces tracés en surface ne donnent pas d'information sur le sens du plongement des failles. Seule l'orientation (entre 0 et 180°) est disponible et non pas l'azimut (entre 0 et 360°). Ce problème peut être surmonté en première approche en choisissant aléatoirement le sens du plongement. Par exemple, si l'orientation de la faille est de 75°, l'azimut est égal à 75° ou à $75 + 180 = 255^\circ$.

Un exemple de paramétrisation des régions est présentée Section 5.2.2.2 pour une application à la France métropolitaine.

Dans cette thèse, la régionalisation sert à limiter la répartition spatiale des séismes générés les plus puissants et à définir les caractéristiques de leur plan de rupture. Ainsi, cette régionalisation n'est pas utilisée dans le but de définir une distribution fréquence-magnitude de séismes principaux dans chaque région. Elle est donc à différencier du zonage sismotectonique défini Section 1.4.1.1. En effet, la régionalisation n'a pas besoin d'être définie par la composante temporelle de la sismicité contrairement au zonage sismotectonique. De ce fait, des régionalisations moins complexes que les zonages sismotectoniques peuvent être utilisées. Une régionalisation peut par exemple être formulée uniquement par la tectonique globale : compression, extension, volcanisme, croûte océanique ou encore croûte continentale stable [Woessner et al., 2015, Drouet et al., 2020].

L'Algorithme 4 présente la structure générale du générateur de séismes principaux.

Algorithme 4 : GÉNÉRATEUR DE SÉISMES PRINCIPAUX

Données : DFM de séismes principaux (Chapitre 2),
 Densité spatiale,
 Régionalisation

Résultat : Liste des séismes principaux générés

```

1 pour chaque magnitude  $M_w$  allant de  $M_{maj}$  à  $M_c$  faire
2   pour chaque annee  $i$  allant de 1 à  $N_a$  faire
3     Tirage temporel :
4      $\lambda_{i, M_w}$  est tiré dans la distribution de fréquences annuelles de séismes
5     principaux associée à  $M_w$  (Équation 3.1) ;
6     Le nombre  $n_{i, M_w}$  de séisme de magnitude  $M_w$  à l'année  $i$  est tiré dans
7     une distribution Poissonienne de paramètre  $\lambda_{i, M_w}$  (Équation 3.2) ;
8     Tirage spatial :
9     si  $n_{i, M_w}$  est supérieur à 0 alors
10      Toutes les régions dont  $M_{max\ Region} \geq M_w$  sont considérées comme
11      éligibles ;
12      La localisation du/des séisme(s) généré(s) est tirée dans la densité
13      spatial au sein des régions éligibles uniquement ;
14      La profondeur, l'azimut, le pendage et le régime de faille sont tirés
15      aléatoirement dans leur intervalle de valeurs ou bien dans leur
16      densité de probabilité correspondant à la région ;
17      L'ensemble des paramètres (année, magnitude, localisation,
18      profondeur, azimut, pendage, régime et région) du/des séisme(s)
19      généré(s) sont conservés dans la Liste des séismes principaux ;
20    fin
21  fin
22 fin

```

3.3 Générateur de répliques

Cette partie a pour objectif de présenter la production de répliques autour de chaque séisme principal généré en amont. Cette production se veut simple, réaliste et dépendante du moins de paramètres possible. Elle s'appuie tout d'abord sur la loi de Båth [Richter, 1958, Båth, 1965] afin de trouver la borne supérieure de la magnitude des répliques (Section 3.3.1). Le nombre de répliques (Section 3.3.2) produites par un séisme principal est gouverné par la magnitude via l'utilisation de la distribution proportion-magnitude de répliques, complémentaire de celle de séismes principaux.

3.3.1 Magnitude maximale des répliques

La loi de Båth [Richter, 1958, Båth, 1965] (Équation 3.3) stipule que le séisme principal d'une séquence sismique est le séisme ayant la plus forte magnitude. Aussi, cette loi définit une borne supérieure $M_{r\ max}$ aux magnitudes des répliques provoquées par un séisme principal. Cette valeur est égale à la magnitude M_{sp} du séisme principal corrigée d'un terme ΔM :

$$M_{r\ max} = M_{sp} - \Delta M . \quad (3.3)$$

Le terme ΔM est en moyenne égal à 1,2 pour les séismes crustaux [Richter, 1958] dont la profondeur est inférieure à 50 km. Concrètement, selon cette relation, un séisme de magnitude 7,2 ne pourrait engendrer que des répliques de magnitudes inférieures ou égales à 6. Certains auteurs [par ex., Helmstetter et Sornette, 2003a, Console et al., 2003] ont montré que l'indépendance de $\Delta M = 1,2$ vis-à-vis de la magnitude M_{sp} du séisme principal ne s'observe que lorsque $M_{sp} \geq M_c + 2$. Il est donc nécessaire de définir une magnitude minimale $M_{sp\ min}$ de séisme principal capable de produire des répliques. Dans cette thèse, nous choisissons $M_{sp\ min} = M_c + 1$.

L'Équation 3.3 est un modèle empirique et donc ne se vérifie pas toujours. Par exemple, le séisme $M_w 6,5$ d'août 2016 à Amatrice (Italie) a provoqué une réplique de magnitude 6,1 en octobre 2016. Afin de prendre en compte la variabilité de ΔM dans la génération des répliques, les travaux de Zakharova et al. [2013] sont exploités. Ces auteurs calculent ce terme tel que :

$$\Delta M = -\frac{\log_{10}(R)}{1,5} , \quad (3.4)$$

où R correspond au moment sismique libéré par l'ensemble des répliques $\sum M_{0r}$ soustrait du moment sismique libéré par l'ensemble des séismes précurseurs $\sum M_{0p}$ et divisé par le moment sismique $M_{0\ sp}$ libéré par le séisme principal :

$$R = \frac{\sum M_{0r} - \sum M_{0p}}{M_{0\ sp}} \approx 0,05 . \quad (3.5)$$

D'après ces auteurs, R vaut en moyenne 5%. Le moment sismique M_0 (N.m) est obtenu à partir de la magnitude de moment M_w d'après Hanks et Kanamori [1979] :

$$M_0 = 10^{1,5 M_w + 9,1} . \quad (3.6)$$

L'échantillonnage de R dans une loi normale de moyenne 5% et d'écart-type 1,25% [Zakharova et al., 2013] aboutit à une distribution de ΔM à l'aide de l'Équation 3.4. La Figure 3.2 illustre une telle distribution avec 1000 échantillons. La moyenne et la médiane de cette distribution oscillent autour de 0,87. 90% des ΔM sont compris entre 0,77 (Q5) et 1,02 (Q95). Ces valeurs sont inférieures à ce que préconisent la loi de Bâth originelle : 1,2 [Richter, 1958].

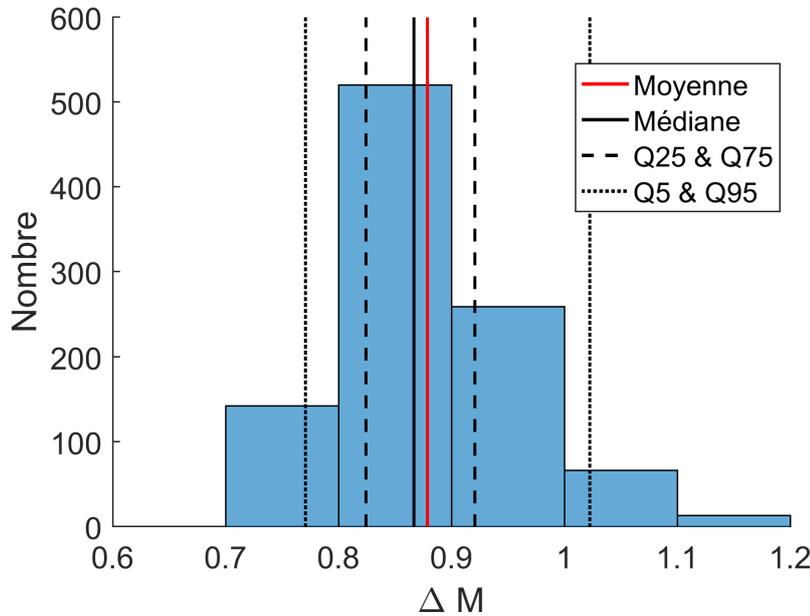


Figure 3.2: CALCUL DE ΔM À L'AIDE DU RATIO R DES MOMENTS SISMIQUES. ΔM est calculé à l'aide de l'Équation 3.4. Pour se faire, le ratio R entre le moment sismique libéré par l'ensemble des répliques et celui libéré par le séisme principal est échantillonné 1000 fois dans une loi normale de moyenne 5% et d'écart-type 1,25%.

Zakharova et al. [2013] introduisent une correction au terme R afin de prendre en compte la dépendance de ΔM envers M_{sp} lorsque cette dernière se rapproche de la magnitude de coupure [par ex., Helmstetter et Sornette, 2003a, Console et al., 2003]. Cette correction impacte peu la distribution de ΔM : il vaut en moyenne 0,77 à $M_{sp} = M_c + 1$ avec la correction contre 0,87 sans la correction. De plus, des paramètres tels que le régime de faille peuvent légèrement influencer la valeur de R et donc directement ΔM [Zakharova et al., 2013]. Dans cette thèse, ces variations ne sont pas prises en compte : R suit toujours une loi normale de moyenne 5% et d'écart-type 1,25%. Ainsi, la distribution de ΔM visible Figure 3.2 est utilisée dans cette thèse.

3.3.2 Tirage des répliques produites par un séisme principal

Le tirage des répliques produites par un séisme principal de magnitude $M_{sp} \geq M_{spmin}$ se base sur la proportion de répliques en fonction de la magnitude. Le tirage prend fin lorsque le moment sismique cumulé par l'ensemble des répliques produites atteint une valeur seuil définie à partir de la magnitude du séisme principal. Ce tirage s'articule autour de trois étapes décrites ci-dessous.

Gamme de magnitudes possibles pour les répliques La première étape consiste à définir l'ensemble des magnitudes auxquelles les répliques du séisme principal peuvent être associées. Cette gamme de magnitudes est comprise entre la magnitude de coupure M_c et la magnitude maximale des répliques $M_{r\max}$. Cette dernière est définie par les travaux réalisés dans la Section précédente. Tout d'abord, un R est échantillonné dans une loi normale de moyenne 5% et d'écart-type 1,25%. ΔM est ensuite obtenu à l'aide de R via l'Équation 3.4 (Figure 3.2). Enfin, la magnitude maximale $M_{r\max}$ des répliques est calculée à l'aide de l'Équation 3.3.

Tirage des répliques en fonction de leur magnitude Le tirage se base sur l'hypothèse que la probabilité de tirer une réplique de magnitude M_w est égale à la proportion $p_r(M_w)$ de répliques de cette magnitude. Cette proportion est directement obtenue via la DPM de séismes principaux calculée Chapitre 2. En effet, la proportion $p_r(\geq M_w)$ de répliques de magnitude supérieure ou égale à M_w est complémentaire de $p_{sp}(\geq M_w)$, la proportion de séismes principaux :

$$p_r(\geq M_w) = 1 - p_{sp}(\geq M_w). \quad (3.7)$$

En première approche, nous faisons l'hypothèse que $p_r(M_w) = p_r(\geq M_w)$.

Pour réaliser le tirage, une magnitude M_w est tirée uniformément entre M_c et $M_{r\max}$. Une réplique de magnitude M_w est associée au séisme principal si et seulement si $p_r(M_w)$ vérifie la condition suivante :

$$p_r(M_w) > rand(0, 1), \quad (3.8)$$

avec $rand(0, 1)$ un réel tiré uniformément entre 0 et 1. Cette opération qui consiste à tirer une magnitude M_w puis à vérifier l'Équation 3.8 afin de générer ou non une réplique de magnitude M_w est répétée jusqu'à ce que la condition d'arrêt présentée ci-dessous soit atteinte.

Dans cette thèse, de par l'approche stochastique proposée Section 2.3, $p_r(M_w)$ n'est pas constante mais est définie par une densité de probabilité. Pour une même magnitude, $p_r(M_w)$ varie et sa valeur à une probabilité $P(p_r(M_w))$ d'être tirée dans la DPM stochastique de répliques. D'après la condition de tirage (Équation 3.8), la probabilité $P_r(M_w)$ de tirer une réplique de magnitude M_w est donc égale à l'intégrale des $p_r(M_w)$ multipliées par leur probabilité $P(p_r(M_w))$ d'être tirée dans la DPM stochastique de répliques :

$$P_r(M_w) = \int_0^1 p_r(M_w) \times P(p_r(M_w)) dp_r(M_w). \quad (3.9)$$

Condition d'arrêt du tirage des répliques Comme le tirage des répliques est répété en boucle, il est nécessaire d'introduire une condition d'arrêt afin de ne pas répéter à l'infini cette opération. Cette condition est définie par le moment sismique maximal libéré par l'ensemble des répliques $\sum M_{0r}$, obtenu à l'aide des travaux réalisés dans la Section précédente. En effet, en considérant uniquement les répliques dans l'Équation 3.5 ($\sum M_{0p} = 0$), $\sum M_{0r}$ est calculé en multipliant le moment sismique du séisme principal M_{0sp} par le ratio R tiré précédemment. Le tirage des répliques du

séisme principal prend fin une fois que la somme des moments sismiques de l'ensemble des répliques produites a atteint ou dépassé la valeur limite $\sum M_{0r}$. Autrement dit, pour chaque nouvelle réplique produite, son moment sismique est additionné à l'ensemble des moments sismiques libérés par les répliques déjà produites. Si cette somme atteint ou dépasse la valeur limite $\sum M_{0r}$, alors le tirage des répliques prend fin. Cela signifie que si la première réplique à être tirée a une magnitude égale à la magnitude maximale des répliques (M_{rmax}), alors seule cette réplique est générée. La méthode proposée par Zakharova et al. [2013] permet donc d'estimer à la fois la magnitude maximale des répliques qu'un séisme principal peut produire (Section 3.3.1) et le nombre de répliques que ce séisme peut produire.

Cette façon originale de tirer les répliques d'un séisme principal s'affranchit d'un nombre conséquent de paramètres. En effet, seuls le ratio R (Équation 3.5) et la DPM de répliques (Chapitre 2) sont nécessaires. De plus, ces deux paramètres sont définis de manière stochastique, permettant la prise en compte de leur variabilité.

Exemples de tirage de répliques La Figure 3.3 illustre deux exemples de tirage de répliques produites par un séisme principal de magnitude 6,8. Dans le premier exemple (a), la magnitude M_{rmax} maximale des répliques est égale à 5,9. Ainsi, la valeur limite $\sum M_{0r}$ du cumul du moment sismique libéré par les répliques est d'environ $8,9 \times 10^{17}$ N.m. La première réplique tirée a une magnitude égale à 5,8. Le moment sismique libéré par cette seule réplique ($6,3 \times 10^{17}$ N.m) atteint presque la valeur maximale $\sum M_{0r}$. Le tirage de cette forte réplique empêche donc de tirer un grand nombre d'autres répliques $M_w \geq 4$: seulement quatre. À l'inverse, dans le deuxième exemple (b), la magnitude M_{rmax} maximale des répliques est égale à 6 et toutes les répliques tirées ont une magnitude inférieure ou égale à 5,6. Ces relativement faibles magnitudes permettent de produire une quantité plus importante de répliques $M_w \geq 4$ (39).

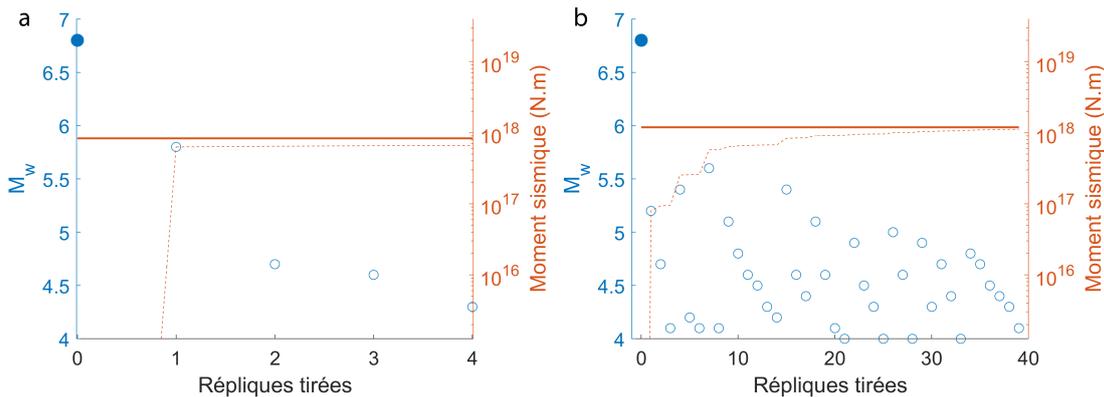


Figure 3.3: EXEMPLES DE TIRAGE DES RÉPLIQUES D'UN SÉISME PRINCIPAL $M_{sp}6,8$. Un séisme de magnitude 6,8 (cercle plein) produit dans ces deux exemples quatre (a) et 39 (b) répliques de magnitudes supérieures ou égales à 4 (cercles vides). Le tirage de ces répliques s'arrête lorsque le cumul du moment sismique des répliques produites (tirets) dépasse la valeur limite $\sum M_{0r}$ (trait plein).

Dans les deux exemples visibles Figure 3.3, la cinquième (a) et la quarantième (b) répliques produites amènent le moment sismique cumulé à dépasser la valeur maximale $\sum M_{0r}$. Ces dernières répliques ne sont donc pas générées et le tirage des répliques prend fin pour ce séisme principal.

3.3.3 Définition des paramètres des répliques

La résolution temporelle du générateur de séismes est l'année (Équation 3.2) compte tenu du contexte opérationnel de cette thèse. Les répliques sont générées la même année que le séisme principal auquel elles sont liées. Ainsi, pour un séisme principal donné, chacune de ses répliques ne possède pour l'heure que deux paramètres : sa magnitude M_w et son année d'occurrence.

En ce qui concerne les autres paramètres tels que la localisation ou encore les paramètres de faille, nous faisons l'hypothèse que les répliques ont lieu sur la même faille que le séisme principal. Dans cette optique, le régime de faille de la réplique est considéré comme étant identique à celui du séisme principal.

Pour ce qui est de la localisation, il est courant (par ex., la majorité des applications du modèle ETAS) de placer les répliques de manière isotrope autour du séisme principal. Ceci étant dû au fait qu'une grande partie des algorithmes de *declustering* appliqués dans les études de l'aléa sismique utilisent des fenêtres isotropes [par ex., Gardner et Knopoff, 1974, Grünthal, 1985]. Néanmoins, certains algorithmes [par ex., Reasenber, 1985] prennent en compte l'anisotropie parfois observée sur les données réelles [par ex., Shearer, 2005] et principalement induite par la direction du plan de rupture. Nous pensons qu'il est préférable de prendre en compte cette anisotropie lors de la localisation des répliques [par ex., Hainzl et al., 2010]. Nous considérons donc le modèle géométrique suivant : la localisation (x_r, y_r) des répliques est tirée à plus ou moins $0,75 L$ de l'épicentre (x, y) du séisme principal dans la direction de son azimut (azi) à 10° près :

$$\begin{aligned} x_r &\sim N(x; 0,75 L \cos(-(azi \pm 10 - 90))) \\ y_r &\sim N(y; 0,75 L \sin(-(azi \pm 10 - 90))) \end{aligned} \quad (3.10)$$

où L est la longueur du plan de rupture associé au séisme principal. Cette longueur est estimée en fonction de la magnitude du séisme principal [par ex., Wells et Coppersmith, 1994, Leonard, 2014] :

$$L = 10^{(M_w - l_1)/l_2}, \quad (3.11)$$

avec l_1 et l_2 des constantes. La correction de 90° dans l'Équation 3.10 permet de passer de l'espace trigonométrique à l'espace géographique. Des exemples fictifs de distribution spatiale de répliques autour du séisme principal sont illustrés Figure 3.4. À noter que ni le pendage, ni la profondeur du séisme principal ne sont pris en compte dans la répartition spatiale des répliques.

La profondeur z_r (km), l'azimut azi_r ($^\circ$) ainsi que le pendage dip_r ($^\circ$) du plan de rupture associé à chaque réplique sont, quant à eux, échantillonnés dans des lois normales. Ces lois ont comme moyenne la valeur associée au séisme principal (z , azi et

dip) et un écart-type dépendant du paramètre en question :

$$\begin{aligned} z_r &\sim N(z; 2, 5) \geq 0 \\ azi_r &\sim N(azi; 5) \geq 0 \\ dip_r &\sim N(dip; 2, 5) \geq 0 \end{aligned} \quad . \quad (3.12)$$

Une condition est ajoutée à ce calcul afin de ne pas aboutir à des valeurs négatives. Si c'est le cas, la valeur est retirée jusqu'à ce qu'elle soit positive. La légère variabilité apportée dans l'Équation 3.12 traduit le fait que l'azimut et le pendage d'une faille peuvent évoluer le long de celle-ci étant donné que les failles ne sont pas des objets linéaires mais tortueux et complexes.

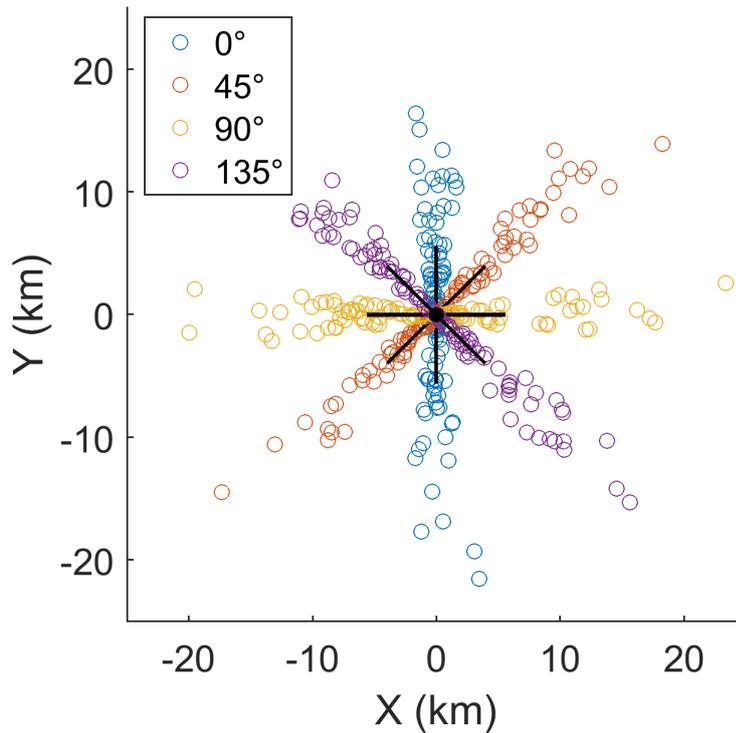


Figure 3.4: DISTRIBUTION SPATIALE DE RÉPLIQUES AUTOUR DU SÉISME PRINCIPAL. Exemples fictifs de localisation épiscopentrale de cent répliques autour d'un séisme principal de magnitude 6 dont l'épicentre est signalé par le point noir. La localisation des épiscopentres des répliques est réalisée via l'Équation 3.10 et à partir de quatre azimuts différents. Les lignes noires représentent les longueurs des plans de rupture associés au séisme principal (Équation 3.11 Leonard [2014]) et orientés suivant les quatre azimuts.

L'Algorithme 5 résume la génération des répliques à partir des séismes principaux.

Algorithme 5 : GÉNÉRATEUR DE RÉPLIQUES

Données : Liste des séismes principaux (Section 3.2),
 M_{spmin}

Résultat : Liste des répliques générées

1 **pour chaque** *séisme principal dans la Liste des séismes principaux* **faire**

2 M_{sp} reçoit la magnitude du séisme principal ;

3 **si** M_{sp} est supérieure ou égale à M_{spmin} **alors**

4 **Initialisation :**

5 R est tiré dans une loi normale de moyenne 5% et d'écart-type 1,25% ;

6 M_{0sp} est calculé à partir de M_{sp} (Équation 3.6) ;

7 $\sum M_{0r}$ est obtenu en multipliant R et M_{0sp} (Équation 3.5) ;

8 ΔM est estimé en fonction de R (Équation 3.4) ;

9 $M_{r,max}$ reçoit $M_{sp} - \Delta M$ (Équation 3.3) ;

10 **Tirage des répliques :**

11 **tant que** le moment sismique cumulé par les répliques générées est
inférieur à $\sum M_{0r}$ **faire**

12 M_w est tirée uniformément entre M_c et $M_{r,max}$;

13 Tirage de la proportion $p_{sp}(M_w)$ de séismes principaux de magnitude
 M_w ;

14 La proportion $p_r(M_w)$ de répliques de magnitude M_w reçoit
 $1 - p_{sp}(M_w)$;

15 Une réplique de magnitude M_w est tirée si la condition de
l'Équation 3.8 est vérifiée avec $p_r(M_w)$;

16 **Tirage des paramètres des répliques :**

17 **si** une réplique de magnitude M_w est tirée **alors**

18 L'année d'occurrence et le régime de faille de la réplique sont
identiques à ceux du séisme principal ;

19 La localisation (Équation 3.10), la profondeur, l'azimut et le
pendage (Équation 3.12) de la réplique sont tirés autour des
valeurs du séisme principal ;

20 Tous ces paramètres ainsi que M_w sont enregistrés dans la Liste
des répliques générées ;

21 **fin**

22 **fin**

23 **fin**

24 **fin**

3.4 Discussions et Conclusion

L'approche présentée dans ce Chapitre permet de produire des catalogues synthétiques de sismicité (séismes principaux et répliques) à partir de cinq paramètres. Ces paramètres sont des distributions proportion - magnitude (DPM) et fréquence - magnitude (DFM) de séismes principaux, une régionalisation, une densité spatiale caractérisant la sismicité et un ratio R entre les moments sismiques libérés par l'ensemble des répliques et par le séisme principal. La DFM de séismes principaux, la régionalisation et la densité spatiale servent à produire les séismes principaux tandis que la DPM de séismes principaux et le ratio R sont utilisés pour générer les répliques.

Un couplage original entre les caractéristiques temporelles et spatiales est appliqué afin de produire les séismes principaux. Ce couplage consiste à définir l'occurrence de ces séismes à partir d'une DFM de séismes principaux calculée à l'échelle du territoire d'étude tandis que leur localisation est définie à l'échelle des régions à l'aide de leur magnitude et d'une densité spatiale. Ainsi, la magnitude joue le rôle de liant entre la définition des caractéristiques temporelles et celle des caractéristiques spatiales des séismes principaux. L'estimation de la DFM de séismes principaux à l'échelle du territoire entier permet d'observer un maximum de données et donc de diminuer les périodes de retour comparé à une analyse réalisée à l'échelle des régions/zones. En supposant le catalogue sismique complet pour l'ensemble des magnitudes, cette approche permet d'aboutir à une distribution des fréquences de séismes principaux représentative de l'ensemble du territoire d'étude et ce pour chaque magnitude.

À l'image du modèle ETAS, les séismes principaux sont produits en premier lieu puis des répliques leur sont associées. En revanche, dans notre approche les répliques ne sont pas produites en cascade, seuls les séismes principaux peuvent en produire. De plus, ces séismes principaux ont forcément une magnitude supérieure à celles des répliques qu'ils produisent. En effet, la magnitude du séisme principal ainsi que le ratio R permettent de définir la magnitude maximale autorisée de ses répliques. Le nombre et la magnitude des répliques qu'un séisme produit dépendent, quant à eux, de ce même ratio R et de la DPM de séismes principaux fournie en entrée.

Ce générateur tend à proposer une approche alternative aux approches classiquement utilisées. Néanmoins, il constitue une première étape qui nécessite des améliorations à plusieurs niveaux.

Tout d'abord, une seule DFM de séismes principaux est utilisée afin de représenter la caractéristique temporelle de la sismicité à l'échelle du territoire entier. Cet aspect est expliqué par le souhait d'obtenir la DFM de séismes principaux la plus représentative possible de l'ensemble de la sismicité du territoire d'étude. Cependant, cela a pour conséquence d'aboutir à des catalogues synthétiques partageant la même pente (b) de DFM quelle que soit la région. Une approche par macro-régions [par ex., Grünthal et al., 2018] peut être envisagée afin de conserver une quantité de données suffisante tout en respectant la variabilité spatiale de la temporalité sismique.

Ensuite, la magnitude $M_{max\ Région}$ maximale autorisée dans chaque région joue un rôle important dans l'estimation de l'aléa sismique. Pour autant, tout comme pour la magnitude M_{maj} maximale de l'étude (Chapitre 2), $M_{max\ Région}$ est égale la magnitude

maximale du catalogue observée dans la région, à laquelle est ajoutée son incertitude. Chaque région est donc caractérisée par une seule valeur de $M_{max Région}$. La variabilité de $M_{max Région}$ n'est pour l'instant pas prise en compte.

Enfin, si nous avons tenté de reproduire les corrélations et les liens existants entre les paramètres décrivant la sismicité, nous n'avons pas pu tous les honorer. Certaines corrélations ne sont pas suffisamment bien connues et certains paramètres tels que les paramètres de failles, restent simulés de façon aléatoire et indépendante. Une relation entre le régime de faille et le pendage pourrait par exemple être envisagée. L'implémentation de telles relations permettrait de rendre le générateur plus réaliste.

Chapitre 4

Des séismes générés à l'aléa sismique

Sommaire

4.1	Introduction	66
4.2	Calcul de l'aléa sismique	67
4.2.1	Des séismes générés aux plans de rupture	67
4.2.2	Des plans de rupture à l'aléa sismique	69
4.3	Pondération des GMPE	72
4.3.1	Méthodologie	72
4.3.2	Calcul des poids	74
4.4	Discussions et Conclusion	75

Résumé Ce Chapitre présente comment l'aléa sismique est calculé dans cette thèse. Chaque séisme généré est tout d'abord transformé en plan de rupture à partir des paramètres qui lui sont associés. Ensuite, l'aléa sismique est calculé à l'aide d'un ensemble de modèles de prédiction des mouvements du sol pré-sélectionnés. Une méthodologie permettant de pondérer cet ensemble est proposée dans ce Chapitre. Cette pondération est dépendante des deux paramètres gouvernant le plus l'estimation de l'aléa sismique, à savoir la magnitude du séisme et la distance séisme-site. La pondération est réalisée à l'aide de données d'accélération. Chaque point de données est comparé avec l'aléa médian calculé par chaque modèle de prédiction afin d'estimer un poids à chacun d'entre eux. Une interpolation du plus proche voisin permet d'estimer les poids dans les domaines magnitude - distance pour lesquels aucune donnée n'est disponible.

4.1 Introduction

Dans cette thèse, l'aléa sismique considéré est l'accélération maximale du sol (*Peak Ground Acceleration*, PGA) estimée au rocher. L'aléa produit en surface par un séisme est calculé en chaque point d'une grille prédéfinie. Pour se faire, un modèle de prédiction des mouvements du sol ou *Ground-Motion Prediction Equation* (GMPE) est utilisé. Une GMPE permet de traduire les paramètres d'un séisme (*a minima* sa magnitude et sa distance au site) en aléa sismique en surface. La forte variabilité naturelle observée sur les PGA tend à associer à la valeur médiane du modèle de prédiction des mouvements du sol de fortes incertitudes aléatoires, appelées σ . De plus, de par les données utilisées, les hypothèses faites et le modèle de régression employé, chaque GMPE est unique et aboutit à une PGA médiane et un σ différents des autres GMPE. Ces différentes induisent des incertitudes épistémiques qu'il est nécessaire de prendre en compte en utilisant non pas une GMPE mais plusieurs.

Si plusieurs GMPE sont employées, il est indispensable de les hiérarchiser. Pour se faire, des analyses comparatives [voir Mak et al., 2017, pour une liste non exhaustive] peuvent être appliquées sur ces modèles. Si ces analyses comparatives sont utilisées lors de l'élaboration d'une nouvelle GMPE pour la comparer avec d'autres modèles de référence, cette pratique est en revanche moins courante dans les études PSHA, où les poids des GMPE sont souvent définis à dire d'experts. Des approches hybrides permettent de prendre en compte à la fois le jugement des experts et les résultats des analyses comparatives [par ex., Delavaud et al., 2012]. De plus, lorsque ces analyses comparatives sont réalisées, elles sont rarement dépendantes de la magnitude (par ex., analyses séparées $M_w < 5$ et $M_w \geq 5$ par Drouet et Cotton [2015]) et encore moins dépendantes de la distance. Ces paramètres sont pourtant ceux qui influencent le plus le calcul de l'aléa sismique.

La Section 4.3 s'applique à présenter une méthode de pondération s'appuyant sur une analyse résiduelle en fonction de la magnitude du séisme et de la distance séisme-site. Le calcul de l'aléa sismique est mis en contexte au sein du générateur de séismes synthétiques dans la Section 4.2.

4.2 Calcul de l'aléa sismique

4.2.1 Des séismes générés aux plans de rupture

Dans cette thèse, un séisme est considéré comme un point dont la position spatio-temporelle ainsi que les paramètres caractérisant sa rupture associée sont connus. Ces paramètres sont la magnitude de moment et l'azimut, le pendage et le régime de faille. À partir de cette représentation ponctuelle, le calcul de l'aléa sismique évalue l'impact du séisme en surface. Lors de ce calcul, la géométrie du plan de rupture associé à ce séisme n'est pas considérée. Elle peut cependant jouer un rôle important dans l'estimation de l'aléa en surface lorsqu'elle est prise en compte (Figure 4.1b). Il est donc préférable de tenir compte de cette géométrie à l'aide des paramètres décrits ci-dessus.

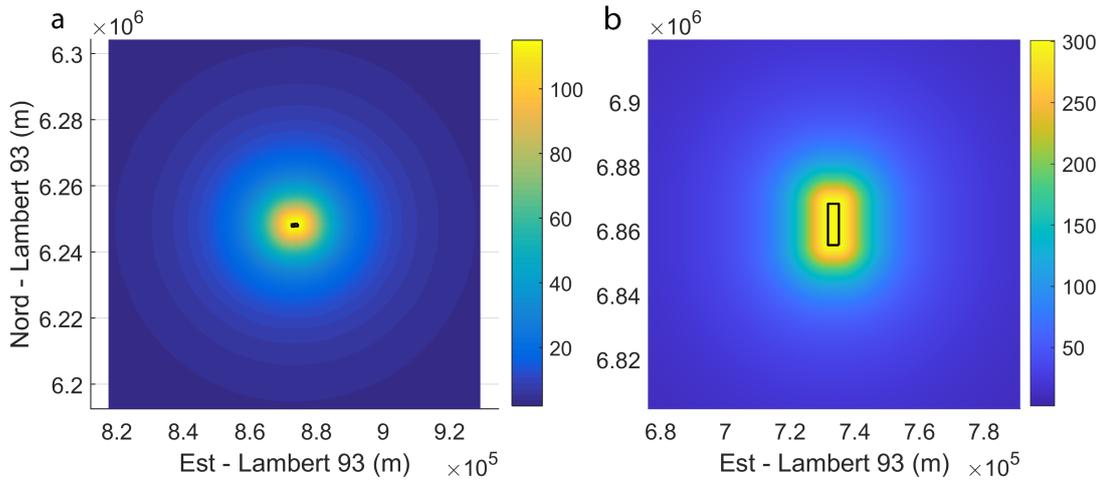


Figure 4.1: COMPARAISON DE PGA CALCULÉE POUR DEUX MAGNITUDES DIFFÉRENTES.

Comparaison de la PGA médiane (cm/s^2) produite en surface par deux séismes à régime décrochant de magnitude 4,7 (a) et 6,1 (b). Le calcul est réalisé à l'aide de la GMPE d'Akkar et al. [2014a] et de la distance Joyner-Boore (Section 4.2.2).

Les rectangles noirs représentent les projetés en surface des plans de rupture des séismes. Les longueurs de ces plans de rupture sont environ égales à 2 (a) et 13 km (b) (Équation 4.1). L'étendue spatiale de a et b est définie à l'aide de la distance d'emprise (Équation 4.3) associée à $M_w 4,7$ (111×111 km).

L'augmentation de la magnitude entraîne une élévation de la PGA (de 115 à 300 cm/s^2) mais également une accentuation de l'impact de la géométrie du plan de rupture sur la répartition spatiale de la PGA.

Pour cela, le plan de rupture associé à ce séisme est construit Figure 4.2. Le parti pris dans cette thèse est de construire des plans de rupture rectangulaires. Pour autant, d'autres géométries peuvent être prises en compte si nécessaire (faille listrique, ellipsoïde, etc.). Tout d'abord, la longueur L du plan de rupture est estimée en fonction de la magnitude M_w du séisme :

$$L = 10^{(M_w - l_1)/l_2}, \quad (4.1)$$

d'après Leonard [2014]. Les coefficients l_1 et l_2 de l'Équation 4.1 varient suivant le régime de faille. Les valeurs de ces coefficients applicables aux régions stables ainsi que

leurs intervalles de validité sont résumés dans le Tableau 4.1. La Figure 4.3 illustre la distribution de la longueur L du plan de rupture en fonction de la magnitude et du régime de faille.

Régime	l_1	l_2	Intervalles de validité	
			L (km)	M_w
R & N	4,32	1,667	>2,5	>4,9
S	4,25	1,667	1,6 - 70	4,6 - 7,3

Table 4.1: COEFFICIENTS ET INTERVALLES DE VALIDITÉ DE L'ÉQUATION 4.1 [Leonard, 2014].

Valeurs applicables aux régions continentales stables.

Régimes : R=inverse / N=normal / S=décrochant.

Ensuite, un plan de rupture rectangulaire est créé dans l'espace grâce à L , le ratio de forme (longueur sur largeur), l'azimut et le pendage (Figure 4.2b). L est calculé à partir de M_w , le ratio de forme est choisi par l'opérateur tandis que tous les autres paramètres sont fournis par le générateur de séismes.

Dans cette thèse, l'hypocentre (x,y,z) est placé au centre de l'arrête sommitale du plan de rupture (Figure 4.2).

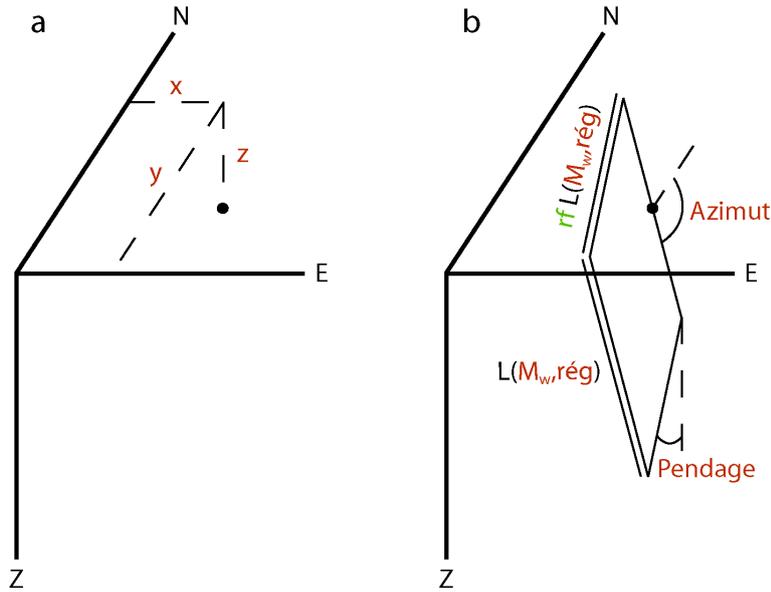


Figure 4.2: CRÉATION DU PLAN DE RUPTURE.

a : Localisation de l'hypocentre (x,y,z) du séisme.

b : Création du plan de rupture à l'aide du ratio de forme rf , de la magnitude M_w et de l'azimut, du pendage et du régime $rég$ de faille.

Les paramètres signalés en rouge sont renseignés par le générateur de séismes.

Le paramètre signalé en vert (ratio de forme rf) est à choisir lors de cette étape.

D'un côté, la Figure 4.1a illustre une distribution spatiale pratiquement circulaire de la PGA produite par une séisme $M_w 4,7$ et ce malgré la prise en compte de la géométrie du plan de rupture. L'impact de cette géométrie semble négligeable dans ce cas. Ainsi,

la géométrie du plan de rupture peut être négligée pour les faibles séismes ($M_w < 5$). D'un autre côté, d'après le Tableau 4.1, tous les séismes à régime normal ou inverse de magnitude inférieure à 5 et tous les séismes à régime décrochant de magnitude inférieure à 4,6 ne peuvent pas être associés à un plan de rupture à partir de l'Équation 4.1. Le constat fait Figure 4.1a nous permet de considérer ces cas particuliers comme de simples points sources et donc de ne pas leur associer de plan de rupture. Pour ces cas particuliers, aucun plan de rupture n'est construit (Figure 4.2a) puisqu'on considère alors la géométrie de la rupture assimilable à un point source.

4.2.2 Des plans de rupture à l'aléa sismique

L'aléa $PGA_g(M_w, D)$ que produit sur un site un séisme de magnitude M_w éloigné d'une distance D est calculé à l'aide d'une GMPE g comme suit :

$$\forall M_w \in [M_{min}, M_{max}], \forall D \in [D_{min}, D_{max}],$$

$$PGA_g(M_w, D) = \begin{cases} 10^{\log_{10}(PGA(M_w, D)) + \epsilon\sigma} \\ e^{\ln(PGA(M_w, D)) + \epsilon\sigma} \end{cases}, \quad (4.2)$$

où $PGA(M_w, D)$ est la PGA médiane calculée par la GMPE g et $\epsilon\sigma$ représente l'incertitude aléatoire à distribution normale centrée en zéro et d'écart-type σ . La variabilité σ des GMPE étant très forte, la communauté scientifique a coutume de tronquer la distribution normale à l'aide d'un ϵ_{max} afin de ne pas calculer de rares PGA extrêmes aberrantes. Cette valeur a un impact fort sur le calcul des courbes probabilité-aléa finales [par ex., Abrahamson, 2000, Bommer et al., 2004]. Dans cette thèse, nous n'évaluons pas l'impact de ϵ_{max} sur l'aléa final. Nous choisissons un ϵ_{max} égal à 3 comme dans les récentes études faites en Europe [Woessner et al., 2015] ou encore en Allemagne [Grünthal et al., 2018]. Les distances minimale D_{min} et maximale D_{max} ainsi que les magnitudes minimale M_{min} et maximale M_{max} sont choisies par l'utilisateur suivant le domaine d'étude envisagé. La Figure 4.3 montre le domaine d'étude utilisé dans cette thèse : la magnitude et la distance s'étendent respectivement de 4 à 7,5 et de 0 à 200 km.

En ce qui concerne les distances, la liste ci-dessous énumère les quatre types les plus couramment utilisés suivant si la source est considérée comme :

- un point :
 - hypocentrale : distance entre le point source en profondeur et le site.
 - épiscopentrale : distance entre le projeté du point source en surface et le site.
- un plan de rupture :
 - rupture : distance la plus proche entre le plan de rupture et le site.
 - Joyner-Boore (JB) : distance la plus proche entre le projeté du plan de rupture en surface et le site.

Chaque modèle de prédiction des mouvements du sol utilise un ou plusieurs type(s) de distance listé(s) ci-dessus. Si la distance de rupture ou JB est utilisée, la géométrie du

plan de rupture est automatiquement prise en compte (par ex., Figure 4.1b). Si en revanche, une distance ponctuelle est utilisée (hypocentrale ou épacentrale), la géométrie du plan de rupture ne sera pas prise en compte.

Dans cette thèse nous utilisons la distance JB pour des raisons expliquées Section 5.2.3. Nous distinguons deux cas de figure dont les spécificités sont résumées dans le Tableau 4.2.

Nom	Conditions		Représentation	Distance utilisée
	Régime	M_w		
Cas 1	R & N	<5	Point	épacentrale
	S	$<4,6$		
Cas 2	R & N	≥ 5	Plan de rupture	Joyner-Boore
	S	$\geq 4,6$		

Table 4.2: SPÉCIFICITÉS DES DEUX SOURCES POSSIBLES DANS CETTE THÈSE.

Nom, conditions d'utilisation [Leonard, 2014], représentation conceptuelle et distance associée aux deux types de source utilisés dans cette thèse.

Régimes : R=inverse / N=normal / S=décrochant.

Le premier concerne les séismes à magnitudes trop faibles pour pouvoir calculer la longueur L de leur plan de rupture associé (Équation 4.1 et Tableau 4.1). Pour ces cas particuliers, aucun plan de rupture n'est créé. Les séismes sont donc considérés comme des points à l'image de la Figure 4.2a. Dans de tels cas, la distance D utilisée dans l'Équation 4.2 est une distance épacentrale (Cas 1 Tableau 4.2).

Le deuxième cas concerne tous les autres séismes, i.e., tous ceux pouvant être associés à un plan de rupture. Ces séismes sont quant à eux considérés comme des plans de rupture à l'image de la Figure 4.2b. Dans de tels cas, la distance D utilisée dans l'Équation 4.2 est une distance JB (Cas 2 Tableau 4.2).

Une fois chaque source définie, l'aléa est calculé à partir de l'Équation 4.2. En plus de prendre un temps de calcul considérable, les PGA calculées à de trop grandes distances tendent vers zero et donc ne participent pas au risque sismique. De plus, la plupart des modèles de prédiction des mouvements du sol ne peuvent pas calculer l'aléa au-delà d'une certaine distance (généralement de 100 à 400 km).

Dans cette thèse, nous considérons comme négligeable vis-à-vis du risque sismique toute intensité macrosismique inférieure ou égale à IV sur l'échelle de Mercalli modifiée (*Modified Mercalli Intensity*, MMI). Cette intensité correspond à un ressenti léger et à aucun dommage sur le bâti. D'après la loi de conversion MMI - PGA de Worden et al. [2010], une intensité MMI de IV correspond à une PGA comprise entre 15 et 40 cm/s^2 . Nous considérons comme négligeable vis-à-vis du risque sismique toute PGA au rocher inférieure à 20 cm/s^2 . Même avec une multiplication de la PGA par deux due aux effets de site, cette dernière reste inférieure à 40 cm/s^2 et donc liée à une intensité égale à IV. Afin de limiter l'étendue spatiale autour d'un séisme, la notion d'emprise est définie. Cette emprise correspond à l'étendue géographique que le séisme impacte. Au-delà de cette étendue, l'aléa produit par le séisme est considéré comme négligeable i.e., inférieur

à 20 cm/s^2 , et n'est donc pas calculé. Une distance d'emprise spatiale $D_{emprise}$ (km) est définie en fonction de la magnitude M_w à partir du 99^{ème} centile des PGA calculées par les GMPE utilisées dans cette thèse :

$$D_{emprise} = \begin{cases} d_1 10^{d_2 M_w} & \text{si } M_w < 6,3 \\ 200 & \text{sinon} \end{cases}, \quad (4.3)$$

avec les coefficients $d_1 = 1,095$ et $d_2 = 0,363$ pour une distance Joyner-Boore. Cette distribution de la distance d'emprise spatiale est limitée à la distance maximale de l'étude, à savoir 200 km. La prise en compte du 99^{ème} centile des PGA pour calculer $D_{emprise}$ est due au choix de ϵ_{max} égal à 3. Ainsi, nous sommes assurés d'être exhaustif pour les PGA supérieures ou égales à 20 cm/s^2 . La Figure 4.3 montre la distribution de $D_{emprise}$ en fonction de la magnitude.

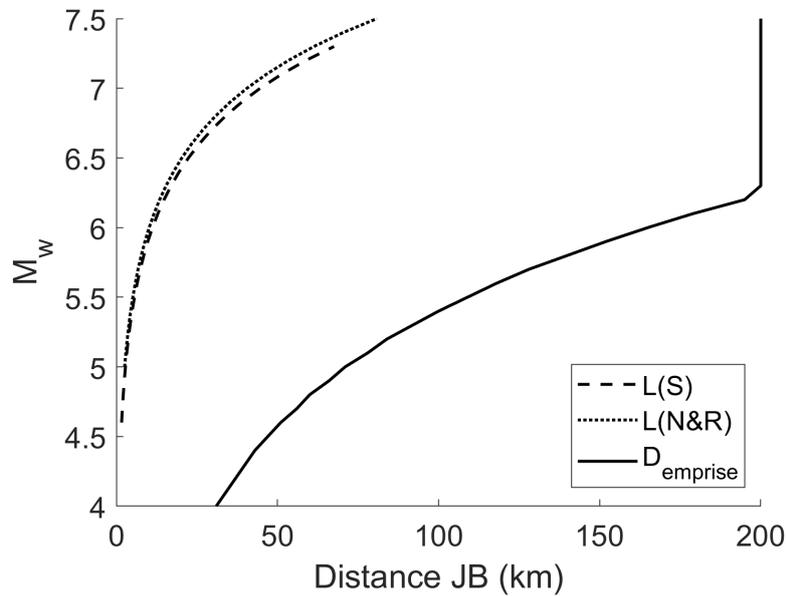


Figure 4.3: DOMAINE DISTANCE - MAGNITUDE UTILISÉ DANS CETTE THÈSE.

Les courbes en tirets et en pointillés correspondent aux longueurs de rupture calculées via l'Équation 4.1 respectivement pour un régime décrochant ($L(S)$) et un régime normal ou inverse ($L(N\&R)$) (Tableau 4.1).

La courbe en trait plein représente la distance d'emprise $D_{emprise}$ calculée à l'aide de l'Équation 4.3. L'aléa produit par les couples distance - magnitude se trouvant à droite de cette courbe est considéré comme négligeable ($< 20 \text{ cm/s}^2$), notamment dans une visée opérationnelle dédiée à l'estimation du risque, et n'est donc pas calculé.

L'emprise spatiale rectangulaire est obtenue en appliquant cette $D_{emprise}$ autour du point source (Cas 1 Tableau 4.2) ou du plan de rupture (Cas 2 Tableau 4.2). La Figure 4.1a illustre l'étendue spatiale obtenue pour un séisme de magnitude 4,7 : $111 \times 111 \text{ km}$.

4.3 Pondération des GMPE

4.3.1 Méthodologie

Pour calculer la PGA, les modèles de prédiction des mouvements du sol (*Ground Motion Prediction Equations*, GMPE) ont besoin d'au minimum la distance séisme-site D et la magnitude de moment M_w du séisme (Équation 1.11). Ainsi, il est possible de représenter la PGA médiane calculée au rocher par une GMPE en fonction de la magnitude et de la distance (Figure 4.4a).

En Figure 4.4b, la distribution de PGA produite par un séisme de magnitude 5,6 est calculée au rocher en fonction de la distance JB par les GMPE d'Améri [2014] (Am14) et de Drouet et Cotton [2015] (DC15, Équation 1.11). Ces distributions de PGA sont comparées avec des données d'accélération enregistrées au rocher et collectées dans la base de données européenne RESORCE [Akkar et al., 2014b]. Deux constats peuvent être faits à la vue de la Figure 4.4b. Le premier est que les accélérations enregistrées possèdent une variabilité que les seuls paramètres distance et magnitude ne peuvent réduire. Cette variabilité se ressent dans les larges écart-types des GMPE. La deuxième observation est que les PGA médianes tout comme les écart-types sont sensiblement différents d'une GMPE à l'autre. Ce constat est la conséquence de données utilisées différentes lors de l'élaboration de ces modèles mais également d'hypothèses et de choix différents. Ainsi, malgré un fort écart-type, l'utilisation d'un seul modèle de prédiction des mouvements du sol ne semble pas suffisant pour capter la variabilité observée sur les accélérations enregistrées (Figure 4.4b).

L'utilisation de plusieurs GMPE applicables au territoire étudié est donc recommandée. Le choix de ces GMPE constitue une incertitude épistémique. Dans la méthode PSHA classique (type Cornell, Section 1.3.2), cette incertitude est traitée au sein d'un arbre logique. Toutes les GMPE considérées sont placées sur une branche de l'arbre à un même niveau de ramification. Chacune de ces branches est associée à un poids témoignant de la confiance accordée à la GMPE en question. Cette approche fait que pour chaque chemin de l'arbre, une unique GMPE est utilisée pour calculer l'aléa. La variabilité apportée par le choix des modèles de prédiction des mouvements du sol est analysée lors du rassemblement de toutes les courbes probabilité-aléa finales pondérées.

Dans cette thèse, comme annoncé Chapitre 1, nous n'utilisons pas d'arbre logique. Ainsi, toutes les incertitudes (aléatoires et épistémiques) sont confondues. En ce qui concerne le calcul de l'aléa sismique, toutes les GMPE sont utilisées en même temps au sein d'une combinaison linéaire dont les coefficients sont les poids de chacune d'entre elles. En posant N_{GMPE} le nombre de GMPE que nous considérons, le calcul de l'aléa devient donc :

$$PGA(M_w, D) = \sum_{g=1}^{N_{GMPE}} a_g(M_w, D) \times PGA_g(M_w, D), \quad (4.4)$$

avec $a_g(M_w, D)$ le poids de la GMPE g et $PGA_g(M_w, D)$ la PGA calculée par la GMPE g via l'Équation 4.2. Contrairement à ce qui se fait classiquement, cette thèse propose de calculer des poids non pas fixes mais dépendant de la magnitude M_w du séisme et de la distance D séisme-site.

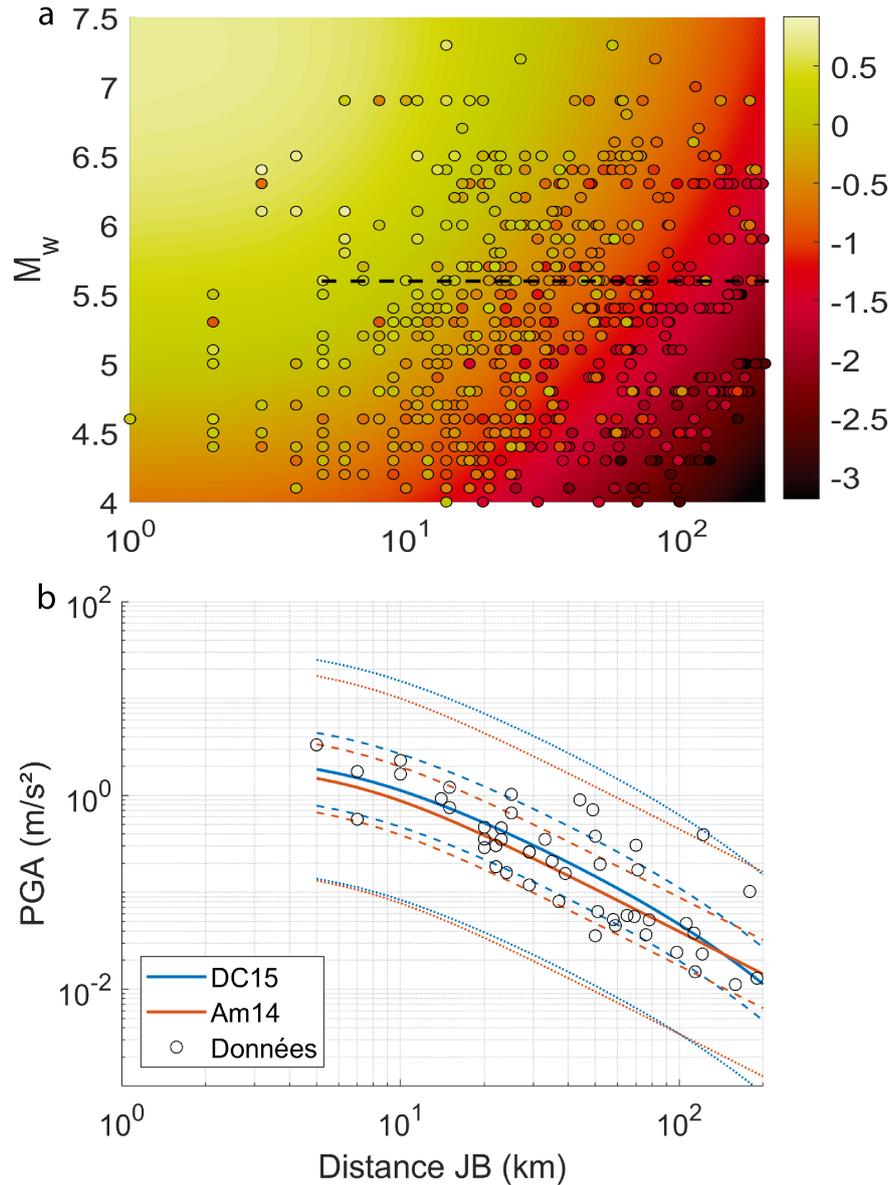


Figure 4.4: DISTRIBUTION DE LA PGA EN FONCTION DE LA MAGNITUDE ET DE LA DISTANCE.

Les cercles représentent des données réelles d'accélération au rocher relevées par les réseaux de sismomètres et contenues dans la base de données RESORCE [Akkar et al., 2014b], présentée Section 5.2.4.

a : Distribution de la PGA sur l'ensemble du domaine défini par $D_{min} = 0$, $D_{max} = 200$ km, $M_{min} = 4$ et $M_{max} = 7,5$. La distribution en arrière plan représente la PGA médiane calculée au rocher à l'aide de la GMPE de Drouet et Cotton [2015] (Équation 1.11). Les tirets localisent la sous-Figure b dans le domaine distance-magnitude. La valeur représentée dans ce graphique est le \log_{10} de la PGA exprimée en m/s^2 .

b : Distribution médiane (trait plein) de la PGA au rocher en fonction de la distance JB, pour $M_w 5,6$. Les deux lois utilisées dans cet exemple sont celles d'Ameri [2014] (Am14) et de Drouet et Cotton [2015] (DC15) avec un régime décrochant. Les tirets et les pointillés représentent respectivement $PGA \pm \sigma$ et $PGA \pm 3\sigma$, avec σ l'incertitude aléatoire des GMPE.

4.3.2 Calcul des poids

Pour calculer les poids de chaque GMPE, des données d'accélération sont utilisées (Figure 4.4). Les PGA provenant de ces données sont considérées comme les PGA de référence et sont notées PGA_{ref} (Équation 4.5). Nous notons PGA_g ($g = 1, \dots, N_{GMPE}$) la PGA médiane calculée par la $g^{ème}$ GMPE via l'Équation 4.2 ($\epsilon = 0$). De plus, les poids de chacune des GMPE dans la combinaison linéaire des PGA_g , calculés pour chaque D et M_w , sont notés $a_g(M_w, D)$. En s'appuyant sur les PGA de référence, les poids $a_g(M_w, D)$ doivent respecter la condition suivante :

$$\sum_{g=1}^{N_{GMPE}} a_g(M_w, D) \times PGA_g(M_w, D) = PGA_{ref}(M_w, D). \quad (4.5)$$

Pour calculer les coefficients a_g , une régression des moindres carrés contrainte est appliquée en chaque point de données comme suit :

$$\min_{a_g} \frac{1}{2} \|a_g PGA_g - PGA_{ref}\|_2^2 \text{ tel que } \begin{cases} \sum_{g=1}^{N_{GMPE}} a_g = 1 \\ 0 \leq a_g \leq 1 \end{cases}, \quad (4.6)$$

où $\|\cdot\|_2$ est la norme de degré 2. Ce calcul est réalisé à l'aide de la fonction *lsqlin* de MATLAB¹.

Notons que chaque GMPE possède une gamme de validité en ce qui concerne la magnitude et la distance. Lorsqu'une GMPE ne peut pas calculer de PGA car la magnitude ou la distance demandée n'est pas dans cet intervalle d'application, son poids est fixé à 0 pour les D et M_w en question.

À partir des poids calculés en chaque point de données, une interpolation du plus proche voisin est réalisée sur l'ensemble des valeurs de magnitude et de distance afin d'obtenir un modèle de pondération sur l'ensemble du domaine distance-magnitude. Cela permet d'estimer des poids également dans les domaines distance-magnitude non couverts par les données. Les pas d'échantillonnage de la magnitude (par ex., tous les 0,1) et de la distance (par ex., tous les 5 km) sont à choisir par l'utilisateur pour appliquer l'interpolation. Le poids de chaque GMPE peut ainsi être obtenu à partir de la magnitude du séisme et la distance séisme-site. Enfin, la PGA finale résulte de la combinaison linéaire des $a_g(M_w, D)$ et $PGA_g(M_w, D)$ (Équation 4.4).

Lors de la pondération, seule la PGA médiane de chaque GMPE est considérée. La variabilité propre à chaque modèle de prédiction n'est donc pas prise en compte dans le calcul des poids. Cette variabilité est en revanche bien prise en compte lors du calcul de l'aléa final via l'Équation 4.4. Elle est exprimée par le terme ϵ , identique pour chaque GMPE, et par le terme σ , propre à chaque GMPE (Équation 4.2).

¹<https://fr.mathworks.com/help/optim/ug/lsqlin.html>

4.4 Discussions et Conclusion

Dans ce Chapitre, j'ai présenté comment l'aléa sismique que produit un séisme, défini par plusieurs paramètres, est estimé à l'aide d'un ensemble de modèles de prédiction des mouvements du sol (GMPE) pondérés. L'accélération maximale du sol (PGA) estimée au rocher est choisie pour représenter l'aléa sismique.

La pondération des modèles de prédiction des mouvements du sol est réalisée à partir d'une simple analyse résiduelle entre les PGA observées et les PGA médianes calculées par ces modèles (Équation 4.6). Les poids sont déterminés en fonction des deux paramètres influençant le plus le calcul de la PGA : la magnitude du séisme et sa distance au site d'étude. Une fois les poids de chaque GMPE calculés en chaque point de données, une interpolation du plus proche voisin est appliquée afin d'obtenir un modèle de pondération dépendant de la magnitude et de la distance.

Le calcul de la PGA finale est réalisé à l'aide d'une combinaison linéaire entre les poids de chaque GMPE, obtenus via le modèle de pondération, et la PGA calculée par chaque GMPE. Ainsi, les poids associés aux modèles de prédiction ne servent pas à définir le poids de chaque branche d'un arbre logique comme il est coutume de le faire dans les études PSHA classique (type Cornell [Cornell, 1968]). Au contraire, les incertitudes aléatoires (Équation 4.2) et épistémiques (Équation 4.4) sont toutes deux utilisées pour calculer la PGA finale.

La stratégie de pondération des GMPE proposée dans ce Chapitre est une première étape. Plusieurs pistes d'amélioration peuvent être envisagées. Tout d'abord, il existe d'autres analyses comparatives plus complexes que l'utilisation d'un simple résidu. Par exemple, Scherbaum et al. [2004, 2009] normalisent le résidu par la variabilité (σ) des GMPE. En supposant que la distribution des résidus normalisés suit une loi normale centrée réduite, ces auteurs proposent de classer les GMPE à partir des valeurs moyenne et médiane de cette distribution ainsi que de son écart-type. Une approche Bayésienne [par ex., Bertin et al., 2019] peut également être envisagée.

Ensuite, l'interpolation appliquée dans ce Chapitre n'est pas optimale. La méthode du plus proche voisin n'est peut être pas la plus adaptée et les incertitudes liées à l'interpolation ne sont pas prises en compte.

Enfin, la grille régulière (pas constants de distance et de magnitude) utilisée pour appliquer l'interpolation peut être modifiée afin de considérer un maillage adaptatif. La grille pourrait par exemple être plus grossière dans les domaines distance - magnitude où il y a peu de données et inversement plus fine dans les domaines où les données sont plus abondantes.

Finalement, le calcul de l'aléa sismique est appliqué au catalogue synthétique produit en Chapitre 3. Pour se faire, la représentation spatiale de chaque séisme synthétique est tout d'abord choisie en fonction de sa magnitude. Les séismes synthétiques de faible magnitude sont considérés comme des points sources tandis que les autres sont transformés en des plans de rupture à l'aide des paramètres qui les définissent (magnitude, localisation 3D, azimut, pendage et régime de faille). Afin d'éviter du temps et de la mémoire de calcul trop conséquents, une distance d'emprise spatiale (Équation 4.3) est appliquée autour de chacun des plans de rupture ou des points sources. Cette distance

est définie de sorte que la PGA calculée dans l'emprise est obligatoirement supérieure ou égale à 20 cm/s^2 . Cette valeur est considérée comme la borne minimale en dessous de laquelle l'aléa sismique est négligeable pour une étude du risque sismique. Ainsi, les PGA calculées dans cette thèse sont considérées comme exhaustives à partir de 20 cm/s^2 .

Pour les besoins de l'estimation de l'aléa sismique, seule la PGA au rocher est calculée puisque c'est cette métrique qui est classiquement utilisée. Pour autant, lors de l'intégration du modèle d'aléa dans le modèle risque, les effets de site doivent être pris en compte. Cela peut être fait en multipliant la PGA au rocher par des coefficients d'effets de site topographique et lithologique.

Chapitre 5

Application : France métropolitaine

Sommaire

5.1	Introduction	78
5.2	Données françaises et paramétrisation	79
5.2.1	Catalogues sismiques	79
5.2.2	Caractéristiques spatiales	82
5.2.3	Modèles de prédiction des mouvements du sol	89
5.2.4	Catalogue d'accélération	90
5.3	Proportions et fréquences de séismes principaux en France	92
5.3.1	Distribution proportion - magnitude de séismes principaux en France	93
5.3.2	Distribution fréquence - magnitude de séismes principaux en France	97
5.4	Génération de catalogues synthétiques français	104
5.4.1	Comparaison des distributions de proportions et de fréquences	104
5.4.2	Moment sismique cumulé au cours du temps	107
5.4.3	Répartition spatiale de la sismicité synthétique	109
5.5	Calcul de l'aléa sismique français	111
5.5.1	Pondération	111
5.5.2	Calcul de la PGA	114
5.6	Estimation de l'aléa sismique en France	118
5.6.1	Courbes probabilité annuelle-aléa	118
5.6.2	Impact des différents modèles utilisés	124
5.6.3	Contributions distance-magnitude	129
5.7	Conclusion	132

Résumé Dans ce Chapitre, j’applique à la France métropolitaine les méthodes proposées dans les Chapitres 2, 3 et 4. Les données françaises sont décrites et traitées afin de paramétrer au mieux le générateur de séismes synthétiques. Malgré une paramétrisation peu conséquente, ce générateur permet de produire une sismicité cohérente avec certains modèles connus. Néanmoins, la sismicité simulée peine à reproduire certains paramètres donnés en entrée. L’aléa produit par l’ensemble des séismes (principaux et répliques) simulés est finalement estimé pour six villes à l’aide d’un ensemble de cinq modèles de prédiction des mouvements du sol pondérés. Ces aléas sont, d’une manière générale, plus faibles que l’aléa réglementaire et peuvent varier suivant les choix fait lors de la paramétrisation du générateur. Cette vision plus adoucie de l’aléa français s’intègre dans une tendance observée depuis 20 ans dans la littérature.

5.1 Introduction

La France métropolitaine est caractérisée par une sismicité faible à modérée. Comparée aux régions actives comme le Japon ou la Californie, les données sismiques sont peu nombreuses. Une des approches permettant d’analyser l’aléa sismique en dépit de ce peu de données est de produire des séismes synthétiques regroupés dans un catalogue. L’aléa produit par ces séismes synthétiques est analysé afin de calculer une courbe de probabilité annuelle de dépasser des seuils d’aléa. Cette méthode est l’approche PSHA (*Probabilistic Seismic Hazard Assessment*) dite des catalogues synthétiques (ou Monte Carlo).

La première étape de cette approche est de traiter et d’analyser les données sismiques françaises. Une présentation de ces données tels que les catalogues de séismes et d’accélération, est réalisée Section 5.2. Dans cette Section, les données sont également traitées et des hypothèses sont formulées afin de paramétrer le générateur de séismes. Ensuite, une caractérisation de la sismicité actuelle est réalisée Section 5.3 à l’aide des données traitées via la méthode présentée Chapitre 2.

La deuxième étape consiste à produire de la sismicité synthétique à l’aide de modèles décrits Section 1.4. Une méthodologie de générateur de séismes adaptée au contexte de sismicité faible à modérée a été présentée Chapitre 3. Dans la Section 5.4, cette méthodologie est appliquée au territoire français.

La troisième étape est le calcul de l’aléa sismique que produisent les séismes synthétiques. Cinq modèles de prédiction des mouvements du sol élaborés en France, en Europe et dans d’autres régions du monde sont pré-sélectionnés Section 5.2. La stratégie de pondération, présentée Chapitre 4, est appliquée sur ces modèles pré-sélectionnés. Les résultats de cette pondération ainsi que la distribution et la variabilité des PGA calculées sont présentés Section 5.5.

Enfin, la dernière étape concerne le calcul de la probabilité annuelle de dépassement de seuils d’aléa. Les résultats obtenus pour six villes françaises sont présentés, analysés et discutés Section 5.6.

5.2 Données françaises et paramétrisation

5.2.1 Catalogues sismiques

5.2.1.1 Catalogue instrumental

En France métropolitaine, la première campagne de couverture du territoire par des sismographes date de 1962 avec l'installation de six appareils par le Laboratoire de Détection et de Géophysique (LDG) du Commissariat à l'Énergie Atomique (CEA). Au début des années 1980, le Réseau National de Surveillance Sismique (RéNaSS¹) est créé dans le but d'offrir une couverture plus importante de sismographes sur l'ensemble du territoire français. Ainsi, dès les années 1980, 75 appareils sont installés en plus du réseau initial. Le réseau sismologique et géodésique français (Résif²) a été lancé en 2009 pour fédérer, moderniser et développer les moyens d'observation géophysique de la Terre interne. La fédération concerne le regroupement des différents réseaux, qu'ils soient français (par ex., CEA-LDG et RéNaSS) ou étrangers (par ex., BGS en Grande-Bretagne, ORB en Belgique, SED en Suisse et IGN en Espagne). La modernisation fait référence aux travaux d'amélioration du matériel existant comme la conversion progressive vers un réseau large-bande. Enfin le développement sous-entend la densification du réseau de sismographes (Figure 5.1). En 2020, le réseau compte 167 sismographes en France métropolitaine (Figure 5.1).

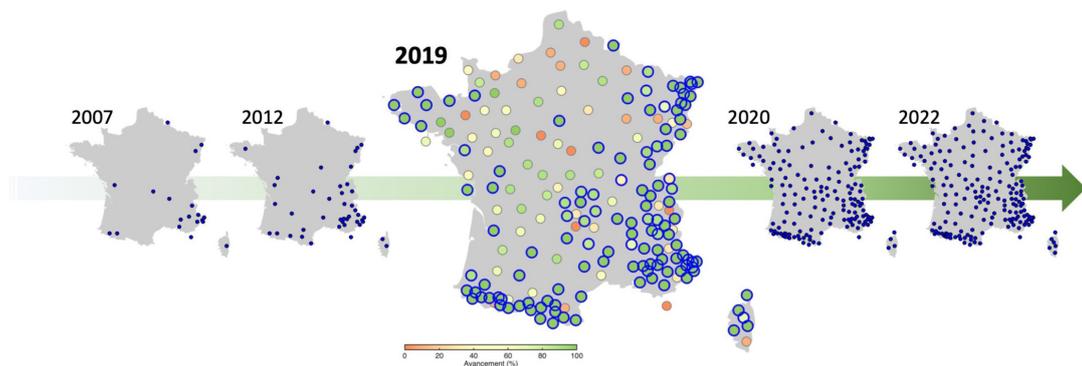


Figure 5.1: ÉVOLUTION SPATIO-TEMPORELLE DU RÉSEAU DE SISMOGRAPHES LARGE BANDE (Résif).

La couleur des cercles indique le degré d'avancement de déploiement de chaque station. Les cercles entourés en bleu correspondent aux stations fournissant des données sur le portail Résif depuis mai 2019.

Cara et al. [2015] créent le catalogue de Sismicité Instrumentale de l'Hexagone (SI-Hex). Ce catalogue vise à homogénéiser l'ensemble des catalogues fournis par les différents réseaux nationaux, locaux (par ex., Sismalp dans les Alpes, Geoazur en Côte d'Azur et OMP dans les Pyrénées) ou encore internationaux (ISC, EMSC). L'homogénéisation concerne la magnitude (toutes converties en magnitude de moment M_w) et la localisation 3D des séismes. La couverture de l'étude comprend les terres émergées situées en France métropolitaine et jusqu'à 20 km de la frontière ainsi que les terres

¹<https://renass.unistra.fr/fr/zones/>

²<https://www.resif.fr/>

immergées comprises dans la zone économique exclusive. Le catalogue SIHex couvre la période allant de 1965 à 2009 et comporte plus de 38 000 séismes. Ce travail d'homogénéisation est poursuivi par le RéNaSS depuis 2010.

Le principal défaut du catalogue instrumental est sa faible profondeur temporelle. En effet, les 56 années que propose le catalogue SIHex prolongé par le RéNaSS (1965-2020) ne représentent que la partie la plus récente de la sismicité française. Ceci explique pourquoi la magnitude la plus élevée du catalogue n'est que de $M_w 5,2$; soit seulement 0,3 de plus que le séisme du Teil de 2019 qui est pourtant loin d'être le plus fort séisme auquel le territoire français peut être exposé.

5.2.1.2 Catalogue historique

Pour pallier la courte profondeur temporelle du catalogue instrumental, un catalogue historique peut être utilisé. À défaut d'être moins exhaustif pour les faibles magnitudes (non ressenties par la population), ce catalogue a le mérite de couvrir une période plus longue et ainsi de mettre en évidence des séismes de plus fortes magnitudes.

En France métropolitaine, les travaux de Traversa et al. [2017] et Baumont et al. [2018] ont permis d'améliorer le catalogue historique SisFrance (BRGM-EDF-IRSN) notamment en ce qui concerne la conversion en magnitude de moment et la relocalisation des profondeurs des séismes. Le catalogue résultant de ces travaux s'appelle FCAT [Manchuel et al., 2017] et s'étend de 463 à 1964. Il est composé de 4 250 séismes dont le plus fort est celui de Ligurie en 1887 avec une magnitude de moment estimée à $6,7 \pm 0,6$.

5.2.1.3 Choix du catalogue sismique

La Figure 5.2 illustre la complémentarité qu'il existe entre le catalogue historique et le catalogue instrumental. Le catalogue historique est composé de moins de séismes que le catalogue instrumental : sur la période 463-2018, les séismes historiques (463-1964) ne représentent que 6,9% du nombre total d'événements. En revanche, la partie historique du catalogue concentre 99,2% du moment sismique libéré sur l'ensemble de la période 463-2018. Ainsi, moins de 7% des séismes ont produit plus de 99% du moment sismique.

Le clivage observé entre les parties historique et instrumental (Figure 5.2) est probablement accentué par une sous-estimation de la magnitude des séismes instrumentaux, une surestimation de la magnitude des séismes historiques ou bien les deux à la fois. Les méthodes d'estimation de la magnitude, paramètre-clé de l'analyse de l'aléa sismique, sont différentes entre les deux catalogues et peuvent expliquer cette différence de comportement. La fusion des catalogues instrumental et historique appelle à une uniformisation des données qui semble actuellement difficile à réaliser à cause des questions soulevées ci-dessus. Ces questions ne trouvant pour l'heure pas de réponse, je préfère ne pas fusionner ces deux catalogues au comportement différent.

La caractérisation temporelle de la sismicité française est basée sur la méthode des temps de latence (Chapitre 2). Cette dernière nécessite une longue profondeur temporelle afin de capturer le maximum d'informations concernant toutes les magnitudes. Ainsi, pour l'analyse temporelle de la sismicité française, le choix s'est porté sur le catalogue historique FCAT (463-1964) car, malgré ses fortes incertitudes, il atteint des magnitudes élevées grâce à la longue période qu'il couvre (1 500 ans). En ce qui concerne

la définition de la magnitude maximale du territoire français, le catalogue FCAT est également utilisé. Enfin, de par sa précision plus élevée, le catalogue instrumental SIHex est utilisé pour définir les gammes de profondeurs dans chaque région d'étude.

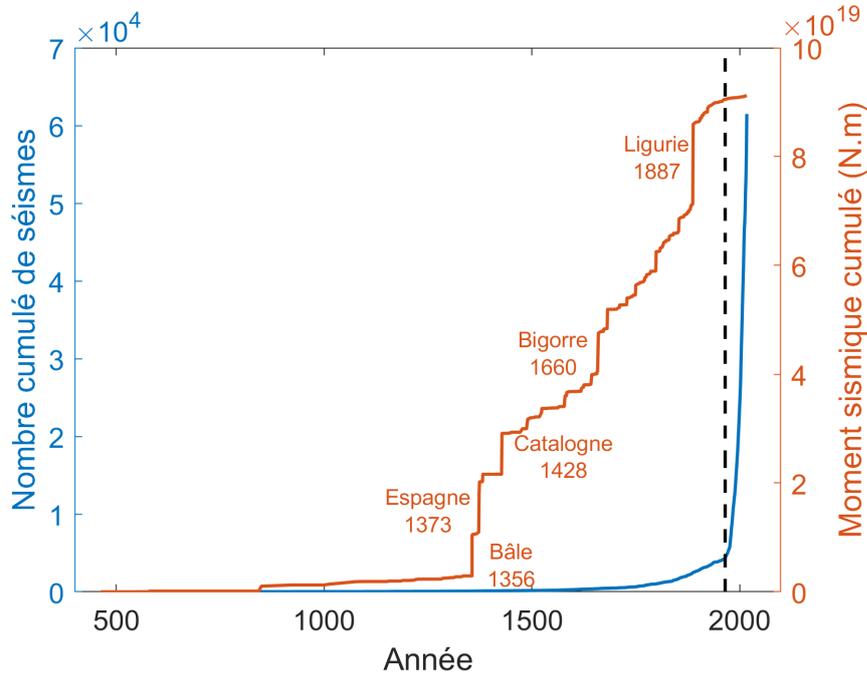


Figure 5.2: COMPLÉMENTARITÉ DES CATALOGUES INSTRUMENTAL ET HISTORIQUE. Nombre cumulé de séismes et moment sismique cumulé (N.m) au cours du temps. Les parties historique (463-1964) et instrumentale (1965-2018) sont séparées par la ligne en tirets. Les cinq séismes ayant une magnitude de moment supérieure ou égale à 6,5 sont indiqués sur la courbe du moment sismique cumulé.

5.2.1.4 Exhaustivité du catalogue sismique

Comme mentionné Sections 1.2.2 et 5.2.1.3, le catalogue historique comporte peu de séismes de faibles magnitudes. Il est donc nécessaire de définir un couple magnitude (M_c) et année (A_c) de coupure afin de rendre le catalogue historique FCAT exhaustif. La capacité du catalogue à répertorier les plus petits séismes s'améliore avec le temps (Section 1.2), ainsi plusieurs couples M_c - A_c sont à définir. Dans cette thèse, nous choisissons les mêmes magnitudes et années de coupure trouvées par Beauval et al. [2020]. Elles ont été définies à l'aide de la visualisation du nombre cumulé de séismes dans le temps en fonction de la magnitude [par ex., Tinti et Mulargia, 1985]. Ces couples M_c - A_c sont visibles Tableau 5.1. Tout comme dans l'étude originelle, l'incertitude autour des périodes d'exhaustivité n'est pas prise en compte.

M_c	4	4,1 - 4,6	4,7 - 5,5	5,6 - 5,8	$\geq 5,9$
A_c	1850	1830	1750	1600	1310

Table 5.1: MAGNITUDES ET ANNÉES DE COUPURE APPLIQUÉES DANS CETTE THÈSE [d'après Beauval et al., 2020].

Dans leur étude, Beauval et al. [2020] utilisent le catalogue historique FCAT ainsi que le catalogue instrumental SIHex pour une application sur une gamme de magnitudes supérieures à 3,2. Dans cette thèse, nous nous intéressons uniquement aux séismes historiques de magnitudes supérieures ou égales à 4. Le catalogue instrumental SIHex (1965-2009) ne comporte que 39 séismes $M_w \geq 4$, soit un taux annuel de 0,87. De leur côté, les données historiques (1850-1964) comportent 288 séismes $M_w \geq 4$, soit un taux annuel deux fois plus élevé (1,8). Nous notons ici une incohérence qui motive la non considération de la faible quantité de séismes $M_w \geq 4$ instrumentaux. Ainsi, les couples $M_c - A_c$ (Tableau 5.1) sont appliqués aux seules données historiques.

De plus, Beauval et al. [2020] obtiennent ces couples $M_c - A_c$ à partir du catalogue FCAT-SIHex résiduel. Ce catalogue résiduel a été obtenu à l'aide des algorithmes de *declustering* de Gardner et Knopoff [1974] et Reasenbergs [1985]. Dans les deux cas, le nombre de séismes dépendants retirés est faible pour les magnitudes supérieures ou égales à 4 (Fig. 7 Beauval et al. [2020]). Tout comme pour la légère différence de catalogue, nous considérons que cette différence n'empêche pas l'application des couples $M_c - A_c$ (Tableau 5.1) au catalogue FCAT non résiduel.

5.2.2 Caractéristiques spatiales

Cette Section a pour vocation de décrire les caractéristiques spatiales de la sismicité française. En d'autres termes, il s'agit de définir un modèle source (Section 1.4.1). Ces caractéristiques sont ensuite utilisées par le générateur afin de produire des catalogues de séismes synthétiques. Pour se faire, une densité spatiale caractérisant la sismicité française est calculée Section 5.2.2.1. Elle servira à tirer la localisation des séismes synthétiques. Ensuite, une régionalisation (Section 5.2.2.2) est nécessaire afin de définir spatialement la magnitude maximale autorisée et les paramètres de faille.

5.2.2.1 Densité spatiale

De par son éloignement des grandes limites de plaques, le territoire français est considéré en grande partie comme stable. De ce fait, des processus géodynamiques complexes [par ex., Mazzotti et al., 2020, pour une liste non exhaustive] jouent un rôle plus important que la tectonique des plaques sur la production de sismicité. Ces processus géodynamiques produisent des déformations intraplaques qui s'accompagnent d'une sismicité. Cependant, la production de séismes nécessite que la croûte soit fragilisée. La sismicité est donc *a priori* produite là où des failles sont présentes. Cela implique que les failles en régions stables peuvent rester inactives sur de très longues périodes [par ex., Liu et al., 2011, Stein et al., 2015], comme en témoigne le récent séisme du Teil (2019). Ce séisme a eu lieu sur la faille de la Rouvière qui date de l'Oligocène (23-34 Ma) et qui n'a pas montré de signe de déplacements depuis au moins plusieurs dizaines de milliers d'années [Ritz et al., 2020]. Une autre conséquence est que la sismicité des régions stables est probablement répartie de manière plus homogène que ce que l'on laisse penser la sismicité actuelle [Calais et al., 2016].

Au vu de ces constats et hypothèses, les failles semblent être des objets géologiques spatialement analogues à la sismicité sur le long terme. Les tracés de failles (Figure 5.3a) sont donc utilisés pour produire une densité spatiale analogue à la sismicité.

Ces tracés de failles proviennent des cartes géologiques numérisées par le BRGM³ (BD Charm). Les constats énoncés ci-dessus ainsi que le manque d'exhaustivité des failles actives et l'absence de définition claire de la période d'activité d'une faille ont conduit à sélectionner l'ensemble des tracés de failles quelles que soient leur importance et activité supposées. Ce choix est discutable et la densité résultante peut être améliorée en incorporant d'autres sources d'information (meilleure exhaustivité des failles actives, etc.).

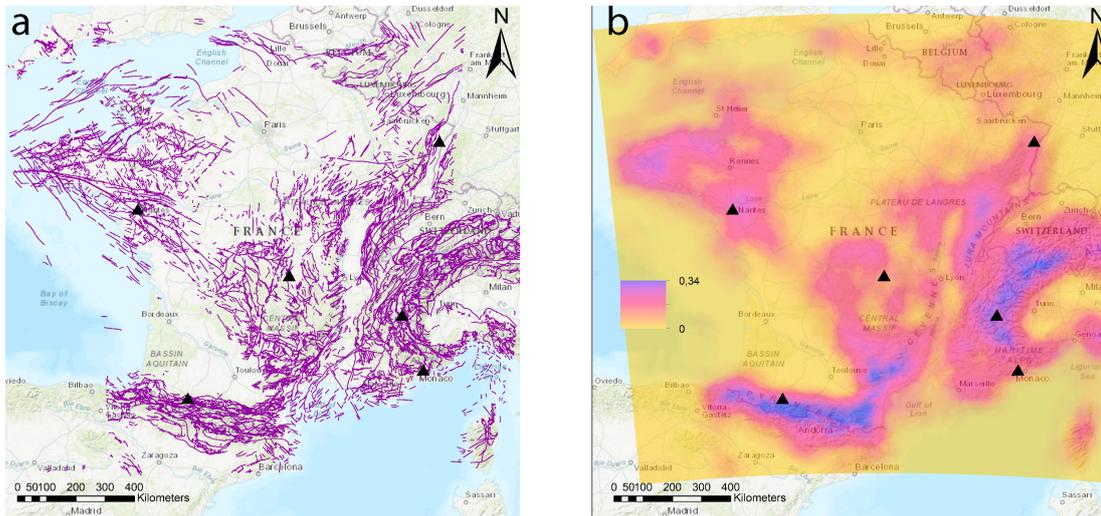


Figure 5.3: DENSITÉ SPATIALE UTILISÉE DANS CETTE THÈSE.

a : Carte des tracés de failles utilisés comme analogue à la sismicité (BD Charm BRGM).
 b : Densité spatiale des tracés de failles calculée en chaque cellule de 5×5 km.
 Les triangles noirs localisent les six villes analysées Section 5.6 : Nantes, Lourdes, Clermont-Ferrand, Briançon, Nice et Strasbourg.

La densité spatiale de tracés de failles est calculée sur une grille cartésienne de résolution 5 km. L'évaluation de la densité se fait pour chaque cellule en divisant la longueur totale des tracés de failles par unité de surface dans un rayon de 25 km autour du centre de la cellule. La valeur obtenue correspond donc à une longueur de tracés de failles par unité de surface. Ce calcul a été fait à l'aide de l'outil "Densité de lignes" d'ArcGIS⁴. La densité spatiale obtenue pour la France métropolitaine après cette opération est visible Figure 5.3b.

La carte répertoriant les tracés de failles (Figure 5.3a) n'est pas exhaustive et les failles cachées par la couverture sédimentaire n'apparaissent pas. L'absence de tracés n'est donc pas synonyme d'absence de faille. Pour pallier cet effet et laisser la possibilité de tirer des séismes n'importe où sur le territoire, nous avons associé aux régions où aucune faille n'a été répertoriée en surface une valeur minimale de densité. Cette valeur minimale est arbitrairement fixée à 1% de la valeur maximale de densité. Ainsi, les régions les moins représentées ont une probabilité 100 fois moindre d'être sélectionnée pour localiser les séismes synthétiques vis-à-vis des régions les plus représentées. Concrètement, d'après la Figure 5.3b, toutes les cellules de 5×5 km ayant une valeur de

³<https://infoterre.brgm.fr/page/telechargement-cartes-geologiques>

⁴<https://desktop.arcgis.com/fr/arcmap/10.3/tools/spatial-analyst-toolbox/how-line-density-works.htm>

densité inférieure au centième de la valeur maximale ($\frac{0,34}{100} = 0,0034$) voient leur densité fixée à cette valeur (0,0034). L'ensemble des cellules concernées par ce réajustement de densité représente environ 15% du territoire et concerne principalement la Méditerranée et les bassins aquitain et parisien (Figure 5.3b). Encore une fois, ce choix est discutable mais nécessaire afin de laisser la possibilité de produire de la sismicité sur l'ensemble du territoire.

5.2.2.2 Régionalisation et paramétrisation

Comme précisé Chapitre 3, la régionalisation sert à limiter la magnitude maximale autorisée lors du tirage des séismes principaux dans l'espace. Cette régionalisation sert également à définir les valeurs des paramètres des séismes synthétiques tels que la profondeur, l'azimut, le pendage et le régime de faille.

Régionalisation À l'image de ce qui a été fait dans le projet européen SHARE, de larges régions peuvent être utilisées pour limiter les magnitudes maximales (Fig. 4a,b Woessner et al. [2015]). La régionalisation tectonique de ce projet européen est utilisée dans cette thèse sous le nom de SHARE (Figure 5.4b). Elle discrétise le territoire français en trois catégories suivant leur tectonique : les régions continentales stables (*Stable Continental Region*, SCR), les régions actives peu profondes (Alpes, Pyrénées et Bassin Rhénan) et la croûte océanique.

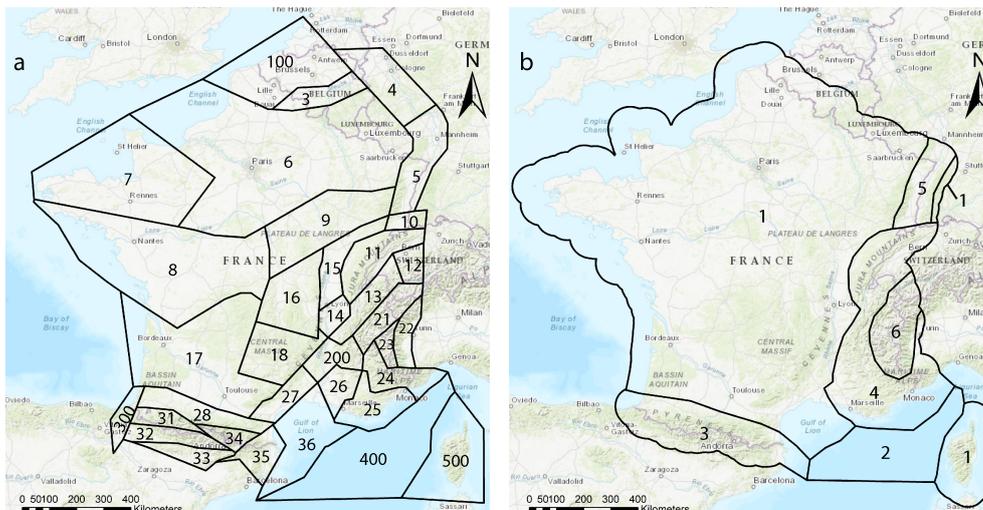


Figure 5.4: RÉGIONALISATION DU TERRITOIRE FRANÇAIS.

a : Régionalisation GEOTER (d'après Martin et al. [2002]) comprenant 35 régions. Les régions 1-2, 19-20, 29-30, 37-38 et 39-40 ont été fusionnées respectivement sous les numéros 100, 200, 300, 400 et 500.

b : Partie française (jusqu'à 50 km de la frontière) de la régionalisation tectonique SHARE comprenant six régions. Les régions 1 correspondent aux régions continentales stables. La région 2 se réfère à la croûte océanique. Enfin, les régions 3 (Pyrénées) et 4 (Alpes) représentent les domaines actifs en compression tandis que les régions 5 (Fossé Rhénan) et 6 (Alpes) correspondent aux domaines actifs en extension.

Le zonage sismotectonique GEOTER [Martin et al., 2002] (Figure 5.4a) a également été choisi comme régionalisation du territoire français. Il est composé initialement 40

zones, appelées régions dans cette thèse. Les données peuvent être peu nombreuses voire inexistantes dans certaines régions GEOTER, empêchant la description des paramètres de faille. Cette observation a conduit à fusionner certaines régions entre elles. C'est le cas des deux régions du Nord (1-2 \rightarrow 100), d'Ardèche-Baronnies (19-20 \rightarrow 200), des Pyrénées occidentales (29-30 \rightarrow 300), de la Méditerranée (37-38 \rightarrow 400) et de la Corse (39-40 \rightarrow 500). Au final la régionalisation GEOTER modifiée comporte 35 régions visibles Figure 5.4a.

Paramétrisation des régions Chaque région est caractérisée par sa magnitude maximale ainsi que des intervalles de profondeur et de paramètres de faille. Tous ces paramètres sont définis à l'aide de différents types de données :

- Le catalogue historique FCAT pour la magnitude maximale.
- Le catalogue instrumental SIHex pour la profondeur.
- La carte des tracés de failles françaises (BD Charm du BRGM) pour l'azimut.
- Les bases de données de mécanismes au foyer européenne EMMA [Vannucci et al., 2009] et mondiale WSM [Heidbach et al., 2016] pour l'azimut, le pendage et le régime (discret) de faille.

La magnitude maximale de chaque région $M_{max\ Région}$ est choisie comme étant la magnitude maximale observée M_{max} dans cette région dans le catalogue FCAT à laquelle est ajoutée son incertitude $ErrMw$. Le plus fort séisme de FCAT possède une magnitude de $6,7 \pm 0,6$ (Ligurie 1887). La magnitude maximale de l'étude est donc fixée à 7,3. Les régions 24 (GEOTER) et 4 (SHARE) où est localisé ce séisme sont donc caractérisées par une $M_{max\ Région}$ égale à 7,3 (Tableaux 5.2 et 5.3).

La valeur minimale des magnitudes maximales des régions est fixée à 5,5 comme cela est par exemple fait dans l'étude française de Drouet et al. [2020]. Autrement dit, pour une région donnée, si les observations dans FCAT ressortent $M_{max} + ErrMw < 5,5$ alors $M_{max\ Région}$ est fixée à 5,5. L'hypothèse qui se cache derrière ce réajustement est qu'un séisme $M_w \leq 5,5$ peut se produire sur l'intégralité du territoire français. L'incertitude sur les magnitudes maximales n'est pas explorée dans cette thèse.

À partir des tracés de failles présentés Figure 5.3a, une densité de probabilité d'orientation de faille est calculée en chaque région avec un pas de 5° . Les densités de probabilité obtenues au sein des régions SHARE sont présentées Figure 5.5. Pour chaque séisme généré, une orientation est tirée au sein de ces distributions en fonction de la région où il a été généré. En choisissant aléatoirement le sens du plongement, l'azimut de ce séisme est obtenu directement à partir de cette orientation (Section 3.2.3). Lorsque le nombre de failles est trop faible au sein d'une région (par ex., la région croûte océanique de SHARE) pour calculer la densité d'orientation de faille, l'azimut est tiré dans un intervalle de valeurs prédéfini pour chaque région. Pour estimer cet intervalle, la distribution de l'azimut dans chaque région est analysée avec un pas de 5° au sein des bases de données de mécanismes au foyer. L'intervalle de valeur est centré sur la valeur d'azimut la plus fréquente. La largeur de cet intervalle (20, 40 ou 60° , Tableau 5.2) dépend de la dispersion de la distribution des valeurs d'azimut.

Région GEOTER	$M_{max\ Région}$	Profondeur		Azimut		Pendage		Régime
		min	max	min	max	min	max	
100	6	0	10	265	285	71	81	S
3	5,9	0	10	85	145	47	57	N
4	6,3	5	15	105	165	58	68	N
5	6,5	5	15	125	185	54	64	N
6	5,6	5	15	115	175	68	78	U
7	5,9	5	15	45	105	63	73	N
8	6,4	5	15	75	135	59	69	S
9	5,9	5	15	-5	15	68	78	S
10	7,1	5	15	110	170	59	69	S
11	6,4	5	15	115	175	61	71	S
12	6,5	5	15	255	315	60	70	S
13	6,6	5	15	265	325	62	72	S
14	5,7	5	15	155	175	65	75	U
15	5,7	10	20	145	185	65	75	U
16	5,9	10	20	145	185	79	89	N
17	5,8	5	15	115	155	55	65	S
18	5,9	5	15	-10	10	69	79	N
200	5,7	5	15	85	145	43	53	R
21	6,6	0	10	20	80	53	63	N
22	6,4	5	15	115	175	46	56	N
23	6,4	0	10	-30	30	51	61	N
24	7,3	5	15	140	200	48	58	S
25	5,8	5	15	165	225	53	63	R
26	6,1	5	15	-5	55	51	61	S
27	5,6	5	15	25	85	68	78	N
28	5,6	5	15	0	60	64	74	S
300	5,5	0	10	80	100	68	78	N
31	7,1	0	10	75	135	60	70	N
32	6,1	0	10	70	130	71	81	N
33	5,5	0	10	-10	50	66	76	R
34	5,9	0	10	135	175	59	69	R
35	7,1	0	10	85	145	65	75	N
36	5,8	5	15	-5	15	75	85	N
400	5,5	10	20	-20	20	54	64	R
500	5,7	5	15	-20	20	51	61	U

Table 5.2: MAGNITUDE MAXIMALE ET PARAMÈTRES DE FAILLE DES RÉGIONS GEOTER.

Magnitude maximale et paramètres de faille définis dans cette thèse pour les régions GEOTER modifiées. Le numéro des régions se réfère à la Figure 5.4a.

Régimes de faille : U = inconnu, N = normal, S = décrochant et R = inverse.

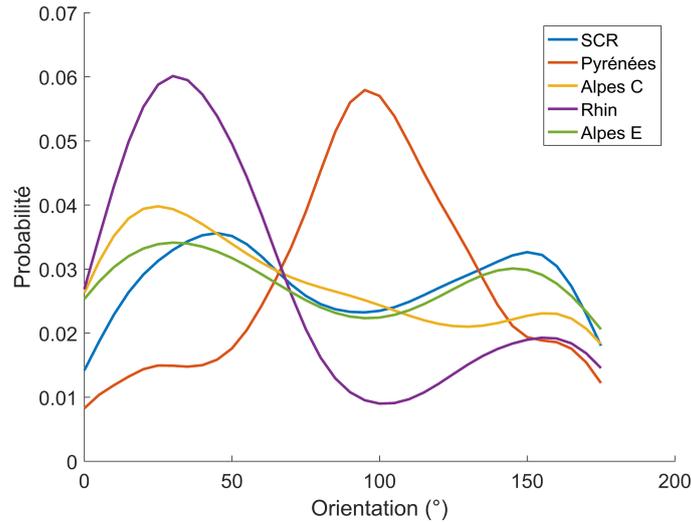


Figure 5.5: DENSITÉS DE PROBABILITÉ DES ORIENTATIONS DE FAILLES.

Ces densités sont calculées pour chaque région de la régionalisation SHARE (Figure 5.4b). La carte des tracés de faille (Figure 5.3a) est utilisée pour calculer ces densités. Alpes C et E réfèrent respectivement les domaines compressif et extensif des Alpes. Cette analyse n'a pas pu être réalisée au sein de la région 2 (croûte océanique) à cause du manque de données. La même analyse est réalisée pour les régions GEOTER. L'azimut est obtenu en ajoutant aléatoirement 0 ou 180° à l'orientation tirée.

La même procédure que pour l'azimut est appliquée pour le pendage mais avec un pas de 1° et une étendue fixée à 10°. Le régime de faille (normal, décrochant ou inverse) le plus fréquent dans ces mêmes bases de données est défini comme étant le régime caractéristique de la région. Si aucune ou trop peu de données sont disponibles dans une région, alors le régime est défini comme inconnu (Tableau 5.2).

La définition des intervalles de profondeur est plus arbitraire. En effet, lors de l'occurrence d'un séisme, les réseaux de sismographes fixent généralement la profondeur par tranche de 5 km afin de pouvoir trouver plus facilement la localisation du séisme. C'est pour cela que les profondeurs qui ressortent le plus souvent dans les catalogues instrumentaux sont 5, 10, 15 et 20 km. Ainsi, il a été préféré de procéder par tranche de 10 km. Suivant la profondeur moyenne des séismes d'une région, une tranche de 0-10 km, 5-15 km, 10-20 km ou 15-25 km est choisie comme intervalle pour cette région.

Pour la régionalisation SHARE, quelques particularités sont à noter. Du fait que les régions sont bien plus étendues que celles de la régionalisation GEOTER, les intervalles de valeurs sont généralement plus importants : jusqu'à 25 km pour la profondeur, 120° pour l'azimut et 40° pour le pendage (Tableau 5.3). De plus, plusieurs régimes de faille peuvent être associés à une seule région. La région continentale stable étant particulièrement vaste, il a été décidé que toutes les valeurs d'azimut (de 0 à 355°) pouvaient être tirées dans cette région. La région de croûte océanique ne comporte, quant à elle, aucune donnée de mécanisme au foyer. Il a donc été également décidé que toutes les valeurs d'azimut (de 0 à 355°) et de pendage (de 20 à 90°) pouvaient être tirées dans cette région.

Les Tableaux 5.2 et 5.3 récapitulent la magnitude maximale, le régime de faille et les intervalles de valeur des paramètres utilisés dans chaque région des deux régionalisations.

Région SHARE	$M_{max\ Région}$	Profondeur		Azimut		Pendage		Régime
		min	max	min	max	min	max	
1	6,5	0	25	0	355	47	87	N & S & R
2	5,5	5	30	0	355	20	90	U
3	7	0	15	20	140	50	82	N
4	7,3	0	20	-10	60	45	77	S & R
5	7	0	20	10	70	40	82	S
6	6,6	0	15	-20	50	37	71	N

Table 5.3: MAGNITUDE MAXIMALE ET PARAMÈTRES DE FAILLE DES RÉGIONS SHARE. Magnitude maximale et paramètres de faille définis dans cette thèse pour les régions SHARE. Le numéro des régions se réfère à la Figure 5.4b.

Régimes de faille : U = inconnu, N = normal, S = décrochant et R = inverse.

Une fois la densité spatiale et la régionalisation choisies (Figure 5.6a), le tirage spatial présenté Section 3.2.2 est réalisé. La Figure 5.6 illustre un exemple de tirage de la localisation d'un séisme M_w 6,5 appliqué à la régionalisation GEOTER. Une fois le séisme localisé dans l'espace, ses paramètres de faille sont tirés dans les gammes de valeurs correspondant à la région dans laquelle il se trouve (ici, région 31 Tableau 5.2).

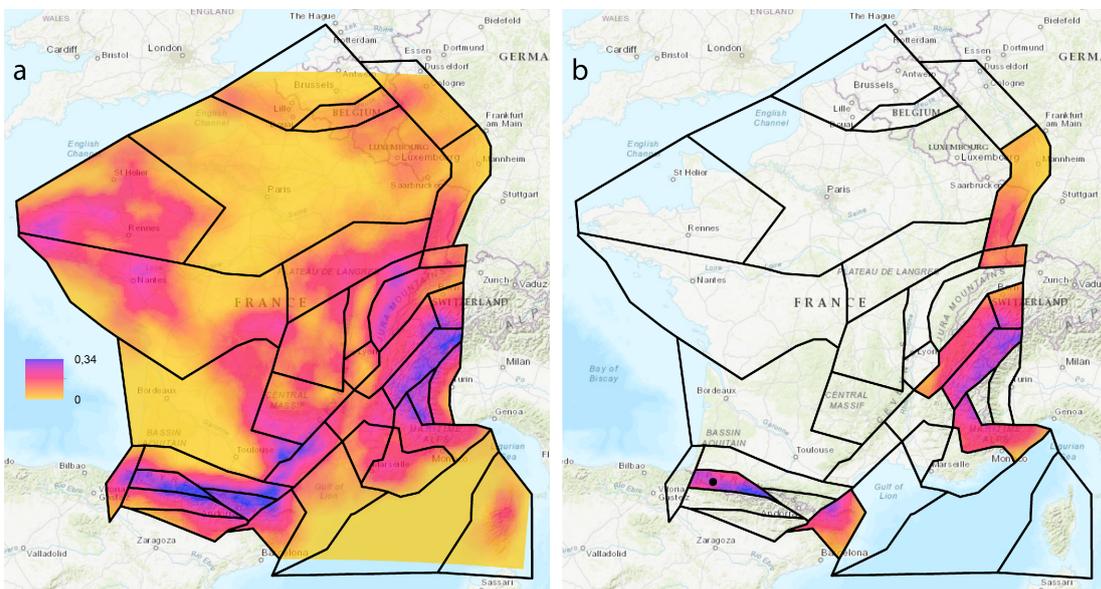


Figure 5.6: EXEMPLE DE TIRAGE SPATIAL D'UN SÉISME DE MAGNITUDE 6,5.

Tirage spatial d'un séisme de magnitude 6,5 dans la régionalisation GEOTER.

a : Densité spatiale de tracés de failles (Figure 5.3b) et régionalisation GEOTER (Figure 5.4a).

b : Limitation du tirage spatial aux seules régions dont $M_{max\ Région}$ (Tableau 5.2) est supérieure ou égale à 6,5. Dans cet exemple, la localisation du séisme (point noir) est tirée dans la région pyrénéenne 31.

5.2.3 Modèles de prédiction des mouvements du sol

Les modèles de prédiction des mouvements du sol (*Ground-Motion Prediction Equations*, GMPE) utilisés dans cette thèse pour calculer l'aléa sismique sont les six modèles utilisés dans une récente étude probabiliste appliquée en France métropolitaine [Martin et al., 2018] : Ak14 [Akkar et al., 2014a], Am14 [Ameri, 2014], Bi14 [Bindi et al., 2014], Bo14 [Boore et al., 2014], C15 [Cauzzi et al., 2015] et DC15 [Drouet et Cotton, 2015]. Am14 calcule uniquement l'accélération maximale du sol (*Peak Ground Acceleration*, PGA). Les cinq autres modèles peuvent de calculer la PGA ainsi que la vitesse maximale du sol (*Peak Ground Velocity*, PGV). La stratégie de pondération présentée Chapitre 4 est appliquée à ces six modèles afin de trouver leurs poids en fonction de la magnitude du séisme et de la distance séisme - site.

Chaque modèle possède ses propres intervalles de validité pour la magnitude M_w et la distance D . De plus, les distances peuvent être de plusieurs type : hypocentrale, épiscopentrale, Joyner-Boore (plus proche distance du plan de rupture projeté en surface) et rupture (plus proche distance du plan de rupture). Ces intervalles de validité ainsi que les types de distance autorisées par chacun des six modèles sont résumés Tableau 5.4. Au vu des plus faibles intervalles de validité ($M_w \geq 4,5$ et $D \leq 150$ km) et surtout de la seule distance de rupture autorisée, la GMPE de C15 n'est finalement pas utilisé dans cette thèse. Le seul type de distance commun aux cinq GMPE restantes est Joyner-Boore (JB Tableau 5.4). C'est donc ce type de distance qui est utilisé dans l'ensemble de la thèse.

Les GMPE de DC15 et d'Am14 ont été formulées à l'aide de données françaises. En ce qui concerne celles d'Ak14 et de Bi14, elles ont été obtenues à partir de données de l'Europe et du Moyen-Orient provenant de RESORCE [Akkar et al., 2014b], base de données présentée Section 5.2.4.1. Enfin, la dernière GMPE (Bo14) est établie à l'aide de données de régions actives de par le monde (NGA-West2 Ancheta et al. [2014]), provenant principalement de Californie, de Tawaï, de Chine et du Japon.

GMPE	M_w	D (km)	Type de D
Ak14	4 - 8	0 - 200	H, E, <u>JB</u>
Am14	3 - 7,5	0 - 300	E, <u>JB</u>
Bi14	4 - 7,6	0 - 300	H, <u>JB</u>
Bo14	3 - 7 (N) 3 - 8,5 (R & S)	0 - 400	<u>JB</u>
C15	4,5 - 7,9	0 - 150	R
DC15	3 - 8	0 - 250	H, E, <u>JB</u> , R

Table 5.4: DOMAINES DE VALIDITÉ ET TYPE DE DISTANCE DES GMPE UTILISÉES. Types de distance : H=hypocentrale / E=épiscopentrale / JB=Joyner-Boore / R=rupture. La distance JB utilisée dans cette thèse est soulignée. La GMPE de C15 n'est finalement pas utilisée car elle n'utilise pas la distance JB et est caractérisée par de plus faibles intervalles d'application. L'intervalle de magnitude de Bo14 dépend du régime de faille normal (N), inverse (R) ou décrochant (S).

5.2.4 Catalogue d'accélération

5.2.4.1 Présentation des données

La base de données d'accélération européennes RESORCE [Akkar et al., 2014b] sont utilisées dans cette étude. Cette base de données comporte plus de 5 600 enregistrements d'accélération (champs d'onde) dans les trois directions de l'espace et couvre un territoire allant de l'Iran aux Açores et de l'Islande à l'Égypte. Comme le montre la Figure 5.7, ces données viennent principalement de l'Ouest de la Turquie, des Appenins, de la Suisse, de la Grèce et de l'Islande. Des informations sur le séisme (localisation, temps d'occurrence, magnitude, profondeur, régime de faille, etc.), sur la station (localisation, altitude, type de sol, etc.) ainsi que les distances séisme-site (épacentrale, hypocentrale, JB, rupture) ou encore la bande de fréquences sont renseignées pour chaque enregistrement.

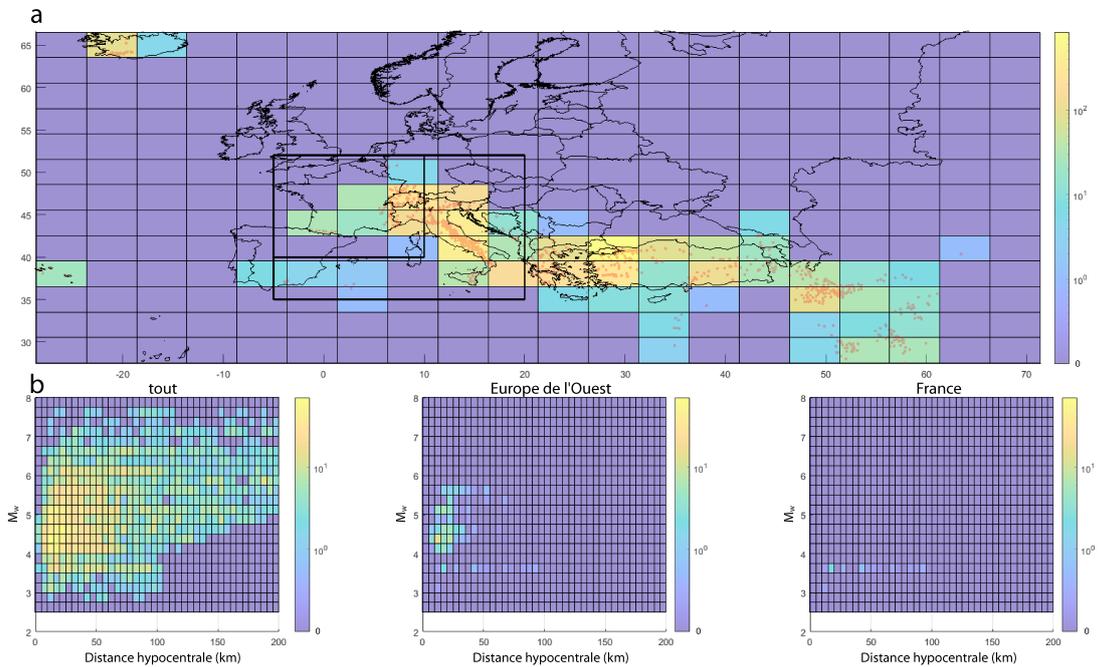


Figure 5.7: RÉPARTITIONS DES DONNÉES D'ACCÉLÉRATIONS CONTENUES DANS RESORCE.

a : Répartition spatiale ($5^\circ \times 3^\circ$) des enregistrements RESORCE (points rouges). Les encadrés noirs représentent les limites spatiales, définies dans cette thèse, de l'Europe occidentale (latitude : $35^\circ N - 52^\circ N$, longitude : $5^\circ O - 20^\circ E$) et de la France métropolitaine (latitude : $40^\circ N - 42^\circ N$, longitude : $5^\circ O - 10^\circ E$).

b : Répartition des données RESORCE en fonction de la distance hypocentrale et de la magnitude de moment pour l'ensemble de la base de données (gauche), l'Europe occidentale (milieu) et la France métropolitaine (droite).

5.2.4.2 Sélection des données d'accélération

Ayant pour objectif final l'estimation du risque sismique, seuls les événements pouvant provoquer des dégâts sont considérés. La magnitude de moment minimale de l'étude est fixée à 4. La distance Joyner-Boore maximale est fixée à 200 km. Du fait que la

France métropolitaine soit une région continentale à sismicité peu profonde, cette étude ne s'intéresse qu'aux séismes ayant une profondeur inférieure ou égale à 30 km. Enfin, l'aléa sismique est classiquement estimé au rocher. Seules les données ayant été enregistrées au rocher sont donc conservées. Le rocher caractérise un sol dont la vitesse moyenne V_{S30m} des ondes S dans les 30 premiers mètres du sol est supérieure à 800 m/s (classe A EuroCode8).

Cette sélection aboutit à seulement 39 accélérations enregistrées et disponibles dans la base RESORCE, produites par onze séismes sur le territoire français. Comme l'illustre la Figure 5.7b et le mentionnent Martin et al. [2018], cet échantillon d'enregistrements français est trop petit pour tenter un test des modèles de prédiction des mouvements du sol qui aiderait à attribuer les poids d'une manière objective.

Je propose de prendre en compte les données de toute l'Europe occidentale en faisant l'hypothèse que les sismicités française et européenne de l'Ouest sont comparables. Ainsi, en considérant les conditions énumérés ci-avant ($M_w \geq 4$, $D_{JB} \leq 200$ km, $z \leq 30$ km et au rocher), 485 enregistrements RESORCE provenant de l'Europe occidentale et produits par 160 séismes sont utilisés dans cette thèse. La Figure 5.8 présente la distribution du nombre de ces accélérations en fonction de la magnitude du séisme et de la distance JB séisme-station.

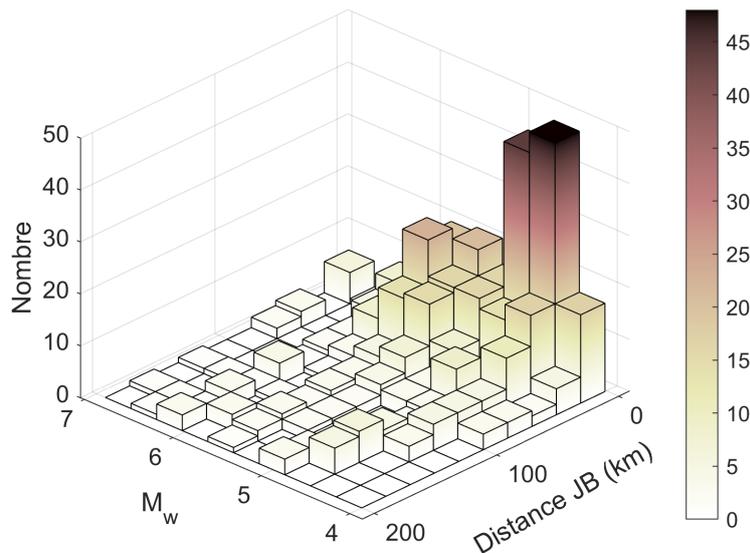


Figure 5.8: DISTRIBUTION MAGNITUDE - DISTANCE DES ENREGISTREMENTS UTILISÉS. Histogramme du nombre d'enregistrements RESORCE provenant de l'Europe occidentale en fonction de la magnitude et de la distance JB (km). Ces données respectent les conditions suivantes : $M_w \geq 4$, $D_{JB} \leq 200$ km, $z \leq 30$ km et au rocher.

Un bon échantillonnage d'enregistrements est observé à une distance JB inférieure à 60 km et à des magnitudes de moment comprises entre 4 et 6. La Figure 5.8 illustre également une nette prédominance des séismes $M_w 4,3 - 4,9$ à moins de 20 km.

5.3 Proportions et fréquences de séismes principaux en France

Dans cette Section sont présentées les distributions proportion - magnitude (Section 5.3.1) et fréquence - magnitude (Section 5.3.2) de séismes principaux estimées à partir des données françaises décrites et traitées ci-dessus (Section 5.2). Ces distributions servent de données d'entrée au générateur de séismes synthétiques détaillé Section 5.4.

Une partie des résultats présentés dans cette Section a été publiée dans *Journal of Seismology* [Gouache et al., 2021].

Pour rappel, la magnitude minimale de l'étude est égale à 4. Le seuil MIMP (Section 2.2) utilisé dans cette thèse est de 0,9. Le nombre N_{cb} de catalogues bruités par le processus de Monte Carlo (Section 2.3) est de 1000 pour l'ensemble des résultats présentés ci-dessous. La propagation de l'incertitude des magnitudes du catalogue initiale se fait à l'aide d'une loi normale de moyenne M_w et d'écart-type $ErrM_w/2$. Enfin, au vu de la longue temporelle et de la précision du catalogue historique, les temps de latence sont calculés au jour près.

Le nombre cumulé de séismes dans le catalogue initial est présenté Figure 5.9 (cercles rouges) en fonction de la magnitude. Le calcul de ce nombre cumulé considère les couples $A_c - M_c$ listés dans le Tableau 5.1. Les changements d'années de coupure sont d'ailleurs visibles aux magnitudes 4,7 5,6 et 5,9. La distribution moyenne du nombre cumulé de séismes dans les 1000 catalogues FCAT bruités est également représentée Figure 5.9 (courbe noire). La propagation de l'incertitude de la magnitude à travers un processus de Monte Carlo augmente légèrement le nombre cumulé moyen de séismes par magnitude.

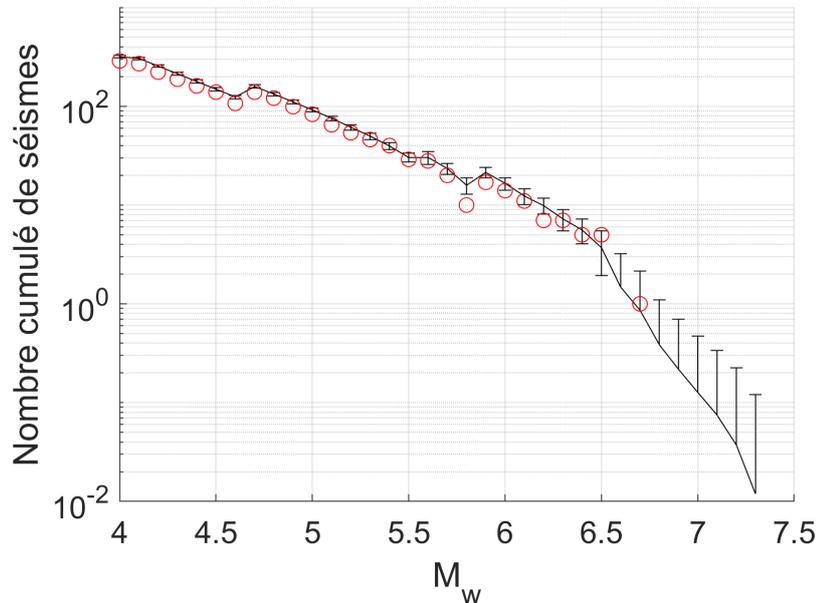


Figure 5.9: NOMBRE CUMULÉ DE SÉISMES DANS LE CATALOGUE D'ÉTUDE.

Le nombre cumulé de séismes contenus dans FCAT est représenté par tranche de magnitude par les cercles rouges. La courbe noire correspond à la moyenne et l'écart-type du nombre cumulé de séismes dans les $N_{cb} = 1000$ catalogues bruités.

5.3.1 Distribution proportion - magnitude de séismes principaux en France

5.3.1.1 Résultats obtenus

À partir du catalogue FCAT exhaustif, les proportions de séismes principaux vont de 0,7 à 0,94 de M_w 4 à M_w 5,2 en passant par 0,6 à M_w 4,8 (Figure 5.10). Les proportions de séismes principaux $M_w \geq 5,3$ sont, quant à elles, égales à 1. Cependant, les proportions ne sont pas calculées pour les magnitudes supérieures ou égales à 5,4 (Figure 5.10) ; la proportion de la magnitude maximale 6,7 ne comptant pas puisqu'elle est automatiquement fixée à 1 d'après l'Algorithme 2. En effet, à partir de M_w 5,4, les proportions de séismes principaux calculées par la méthode des temps de latence sont supérieures à 1, ce qui est impossible. Ces valeurs incohérentes ne sont donc pas représentées Figure 5.10. Nous attribuons ces incohérences au manque de données. Or, d'après la Figure 5.9, 40 séismes ont une magnitude supérieure ou égale à 5,4 et sont donc utilisés pour le calcul de la proportion de séismes principaux. Dans cette étude, 40 semble donc être le nombre minimal de séismes pour l'application de la méthode des temps de latence.

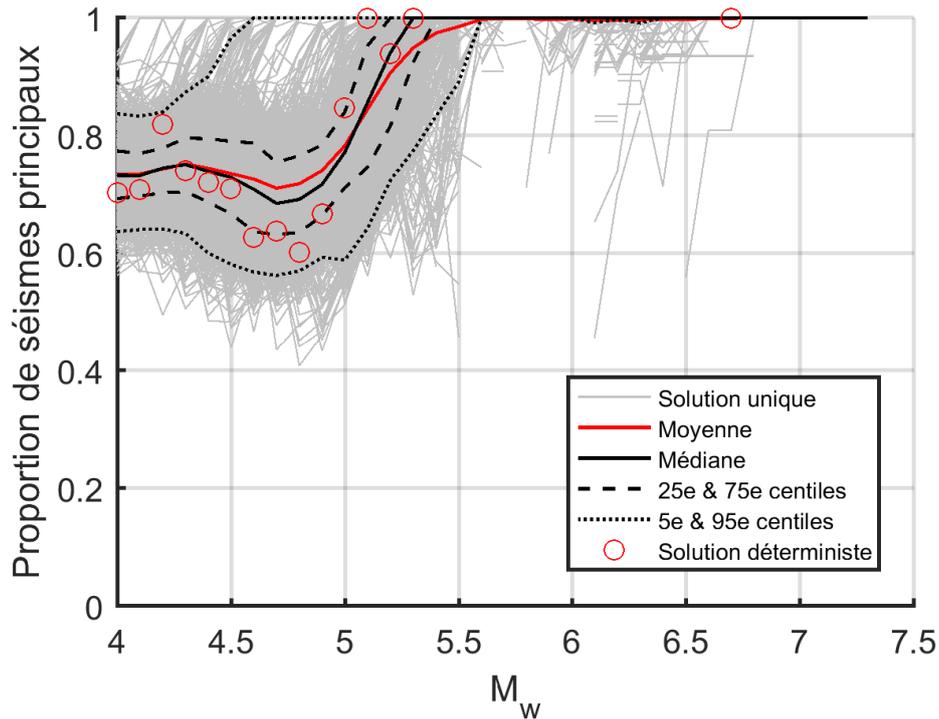


Figure 5.10: PROPORTIONS DE SÉISMES PRINCIPAUX EN FRANCE.
Une solution unique est le résultat obtenu à partir d'un des 1000 catalogues bruités.

Les proportions de séismes principaux obtenues à partir des 1 000 catalogues bruités sont également visibles Figure 5.10. La proportion maximale 1 est atteinte par la valeur médiane à la magnitude 5,3. D'après la distribution médiane, il n'y aurait donc plus de répliques à partir de M_w 5,3. À partir de cette magnitude, la valeur moyenne vaut quant à elle 0,95 et n'atteint 1 qu'à la magnitude 5,6.

Entre les magnitudes 5,8 et 6,6, quelques scénarios proposent des proportions se détachant de la valeur maximale 1 pour osciller entre 0,5 et 1 (Figure 5.10). Ces perturbations sont probablement dues au changement d'année de coupure (de 1600 à 1310) observé à la magnitude 5,9 (Tableau 5.1). La cause de telles perturbations peut également provenir de la sur-représentation de la magnitude 6,5 dans le catalogue FCAT. D'après la Figure 5.9, le catalogue FCAT comporte quatre séismes $M_w 6,5$ alors que les magnitudes proches ne sont représentées que par trois séismes (deux $M_w 6,3$ et un $M_w 6,7$). Quelle que soit leur origine, ces quelques faibles proportions ne représentent néanmoins que peu de résultats. En effet, au moins 95% des proportions calculées sont toujours égales à 1 dans cette gamme de magnitudes (Figure 5.10).

Les proportions calculées de séismes principaux peuvent être comparées à celles observées dans le catalogue FCAT. En effet, pour des magnitudes supérieures ou égales à 5,1 et dans les périodes de complétude associées (Tableau 5.1), pas plus de 65 séismes sont répertoriés dans le catalogue FCAT (Figure 5.9). Les dépendances entre ces séismes peuvent donc être décelées visuellement à la manière d'un algorithme de *declustering* à fenêtres spatio-temporelles. Le Tableau 5.5 résume les proportions de séismes principaux relevées à la main dans FCAT pour chaque pas de magnitude. La proportion de séismes principaux $M_w \geq 5,2$ observée dans le catalogue FCAT (0,93 Tableau 5.5) est proche de la proportion médiane estimée via la méthode des temps de latence (0,94 Figure 5.10). En revanche, pour les magnitudes supérieures, les proportions observées du catalogue FCAT oscillent entre 0,9 et 0,95 (Tableau 5.5) et tandis que nos résultats moyens sont compris entre 0,95 et 1 (Figure 5.10). Il faut attendre la magnitude 5,8 avant d'aboutir à une population constituée uniquement de séismes principaux.

$M_w \geq$	5,1	5,2	5,3	5,4	5,5	5,6	5,7	5,8	5,9
A_c	1750	1750	1750	1750	1750	1600	1600	1600	1310
Nombre	65	54	46	40	29	28	20	10	17
Proportion	0,94	0,93	0,91	0,92	0,9	0,96	0,95	1	1

Table 5.5: PROPORTIONS DE SÉISMES PRINCIPAUX OBSERVÉES DANS FCAT.

Le nombre de séismes contenus dans le catalogue FCAT sur la période associée (Tableau 5.1) et de magnitude supérieure ou égale à chaque pas de magnitude est reporté dans ce Tableau. La proportion de séismes principaux observée visuellement au sein de ces séismes est également indiquée.

La méthode des temps de latence [Molchan, 2004], sans l'implémentation MIMP, a également été appliquée aux données françaises. Les résultats stochastiques de cette méthode sont visibles Figure 5.11b tandis que la Figure 5.11a montre les mêmes résultats que la Figure 5.10. Les proportions médiane et moyenne de séismes principaux obtenues avec ces deux approches sont similaires jusqu'à $M_w 5,1$. Au-delà de cette magnitude, ces proportions sont diamétralement opposées. En effet, si les proportions obtenues à l'aide de l'implémentation MIMP continuent d'augmenter jusqu'à se stabiliser à 1, celles obtenues uniquement à partir de la méthode des temps de latence stagnent autour de 0,9 entre les magnitudes 5,2 et 6,3. Au-delà de $M_w 6,3$, la distribution de proportions chute jusqu'à ce que sa moyenne et sa médiane soient autour de 0,8.

Cette diminution est bien évidemment incohérente et, comme mentionné Section 2.2, reflète la perte de fiabilité des résultats de la méthode des temps de latence pour les fortes magnitudes lorsque les données manquent. D’après la Figure 5.9, à partir de $M_w 5,2$ les catalogues FCAT bruités comportent en moyenne 60 séismes. Dans cette étude, le nombre minimal de données disponibles afin d’appliquer la méthode des temps de latence semble donc être 60. Une valeur seuil de 40 séismes a déjà été relevée lors de l’analyse des résultats déterministes. À l’inverse, l’efficacité de l’implémentation MIMP concernant la stabilisation des proportions aux fortes magnitudes est clairement démontrée avec cette application.

5.3.1.2 Comparaison avec d’autres études

Les proportions de séismes principaux calculées à l’aide de l’implémentation MIMP sont aussi comparées à celles obtenues à partir de l’algorithme de *declustering* par fenêtres spatio-temporelles de Grünthal [1985]. Ce dernier a été développé à partir de données européennes et mis à jour par Burkhard et Grünthal [2009]. Il semble donc être un des algorithmes de *declustering* les plus légitimes à être utilisés pour une application à la France métropolitaine. Il est appelé G85 par la suite.

Les proportions de séismes principaux obtenues via cette approche (Figure 5.11c) sont sensiblement différentes de celles calculées dans cette thèse (Figure 5.11a). La valeur médiane des premières augmente lentement de 0,85 à 1 entre les magnitudes 4 et 5,8 tandis que sa valeur moyenne atteint 1 à $M_w 6,2$. En revanche, la valeur médiane des dernières stagne entre 0,7 et 0,75 pour les magnitudes comprises entre 4 et 4,9 avant d’augmenter fortement de 0,79 à 1 pour $M_w 5-5,3$. Notre distribution proportion-magnitude est donc pratiquement binaire : les proportions associées aux magnitudes inférieures à 5 oscillent autour de 0,7 tandis que les autres plafonnent à 1. Cet aspect binaire semble peu réaliste et ne correspond pas à ce qui est observé à la main dans le catalogue FCAT (Tableau 5.5). Il est probable que cet aspect soit dû à l’implémentation MIMP qui fixe souvent la proportion à 1 lorsque les données sont peu nombreuses.

La distribution proportion-magnitude calculée dans cette thèse atteint des proportions plus faibles que celle calculée via l’algorithme de *declustering* (0,7 VS 0,85). Un facteur pouvant expliquer ceci est que la plupart des algorithmes de *declustering* n’enlèvent pas assez de répliques ; même si l’algorithme G85 est réputé [par ex., van Stiphout et al., 2012] pour être un des algorithmes en supprimant le plus. De plus, comme mentionné Section 1.4.2.3, l’algorithme de *declustering* de Reasenbergh [1985] trouve une proportion de séismes principaux plus élevée dans des sismicités synthétiques simulées par un modèle ETAS que la proportion donnée en entrée du modèle [Beauval et al., 2006, Hainzl et al., 2006]. À l’inverse, ces deux études ont démontré que la méthode des temps de latence aboutit à des résultats plus proches de la proportion de référence du modèle. Nuançons tout de même ces conclusions en rappelant qu’un modèle est par nature une représentation approximée de la réalité. La question sous-jacente est : les conclusions faites à partir d’un modèle peuvent-elles s’appliquer directement sur les données réelles ?

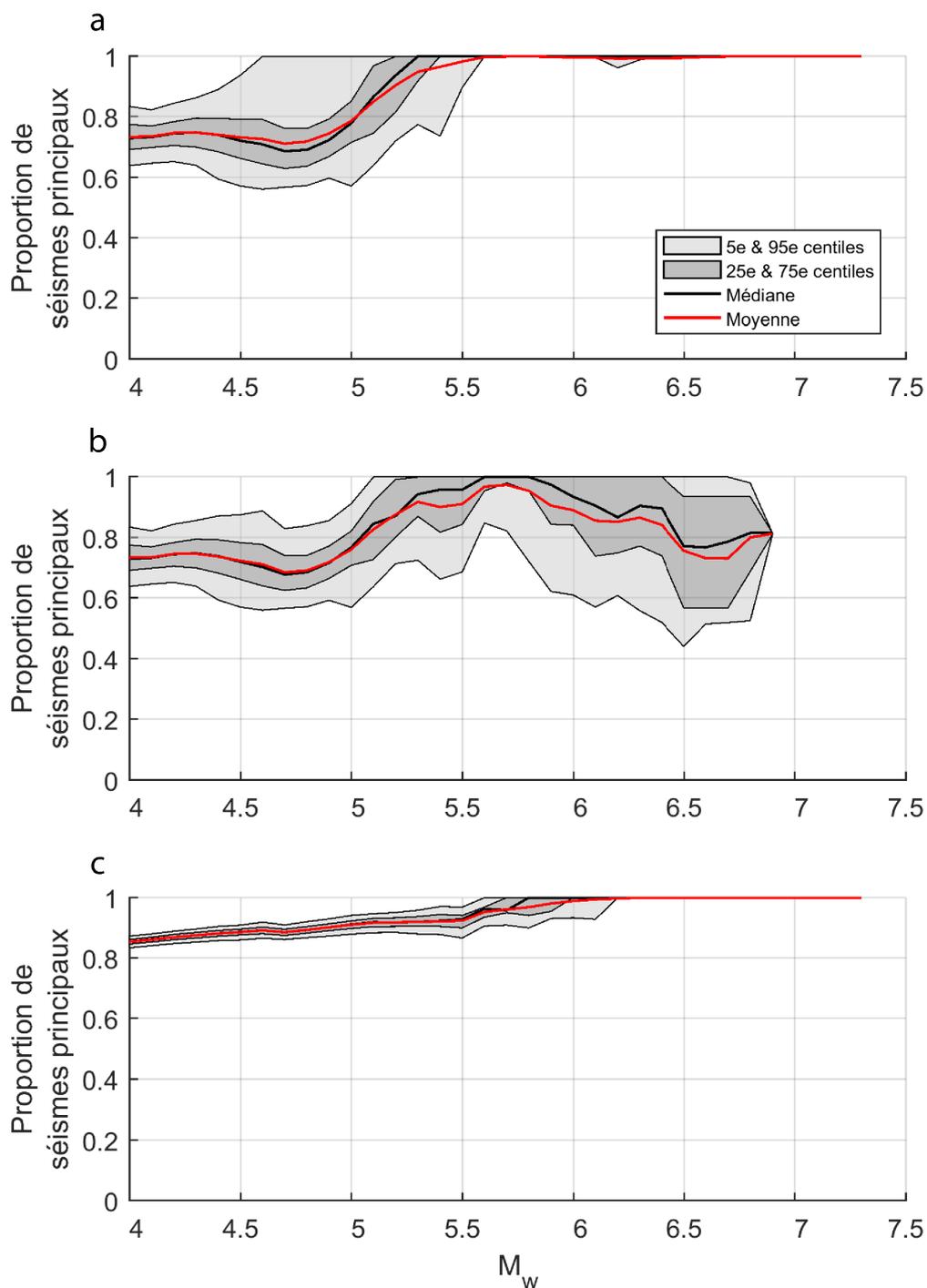


Figure 5.11: COMPARAISON DE PROPORTIONS DE SÉISMES PRINCIPAUX. Proportions de séismes principaux obtenues via trois approches différentes :
 a : Méthode des temps de latence [Molchan, 2004] avec l'implémentation MIMP.
 b : Méthode des temps de latence [Molchan, 2004] sans l'implémentation MIMP.
 c : Algorithme de *declustering* [Grünthal, 1985, Burkhard et Grünthal, 2009].

En ce qui concerne les magnitudes supérieures à 5,2, l'algorithme de *declustering* aboutit à des proportions plus proches que celles observées dans le catalogue FCAT (Tableau 5.5). Pour les magnitudes inférieures à 5,2, il est encore difficile de dire quelle modèle entre celui des temps de latence (Figure 5.11a) et celui du *declustering* (Figure 5.11c) est le plus proche de la réalité. Le modèle proposé dans cette thèse ne permet pas de mieux comprendre le phénomène très complexe qu'est la dépendance sismique. Pour autant, il propose une vision alternative aux algorithmes de *declustering* qui ne détecteraient pas assez de répliques. Il permet surtout d'illustrer les fortes incertitudes liées à ce phénomène difficile à cerner.

5.3.2 Distribution fréquence - magnitude de séismes principaux en France

5.3.2.1 Résultats obtenus

Les distributions fréquence - magnitude (DFM) déterministe et stochastique sont visibles Figure 5.12. Pour la même raison que pour la proportion de séismes principaux (Section 5.3.1), le calcul de la distribution déterministe de la fréquence s'arrête à la magnitude 5,3. Sur la gamme de magnitudes allant de 4 à 5,3, le b déterministe vaut 0,73 (Tableau 5.6). Sur cette même gamme de magnitudes, les b moyen et médian de la distribution stochastique valent 0,75 (Tableau 5.6).

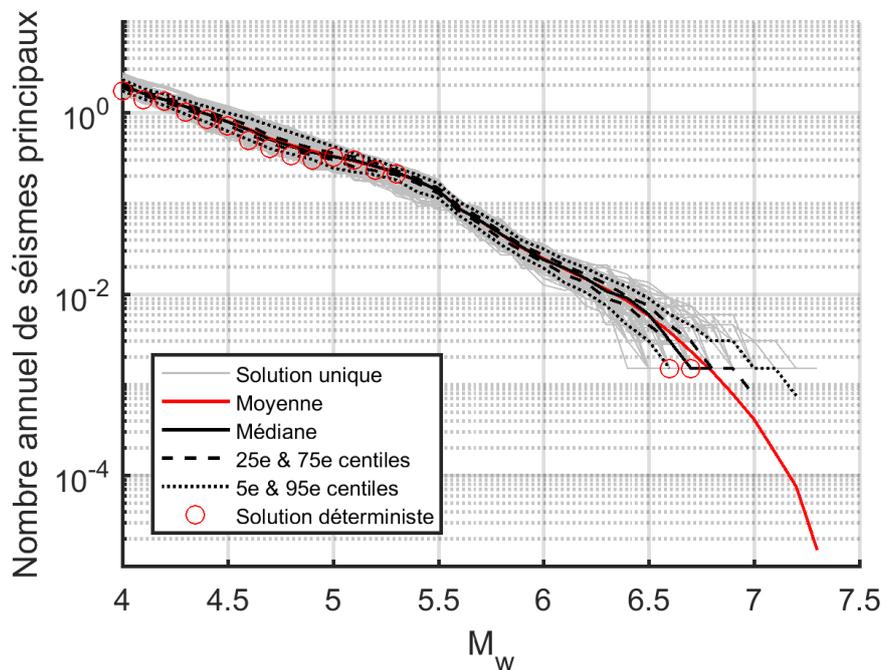


Figure 5.12: FRÉQUENCES ANNUELLES DE SÉISMES PRINCIPAUX EN FRANCE.

Chaque solution unique représente le résultat obtenu à partir d'un des 1 000 catalogues bruités. Autour de la magnitude 6,7 (magnitude maximale de FCAT), la fréquence minimale non nulle qu'il est possible d'obtenir sur les 655 années ($A_c = 1310$ Tableau 5.1) du catalogue est égale à $1/655 = 1,5 \cdot 10^{-3} \text{ an}^{-1}$. Voilà pourquoi seule la moyenne plonge vers la magnitude maximale de l'étude (7,3).

Sur l'ensemble de la DFM, les résultats déterministes sont toujours plus faibles que les résultats médians et moyens. La DFM déterministe est en permanence comprise entre la valeur médiane et le 5^{ème} centile de la DFM stochastique. Il semblerait que la prise en compte des incertitudes sur la magnitude ait tendance à augmenter la fréquence annuelle des séismes principaux. Ce constat fait écho à l'observation faite Figure 5.9 : la prise en compte des incertitudes sur la magnitude tend à augmenter le nombre cumulé de séismes par tranche de magnitude.

		$M_w 4-5,3$	$M_w 5,6-6$
Déterministe		0,73	-
Stochastique	Moyenne	0,75	1,39
	Médiane	0,75	1,40

Table 5.6: COEFFICIENTS b DES DEUX LOIS DE PUISSANCE OBSERVÉES. Valeurs de b calculées sur différents intervalles de magnitudes sur les résultats déterministe et stochastique.

De manière générale, l'écart entre les 5^{ème} et 95^{ème} centiles des résultats stochastiques augmente avec la magnitude. Ainsi, la fréquence annuelle de séismes principaux diminue tandis que son incertitude augmente avec la magnitude, comme le montrent les densités de probabilité des fréquences Figure 5.13. Ce constat est principalement dû à l'augmentation des incertitudes $ErrM_w$ sur la magnitude lorsque cette dernière augmente.

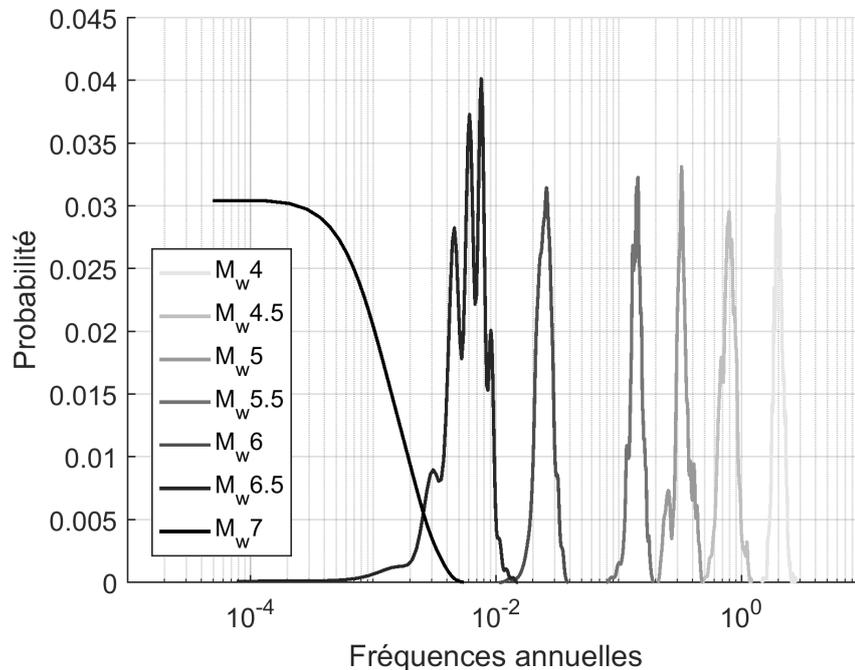


Figure 5.13: DENSITÉ DE PROBABILITÉ DES FRÉQUENCES DE SÉISMES PRINCIPAUX. Ces distributions de probabilité de fréquences de séismes principaux sont pondérées par le nombre de données utilisées pour les calculer. Ces distributions sont ici présentées pour chaque pas 0,5 de magnitude.

5.3.2.2 Deux lois de puissance

La principale observation faite Figure 5.12 est la répartition de la DFM en deux parties bien distinctes. La première est visible entre $M_w 4$ et $M_w 5,5$ tandis que la deuxième s'observe au-delà de $M_w 5,5$. Ce constat fait écho à l'observation de deux lois de puissance déjà discutée Section 1.4.2.2 et illustrée à l'échelle mondiale Figure 1.7. La première partie est caractérisée par une loi de puissance de coefficient moyen $b = 0,75$. La pente moyenne de la deuxième partie est, quant à elle, égale à 1,39 (Tableau 5.6). Cette valeur de pente est cohérente avec celles trouvées par Rundle [1989] (1,5) et Pacheco et al. [1992] (1,39) à l'échelle planétaire.

Schorlemmer et al. [2005] ont observé que b varie en fonction du régime de faille. À l'échelle planétaire, ces auteurs trouvent que b est en moyenne égal à 0,9, 1 et 1,15 respectivement pour les régimes inverse, décrochant et normal. Néanmoins, l'observation d'une deuxième loi de puissance de coefficient autour de 1,39 (Tableau 5.6) ne saurait être expliquée par un changement de régime de faille.

L'origine de ce comportement est traditionnellement attribuée à un changement de la dimensionalité du plan de rupture [voir Sornette et al., 1996, pour une liste non exhaustive]. L'idée est qu'à partir d'une certaine magnitude, la largeur du plan de rupture est aussi grande que l'épaisseur de la zone sismogène. Dans ce cas, la largeur est limitée en haut par la surface et en bas par la base de la zone sismogène comme l'illustre la Figure 5.14. Dans ces conditions, la rupture n'est plus définie par deux dimensions (longueur et largeur) mais par une seule (longueur). D'après Pacheco et al. [1992], ce changement de dimensionalité a des conséquences sur les propriétés d'échelle (par ex., la loi GR) entre les magnitudes faibles et modérées et les fortes magnitudes. Ces auteurs, tout comme Sornette et al. [1996], Sachs et al. [2012] et Corral et González [2019], observent la magnitude de transition entre les deux lois de puissance autour de 7,5 à l'échelle planétaire. Ils associent cette magnitude à une épaisseur de zone sismogène d'environ 60 km, correspondant à l'épaisseur de zone sismogène observée dans les zones de subduction dont est originaire la majorité des séismes $M_w \geq 7,5$ mondiaux. En analysant uniquement les séismes décrochants peu profonds, ils trouvent une magnitude de transition de 5,9-6 qu'ils associent à une largeur d'environ 10 km.

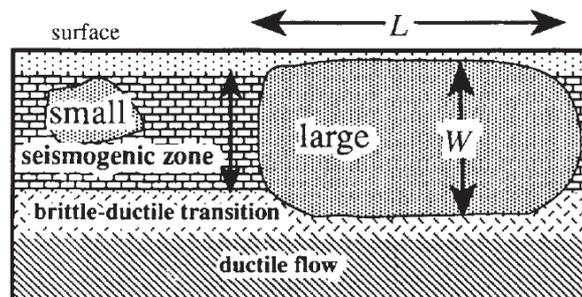


Figure 5.14: LIMITATION DE LA LARGEUR DU PLAN DE RUPTURE [Pacheco et al., 1992]. La largeur (W) du plan de rupture des puissants séismes (*large*) est limitée en haut par la surface et en bas par la transition cassant - ductile de la croûte terrestre. Dans ce cas, le plan de rupture n'est donc plus défini que par une seule dimension : sa longueur (L).

Nos résultats témoignent, quant à eux, d'une magnitude de transition égale à 5,5-5,6 pour une sismicité peu profonde. En se basant sur l'hypothèse de Pacheco et al. [1992], la zone sismogène française serait donc inférieure à 10 km. Cette observation est discutable étant donné que selon le catalogue SIHex, faisant référence à ce jour, 36% des séismes dans la zone d'étude auraient des profondeurs supérieures à 10 km.

L'observation de ces deux lois de puissance est donc due à une ou d'autre(s) explication(s). Nous avons considéré que la longue profondeur temporelle qu'offre le catalogue FCAT permettait de couvrir les magnitudes les plus élevées. Une des explication se trouve probablement dans l'inexactitude de cette hypothèse. Néanmoins, nos résultats parviennent à reproduire le comportement typique des magnitudes les plus élevées, à savoir le plongement asymptotique (Figure 5.12) ; et ce sans utiliser d'équation (par ex., Équation 1.4), seulement à partir de la simple observation des données. Cela apporte du crédit à la méthode utilisée sans pour autant affirmer que ses résultats sont bons. La DFM obtenue dans cette thèse (Figure 5.12) est donc à utiliser avec précaution étant donné qu'elle diffère de la DFM classiquement obtenue avec la loi GR (une seule pente) et utilisée dans les études de risque en France.

5.3.2.3 Impact de la sismicité corrélée sur la DFM

Le but de cette Section est de mesurer l'impact de la méthode de calcul de la proportion de séismes principaux sur la DFM. Pour cela, les coefficients a et b de la loi GR ont été calculés sur la gamme de magnitude 4-5,3 sur trois DFM différentes :

- la DFM moyenne de séismes principaux obtenue sur $N_{cb} = 1\,000$ catalogues FCAT bruités complets (séismes indépendants et dépendants).
- la DFM moyenne obtenue après l'application de l'algorithme de *declustering* G85 sur $N_{cb} = 1\,000$ catalogues FCAT bruités.
- la DFM moyenne de séismes principaux obtenue via la méthode des temps de latence (TL) dans cette thèse (Figure 5.12).

Les trois valeurs de b calculées pour ces trois cas sont visibles Tableau 5.7.

	Total	G85	TL
b	0,82	0,80	0,75
Écart normalisé (%)	9,2	6,9	0

Table 5.7: b OBTENUS À PARTIR DE TROIS DFM DIFFÉRENTES.

b moyens calculés entre $M_w 4$ et $M_w 5,3$ pour trois cas différents : sismicité totale et séismes principaux selon l'algorithme de *declustering* G85 et selon la méthode des temps de latence (TL). Les écarts normalisés par rapport au b de TL sont également renseignés.

Le coefficient b de la loi GR peut être interprété comme la proportion de séismes à faibles magnitudes par rapport aux séismes à fortes magnitudes. Ainsi, plus b est grand, plus la proportion de faibles magnitudes est grande vis-à-vis des fortes magnitudes. Étant donné que la proportion de séismes principaux augmente avec la magnitude (Figure 5.11a et b), le nombre de séismes à faibles magnitudes est plus faible dans

les catalogues résiduels (G85 et TL) que celui dans le catalogue total. Il est donc logique de constater une diminution de la valeur de b lorsqu'on enlève les séismes dépendants (G85 et TL Tableau 5.7). La diminution plus marquée avec la méthode des temps de latence qu'avec l'algorithme de *declustering* s'explique par le fait que la première enlève plus de répliques que le deuxième (Section 5.3.1.2). La diminution de b par la prise en compte des seuls séismes principaux est donc normale mais implique un biais lors de l'extrapolation de la courbe GR pour les plus fortes magnitudes. La Figure 5.15a illustre les différentes droites GR (sans troncature) obtenues en extrapolant aux magnitudes supérieures (M_w 5,4- 7,3) la droite obtenue entre les magnitudes 4 et 5,3. La diminution du coefficient b de la loi de puissance entraîne un dépassement de la fréquence de sismicité totale au niveau d'une magnitude M_x de croisement définie ainsi :

$$M_x = \frac{a_{total} - a}{b_{total} - b}, \quad (5.1)$$

avec a_{total} et b_{total} les coefficients de la loi GR estimés sur la sismicité totale.

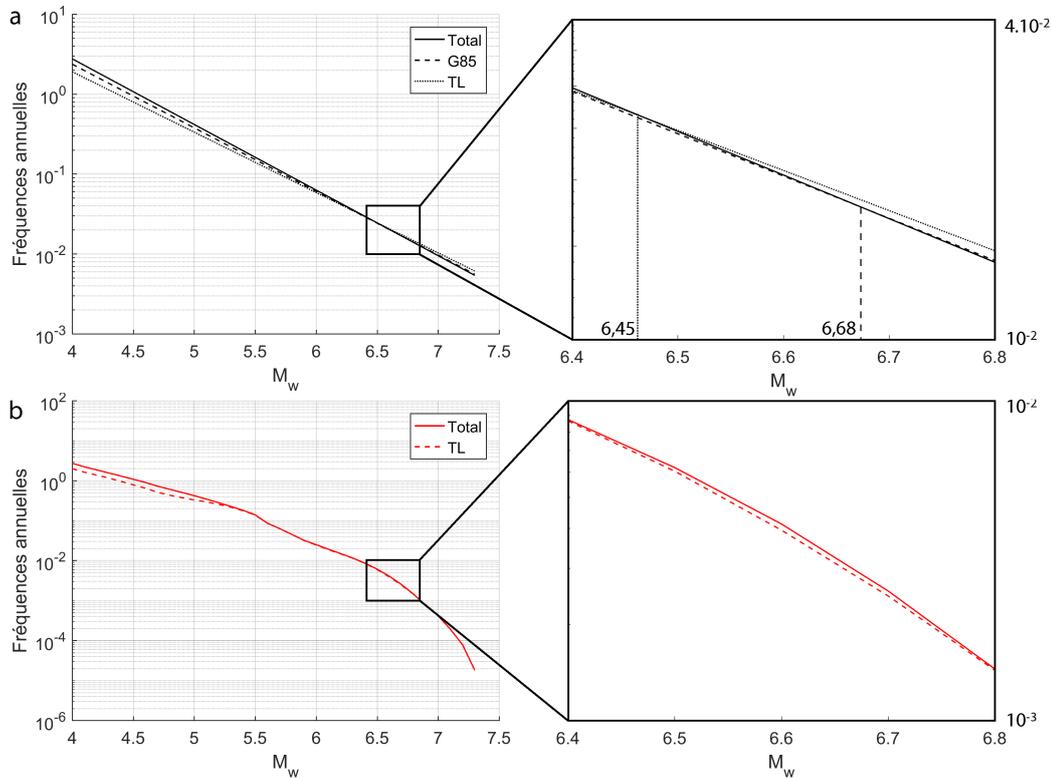


Figure 5.15: DFM OBTENUES VIA DIFFÉRENTES APPROCHES.

a : Les coefficients a et b sont calculés entre les magnitudes 4 et 5,3. Les droites GR sont ensuite extrapolées sans troncature pour les fortes magnitudes. Ces droites sont obtenues à partir de la sismicité totale et des séismes principaux selon l'algorithme de *declustering* G85 et selon la méthode des temps de latence (TL). L'encadré illustre les magnitudes de croisement M_x (Équation 5.1) associées à G85 (6,68) et à TL (6,45).

b : DFM moyenne de l'ensemble de la sismicité observée dans FCAT comparée à celle de séismes principaux calculée via la méthode des temps de latence (TL Figure 5.12). L'encadré montre que les fréquences de la sismicité totale sont bien supérieures ou égales aux fréquences des séismes principaux.

La fréquence des séismes principaux extrapolée se trouve supérieure à la fréquence de toute la sismicité extrapolée pour les magnitudes supérieures à M_x . D'après la Figure 5.15a, M_x vaut 6,45 avec la méthode des temps de latence et 6,68 pour l'algorithme de *declustering*. Cette surestimation de la fréquence, même si elle est restreinte aux très fortes magnitudes et peu flagrante, constitue une incohérence et est due à l'extrapolation de la loi GR. Mizrahi et al. [2020] ont récemment étudié l'impact de différents algorithmes de *declustering* sur la DFM en Californie. L'ensemble de ces algorithmes aboutissent à des b allant de 0,86 à 0,95 contre 1,01 pour la sismicité totale. Ces coefficients conduisent à des magnitudes de croisement variant de 7,2 à 8,8. Bien évidemment, l'application de la troncature sur ces droites tend à diminuer cette surestimation de la fréquence. Pour autant, elle reste présente avec des magnitudes de croisement observées à 6,8 pour la méthode des temps de latence et 7,1 pour G85.

À l'inverse, aucune extrapolation des fréquences de séismes principaux n'est opérée dans l'approche des temps de latence. Ainsi, la fréquence des séismes principaux est toujours inférieure ou égale à celle de la sismicité totale Figure 5.15b.

5.3.2.4 Influence du seuil MIMP

Le seul paramètre de l'implémentation MIMP est le seuil MIMP. Son influence sur les distributions moyenne (a) et médiane (b) de proportion de séismes principaux est présentée Figure 5.16.

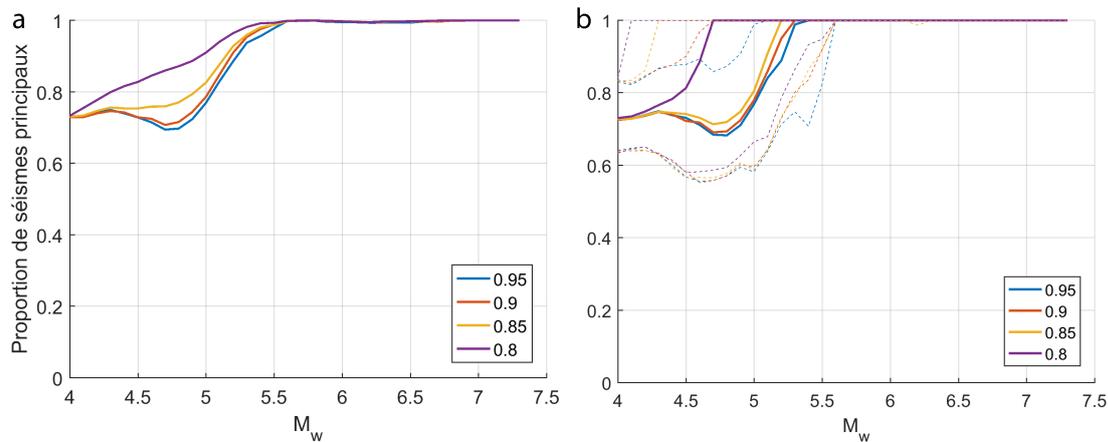


Figure 5.16: IMPACT DU SEUIL MIMP SUR LA PROPORTION DE SÉISMES PRINCIPAUX. a : Moyenne de la distribution proportion - magnitude de séismes principaux pour quatre seuils MIMP différents. b : Centiles (5, 50 et 95) de la distribution proportion - magnitude de séismes principaux pour quatre seuils MIMP différents.

La distribution médiane atteint la proportion maximale 1 aux magnitudes 5,4, 5,3, 5,2 et 4,7 pour des seuils MIMP de 0,95, 0,9, 0,85 et 0,8 respectivement (Figure 5.16b). Ainsi, plus le seuil MIMP est faible, plus la proportion maximale 1 est atteinte rapidement. Ce constat est cohérent avec le fonctionnement de l'implémentation MIMP. En effet, le seuil MIMP est la proportion à partir de laquelle, une fois atteinte, aucune autre proportion ne peut redescendre en dessous de ce seuil. Or, plus ce seuil est bas, plus les magnitudes l'atteignant sont faibles. De plus, ce constat signifie que l'impact

du seuil MIMP est nul pour les magnitudes supérieures ou égales à 5,5. En effet, à partir de cette magnitude les proportions moyennes et médianes de séismes principaux valent 1 quelle que soit la valeur du seuil MIMP (Figure 5.16). Pour cette gamme de magnitude, les distributions des proportions et des fréquences de séismes principaux sont donc identiques à celles observées Figures 5.10 et 5.12.

Qualitativement, en deçà de $M_w 5,5$, les écarts observés entre les distributions (médiane et moyenne) obtenues à l'aide des seuils 0,95 et 0,9 ne semblent pas significatifs. En revanche, le seuil 0,85 aboutit à une distribution moyenne légèrement distincte de celles obtenues par les seuils supérieurs. De plus, les distributions moyenne et médiane obtenues à l'aide du seuil 0,8 diffèrent fortement de celles estimées à l'aide des autres seuils.

En ce qui concerne l'impact du seuil MIMP sur la distribution fréquence - magnitude de séismes principaux, le Tableau 5.8 présente les valeurs de b obtenues par régression des distributions (moyenne et médiane) sur la gamme de magnitudes 4 - 5,3, obtenues à l'aide de quatre seuils MIMP. Les écarts normalisés par rapport au seuil de référence utilisé dans cette thèse (0,9) ne dépassent pas les 6,4% (Tableau 5.8), soit moins que l'écart observé entre les valeurs de b obtenues via G85 et la méthode des temps de latence (6,9% Tableau 5.7). Ces écarts sont même inférieurs à 2% en valeur absolue pour les seuils MIMP compris entre 0,85 et 0,95 (Tableau 5.8).

Seuils MIMP	0,8	0,85	0,9	0,95
b moyen	0,73	0,74	0,75	0,76
Écart normalisé (%)	-3,3	-1,7	0	0,9
b médian	0,70	0,73	0,75	0,76
Écart normalisé (%)	-6,4	-2	0	0,9

Table 5.8: VALEURS b CALCULÉES POUR DIFFÉRENTS SEUILS MIMP. b est calculé entre $M_w 4$ et $M_w 5,3$ sur les distributions moyenne et médiane de fréquences de séismes principaux. Les écarts normalisés par rapport aux valeurs obtenues à l'aide du seuil de 0,9 sont également calculés.

Au-delà d'une certaine valeur (0,8 ici), le choix du seuil MIMP a donc un impact négligeable sur les DPM et DFM de séismes principaux.

5.4 Génération de catalogues synthétiques français

Dans cette Section, le générateur de séismes est utilisé pour générer 50 000 ans de sismicité synthétique à partir des données françaises. Le terme catalogue synthétique fait donc référence à un catalogue de sismicité française synthétique d'une durée de 50 000 ans. Cette sismicité synthétique comprend les séismes principaux et les répliques $M_w \geq 4$.

Cette Section a pour vocation de comparer la sismicité synthétique avec les données d'entrée et le catalogue FCAT initial dans le but de s'assurer du bon fonctionnement du générateur. Dans cette optique, les distributions proportion-magnitude (DPM) et fréquence-magnitude (DFM) de séismes principaux observées dans le catalogue synthétique sont comparées Section 5.4.1 avec celles données en entrée du générateur. Ensuite, les moments sismiques libérés par les séismes synthétiques et les séismes du catalogue FCAT sont comparés Section 5.4.2. Enfin, la Section 5.4.3 présente la distribution spatiale de la sismicité synthétique et la compare avec la densité spatiale donnée en entrée du générateur.

5.4.1 Comparaison des distributions de proportions et de fréquences

Proportions de séismes principaux Étant donné que chaque séisme du catalogue synthétique est étiqueté comme étant une réplique ou un séisme principal, une proportion de séismes principaux synthétiques peut être calculée pour chaque magnitude. La Figure 5.17 compare les DPM observées sur 40 catalogues synthétiques de 50 000 ans à la DPM moyenne donnée en entrée du générateur (Figure 5.10). Ces DPM sont comparables (autour de 1) pour les magnitudes supérieures ou égales à 6,3. En revanche, le générateur ne parvient pas à reproduire la DPM donnée en entrée pour les magnitudes inférieures.

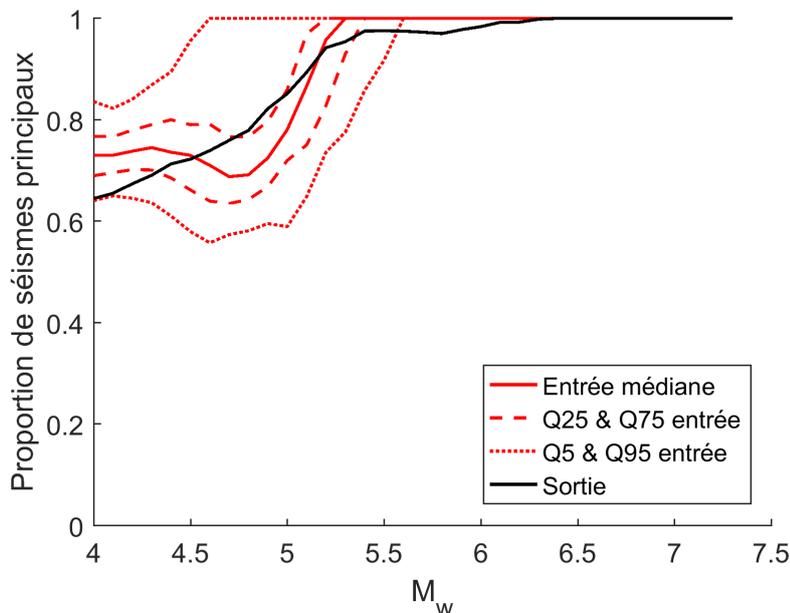


Figure 5.17: COMPARAISON DES DPM D'ENTRÉE ET DE SORTIE DU GÉNÉRATEUR. La DPM médiane (Figure 5.10) fournie comme donnée d'entrée au générateur est comparée avec la DPM observée dans un catalogue synthétique de 50 000 ans.

Ce décalage entre les proportions données en entrée et celles obtenues en sortie du générateur provient probablement de la méthodologie de calcul des densités de probabilité des proportions de séismes principaux.

Nombre de répliques Le nombre de répliques produites en fonction de la magnitude du séisme principal est présenté Figure 5.18. Ce nombre augmente lorsque la magnitude du séisme principal augmente, comme le préconise l'Équation 1.10 [Utsu, 1970]. La régression de cette Équation sur nos résultats entre les magnitudes 5,6 et 7,1 donne les coefficients $k = 0,029$ et $\alpha = 1,116$ avec $M_c = 4$. Certains auteurs ont estimé des α moyens égaux à 0,86 [Guo et Ogata, 1997], 0,83 [Nicolis et al., 2015] et 0,81 [Helmstetter, 2003] respectivement sur des données japonaises, chiliennes et californiennes. Helmstetter [2003] stipule que le cas $\alpha > b$ dans un modèle ETAS n'est probablement pas adéquat pour représenter la sismicité réelle. Or, pour les magnitudes supérieures à 5,6, b vaut en moyenne 1,39 (Tableau 5.6) et est donc supérieur au α calculé dans cette étude. Ainsi, sans utiliser de modèle ETAS, cette étude trouve un paramètre α proche de ce qui se calcule dans la littérature et qui respecte la condition $\alpha < b$.

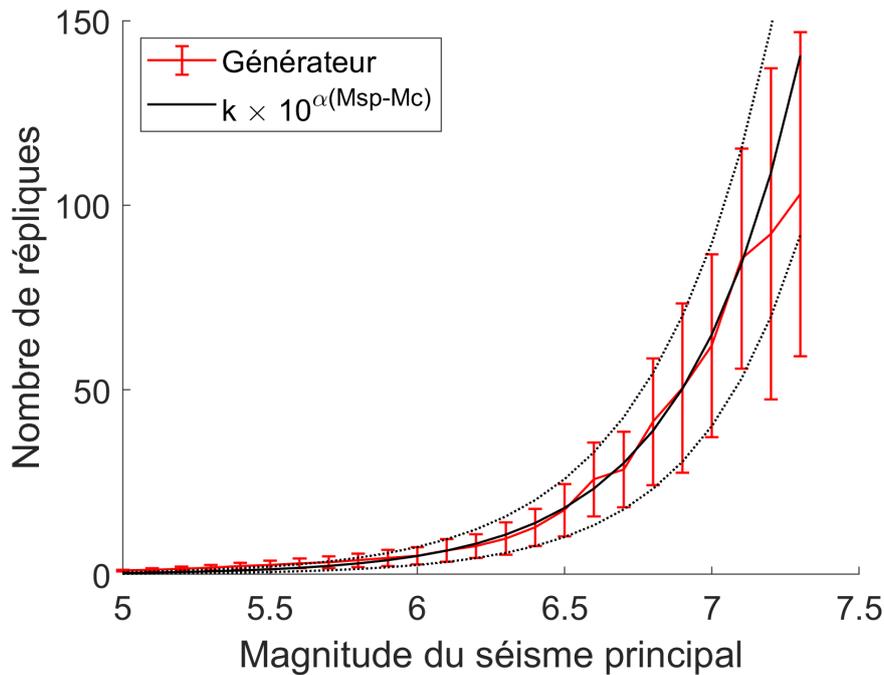


Figure 5.18: NOMBRE DE RÉPLIQUES PAR SÉISME PRINCIPAL.

Nombre de répliques moyen en fonction de la magnitude M_{sp} du séisme principal (rouge). Les barres correspondent à l'écart-type observé dans un catalogue synthétique de 50 000 ans. Une régression de l'Équation 1.10 (noir) réalisée entre M_w 5,6 et M_w 7,1 donne les coefficients $k = 0,029$ et $\alpha = 1,116$ avec $M_c = 4$. Les courbes en pointillés représentent les régressions sur le nombre moyen plus ou moins son écart-type.

Le nombre de répliques $M_w \geq 4$ produites par des séismes principaux $M_{sp} \geq 7,2$ ne suit plus l'augmentation définie par l'Équation 1.10. Ce constat est principalement dû à la sous-représentation (moins de 20) des séismes principaux de magnitude supérieure ou égale à 7,2 dans le catalogue synthétique de 50 000 ans.

Fréquences annuelles de séismes principaux La Figure 5.19 présente les distributions fréquence - magnitude (DFM) de séismes principaux observées dans 40 catalogues synthétiques de 50 000 ans. Ces DFM sont comparées à la DFM moyenne fournie en entrée du générateur. De manière générale, le générateur parvient à reproduire la DFM donnée en entrée. De légères variations sont observées au delà de la magnitude 5,6. À partir de cette magnitude, la DFM moyenne de séismes principaux observée dans les catalogues synthétiques est plus chaotique que celle donnée en entrée. La première est toujours (excepté pour les magnitudes 6,5 et 6,6) légèrement inférieure ou bien égale à la deuxième. Ces légers écarts sont probablement dus à la méthodologie retenue dans cette thèse (Équations 3.1 et 3.2) pour tirer les séismes principaux ainsi que sa mise en œuvre (par ex., calcul des densités de probabilité des fréquences).

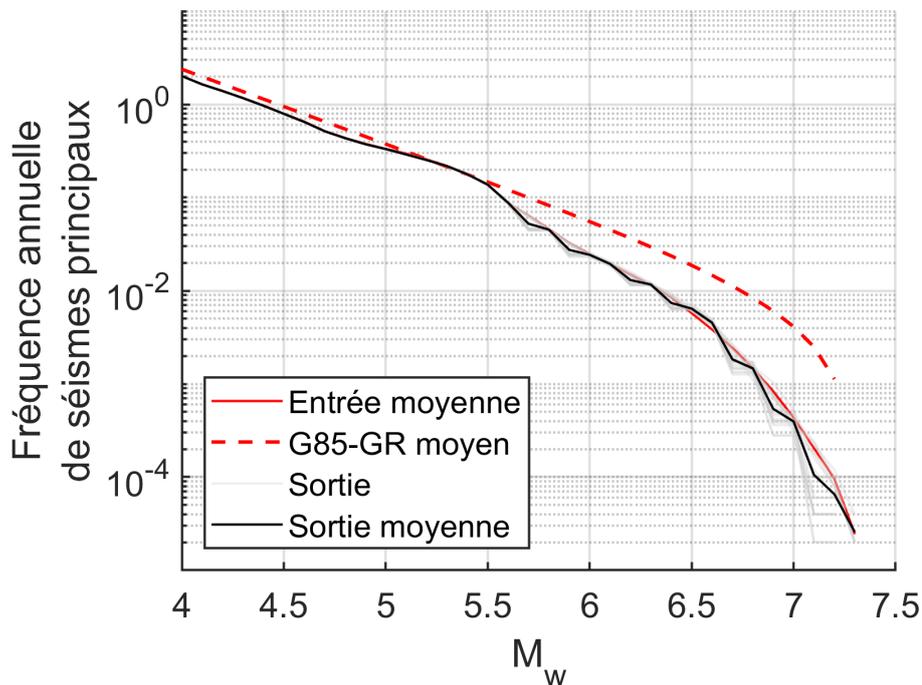


Figure 5.19: COMPARAISON DES DFM D'ENTRÉE ET DE SORTIE DU GÉNÉRATEUR. La DFM moyenne de séismes principaux (Figure 5.12) fournie comme donnée d'entrée au générateur est comparée avec les DFM observées en sortie dans 40 catalogues synthétiques de 50 000 ans. La DFM moyenne obtenue par l'application de l'algorithme de *declustering* de Grünthal [1985] (G85) et de la loi GR est également illustrée.

La DFM moyenne de séismes principaux calculée à l'aide de l'algorithme de *declustering* de Grünthal [1985] et de la loi GR est également représentée Figure 5.19.

Le Tableau 5.9 récapitule les périodes de retour des séismes principaux en fonction de magnitude selon les DFM moyennes donnée en entrée, observée dans la sismicité synthétique et obtenue en appliquant l'algorithme de *declustering* de Grünthal [1985] et la loi GR (Figure 5.19). En ce qui concerne les périodes de retour des séismes $M_w \geq 6,5$, un facteur 3 est observé entre la loi GR et la méthode des temps de latence. Ce facteur est de 10 pour les périodes de retour de séismes plus puissants ($M_w \geq 7$).

$M_w \geq$	4	4,5	5	5,5	6	6,5	7
G85-GR	0,4	1	2,7	6,8	18	53	246
Entrée	0,5	1,2	3	7,2	40	168	2 509
Sortie	0,5	1,3	3	7,3	41	156	2 528

Table 5.9: PÉRIODES DE RETOUR MOYENNES EN FONCTION DE LA MAGNITUDE. Périodes de retour (années) des séismes principaux en fonction de magnitude selon les DFM moyennes obtenue en appliquant l'algorithme de *declustering* de Grünthal [1985] et la loi GR (G85-GR), donnée en entrée et observée dans la sismicité synthétique.

5.4.2 Moment sismique cumulé au cours du temps

Le moment sismique produit par chaque séisme (principal et réplique) généré est calculé grâce à l'Équation 3.6. Le cumul dans le temps du moment sismique produit par les séismes générés est comparé avec celui des séismes contenus dans le catalogue FCAT. Pour se faire, les cinq périodes d'exhaustivité (1310 - 1964, 1600 - 1964, 1750 - 1964, 1830 - 1964 et 1850 - 1964, Section 5.2.1.4) sont analysées individuellement. Pour chacune de ces périodes, seuls les séismes à magnitude supérieure ou égale à la magnitude de coupure associée à cette période d'exhaustivité (Tableau 5.1) sont conservés dans le catalogue synthétique et dans le catalogue FCAT. Ensuite, le catalogue synthétique de 50 000 ans est divisé en une multitude de sous-catalogues de profondeur temporelle égale à la période en question. Chacun de ces sous-catalogues représente une réalisation comparable au catalogue FCAT. Le cumul du moment sismique de chacun de ces sous-catalogues est comparé Figure 5.20 à celui du catalogue FCAT sur les cinq périodes d'exhaustivité.

Une tendance s'observe clairement Figure 5.20 : plus la période d'exhaustivité est longue (par ex., 1310 - 1964 et 1600 - 1964), plus les moments sismiques médians cumulés par les sous-catalogues synthétiques sont élevés vis-à-vis de ceux observés dans le catalogue FCAT. À l'inverse, pour la période d'exhaustivité la plus courte (1850 - 1964), le moment sismique médian cumulé par les sous-catalogues synthétiques est inférieur à celui associé au catalogue FCAT. Enfin, pour les périodes d'exhaustivité 1750 - 1964 et 1830 - 1964, les résultats du générateur sont comparables à ceux du catalogue FCAT.

La surestimation du moment sismique produit par les catalogues synthétiques pour les périodes d'exhaustivité les plus longues a diverses origines. La raison principale est que les DFM observées sur les catalogues synthétiques sont supérieures à celle donnée en entrée pour les séismes principaux $M_w \geq 6,3$ (Figure 5.19). La sur-représentation de séismes $M_w \geq 6,3$ dans les catalogues synthétiques implique une surestimation du cumul du moment sismique vis-à-vis de celui du catalogue FCAT. De plus, aucune réplique $M_w \geq 5,9$ n'est observée dans le catalogue FCAT (Tableau 5.5). Or, d'après la Figure 5.17, quelques répliques $M_w \geq 5,9$ sont produites, ce qui amplifie cette surestimation.

La légère sous-estimation des catalogues synthétiques pour la période d'exhaustivité 1850 - 1964 s'explique par la présence du séisme $M_w 6,7$ (Ligurie, 1887) sur cette courte période de 115 ans alors que la période de retour d'un tel événement se situe autour de 238 ans d'après la DFM moyenne observée dans les catalogues synthétiques (Fig-

ure 5.19 et Tableau 5.9). Ainsi, même si quelques réalisations produisent des séismes de magnitude supérieure ou égale à 6,7 sur cette courte période, cela reste rare.

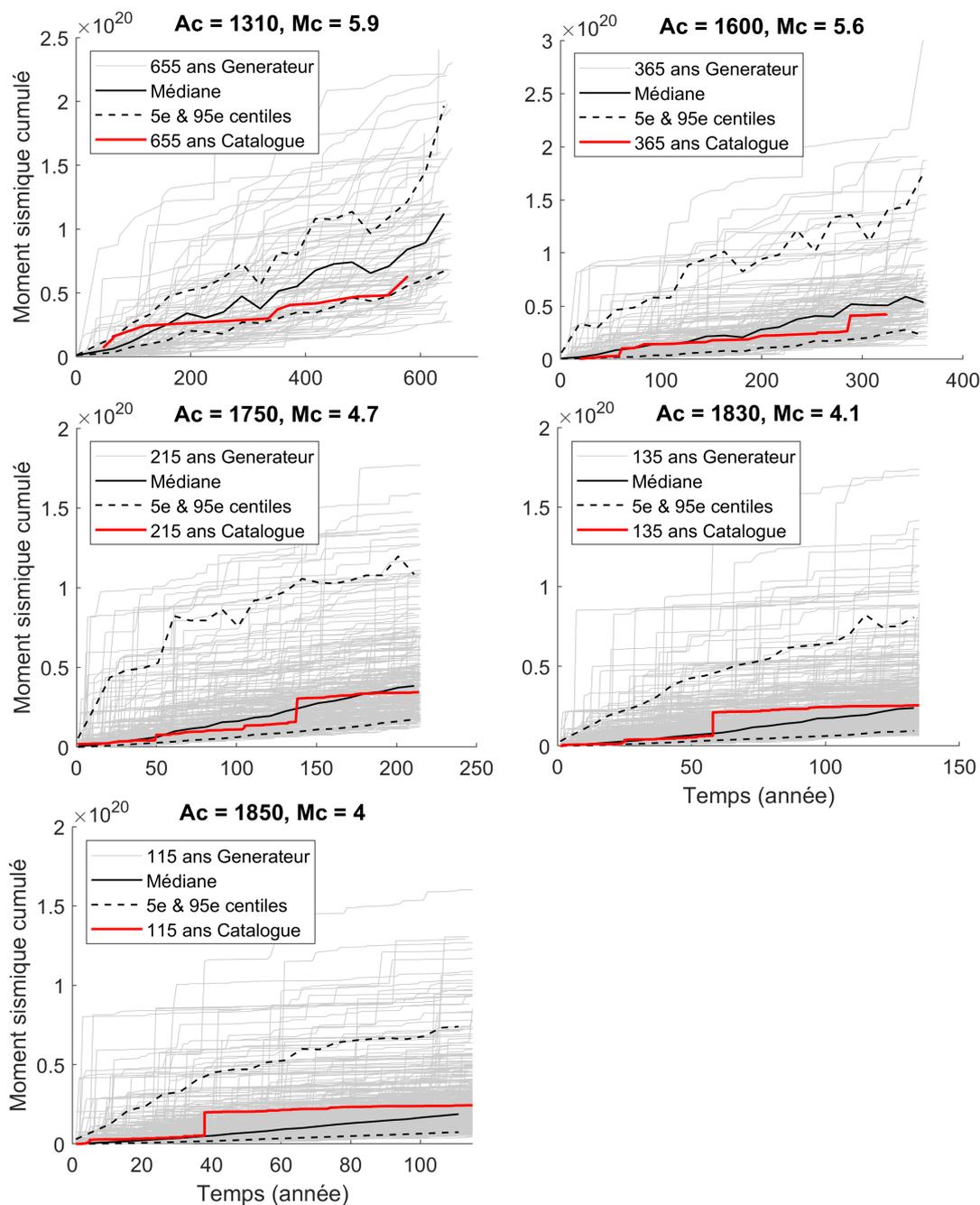


Figure 5.20: MOMENT SISMIQUE CUMULÉ PAR LES SISMICITÉS PASSÉE ET GÉNÉRÉE. Comparaison du moment sismique (M_0 , Équation 3.6) cumulé observé dans les catalogues synthétiques et dans le catalogue FCAT sur les cinq périodes d'exhaustivité (Tableau 5.1). Les courbes statistiques (médiane et centiles) sont calculées sur des fenêtres de 20 ans.

5.4.3 Répartition spatiale de la sismicité synthétique

La répartition spatiale des séismes principaux est réalisée à l'aide d'une densité spatiale et d'une régionalisation qui restreint la répartition spatiale des séismes principaux aux magnitudes les plus élevées (Chapitre 3). Bien que ces forts séismes soient sous-représentés vis-à-vis des séismes de magnitudes plus faibles, ils sont associés à des répliques localisées non pas à partir de la densité spatiale fournie en entrée mais autour du séisme principal. Aussi, ces répliques peuvent être nombreuses (Figure 5.18). Ainsi, la régionalisation a pour impact de concentrer la sismicité dans les régions considérées comme les plus actives. Il est donc attendu que la répartition spatiale de la sismicité synthétique ressemble à la densité spatiale fournie en entrée (Figure 5.3) sans pour autant lui être identique à cause de l'impact de la régionalisation. Il est également attendu que cet impact soit plus marqué par la régionalisation GEOTER qui délimite plus finement les régions actives des régions stables.

La Figure 5.21 compare la densité spatiale des tracés de failles (Figure 5.3b) donnée en entrée du générateur (a) avec les densités spatiales des sismicités synthétiques produites à l'aide des régionalisations GEOTER et SHARE (b). Ces dernières sont calculées en chaque cellule de 5×5 km en divisant le nombre de séismes synthétiques observés dans un rayon de 25 km autour de cette cellule par la surface de l'aire d'étude (outil "Densité de points" d'ArcGIS⁵).

Les densités spatiales obtenues sont proches de la densité spatiale donnée en entrée (Figure 5.21a). Pour autant, quelques différences sont notables entre les densités GEOTER et SHARE et avec la densité d'entrée.

Les régions GEOTER pyrénéennes n°31 et 35 (Figure 5.4), toutes deux limitées par une magnitude maximale égale à 7,1 (Tableau 5.2), sont sur-représentées par rapport aux autres régions pyrénéennes. En revanche, l'uniformité de la vaste région Pyrénées de SHARE produit une sismicité fidèle à la densité spatiale d'entrée.

De la même manière, la vaste région stable de SHARE est exposée à une sismicité plus fidèle à la densité spatiale d'entrée comparée à la discrétisation plus fine des régions GEOTER. Par exemple, la forte densité de tracés de failles observées dans les Cévennes et les Grands Causses (Figure 5.21a) est bien reproduite par la sismicité synthétique SHARE tandis qu'elle est moins exprimée dans la régionalisation GEOTER. Dans cette dernière, les magnitudes maximales des régions n°17, 18 et 27 sont faibles (entre 5,7 et 5,9 Tableau 5.2), faisant baisser le nombre de séismes produits dans ces régions.

Le constat inverse est observé dans les Alpes. La forte densité de tracés de failles est observée principalement dans les régions GEOTER n°21 et 23 et dans la région SHARE Alpes extensif (Figure 5.21a). Si cette densité est bien respectée par la régionalisation SHARE, en revanche cette forte densité est étendue à la région GEOTER n°24 caractérisée par la magnitude maximale autorisée la plus élevée : 7,3. Cette forte magnitude a donc permis de produire une quantité de séismes (principaux et répliques) plus élevée que ce que la densité spatiale fournie en entrée du générateur le laisse supposer.

⁵<https://desktop.arcgis.com/fr/arcmap/10.3/tools/spatial-analyst-toolbox/point-density.htm>

Au vu de ces observations, la régionalisation des magnitudes maximales autorisées se fait donc bien ressentir à l'échelle de l'ensemble de la sismicité synthétique $M_w \geq 4$.

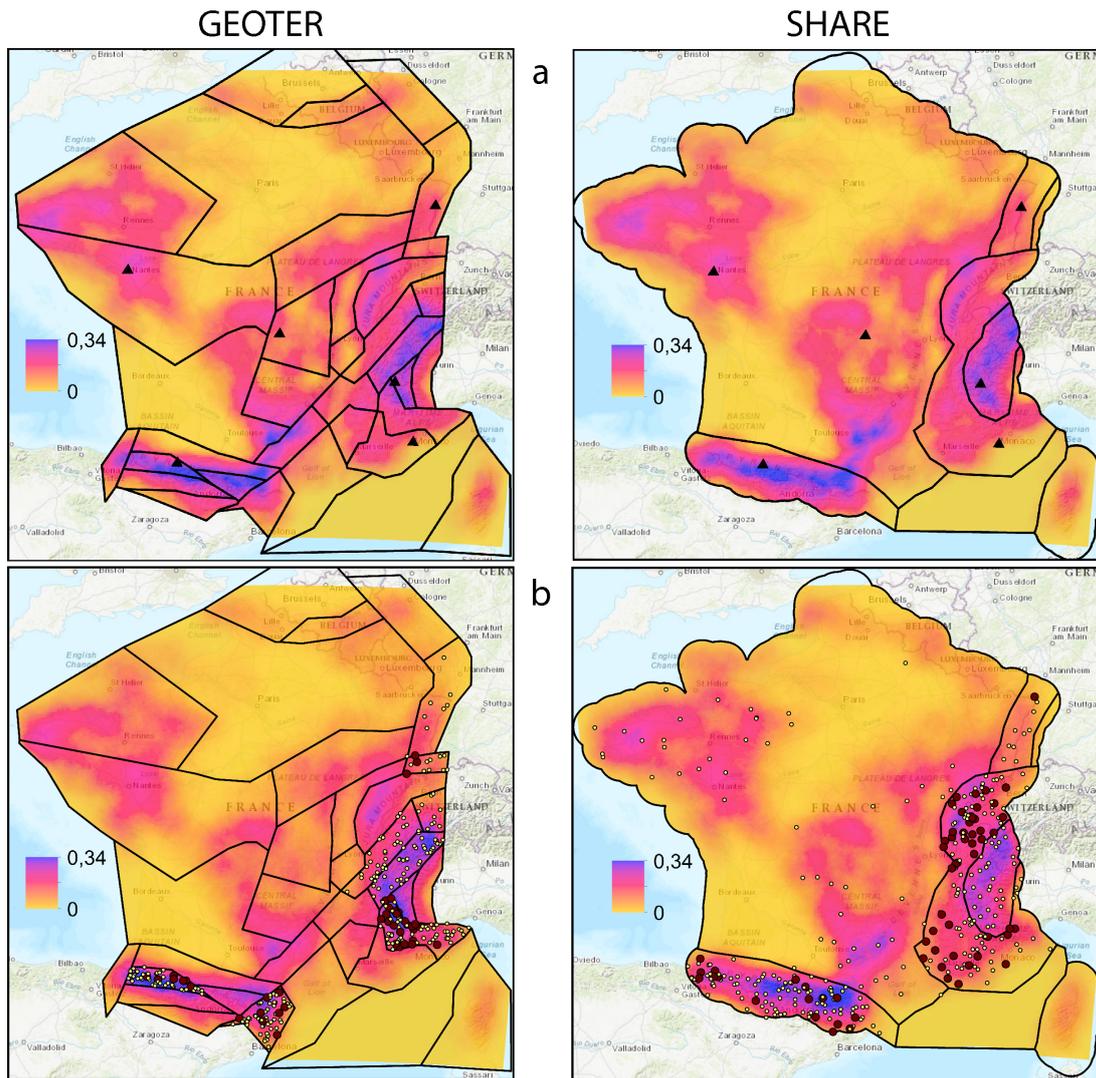


Figure 5.21: DENSITÉS SPATIALES EN ENTRÉE ET EN SORTIE DU GÉNÉRATEUR.
 a : Densité spatiale des tracés de failles fournie comme donnée d'entrée au générateur (Figure 5.3b). Les triangles noirs correspondent aux six villes étudiées Section 5.6.
 b : Densités de point obtenues à partir des séismes synthétiques $M_w \geq 4$ générés sur 50 000 ans. Les séismes $M_w \geq 6,5$ et $M_w \geq 7$ sont respectivement représentés en jaune et en marron.

5.5 Calcul de l'aléa sismique français

Cette Section a pour but d'appliquer aux données françaises la méthode de pondération présentée Chapitre 4. Les poids dépendant de la magnitude et de la distance sont d'abord présentés Section 5.5.1. L'aléa (PGA) provoqué en surface par chacun des séismes générés est calculé Section 5.5.2 à l'aide des cinq modèles de prédiction pondérés. Pour rappel, dans cette thèse les modèles de prédiction des mouvements du sol sont appelés par leur acronyme anglais couramment utilisé par la communauté : GMPE pour *Ground-Motion Prediction Equation*.

5.5.1 Pondération

Résultats Les données RESORCE de l'Europe de l'Ouest s'étendent de 4 à 6,9 en magnitude et de 0 à 200 km en distance Joyner - Boore (JB). Ces données sont utilisées dans le but de pondérer les cinq GMPE choisies Section 5.2.3 sur ces gammes de magnitudes et de distances. L'entièreté des données RESORCE est utilisée dans le processus de pondération pour les magnitudes comprises entre 7 et 7,5. Cet ajout n'apporte que quatre données supplémentaires, faisant passer le nombre de données à 489. La limite $M_w 6,9 - 7$ est matérialisée par une ligne noire dans les Figures ci-après.

La Figure 5.22 montre la distribution du $\log_{10}(\text{PGA})$ (m/s^2) calculé au rocher avec un régime décrochant par chacun des cinq GMPE pour chaque valeur de magnitude et de distance JB. Sont superposées à ces distributions les données RESORCE au rocher. Ces données manquent particulièrement entre $M_w 6,5$ et $6,8$ compris.

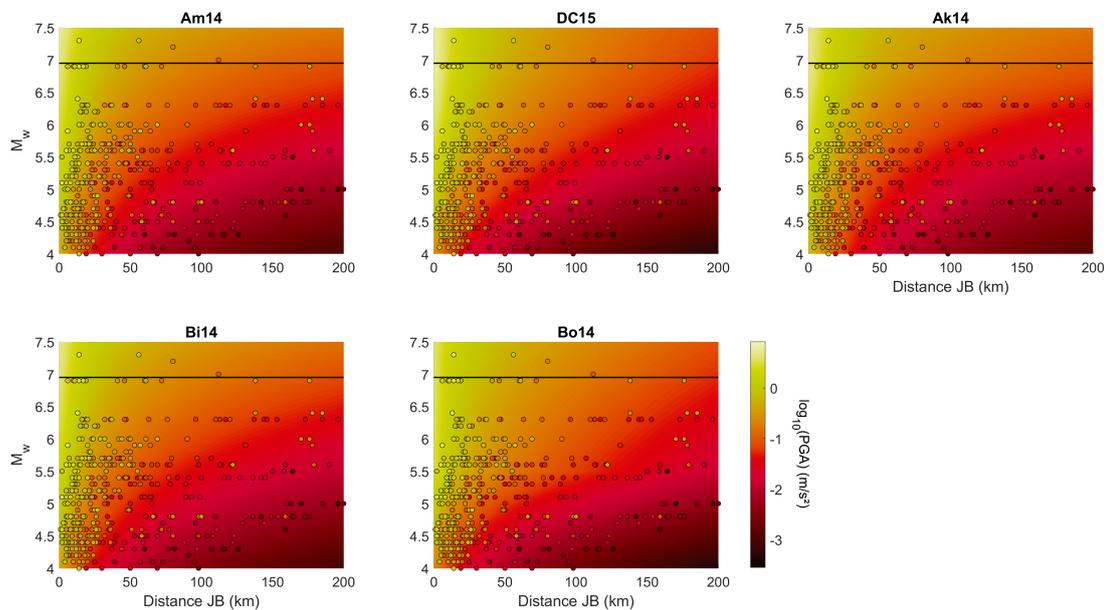


Figure 5.22: COMPARAISON DE PGA DES DONNÉES ET DES GMPE.

Chaque point est un enregistrement RESORCE, sa couleur correspondant au $\log_{10}(\text{PGA})$ (m/s^2) associé. La ligne noire correspond à la transition entre les données provenant de l'Europe de l'Ouest (en dessous) et celles provenant de l'ensemble de RESORCE (au dessus). Chaque GMPE est mentionnée par son abréviation.

Lors du calcul des poids en chaque point de données (Figure 5.23), les informations (magnitude, distance JB et régime de faille) contenues dans la base de données RE-SORCE sont utilisées afin de calculer les PGA à l'aide des GMPE. Le résultat de cette pondération est présenté Figure 5.23a. Une grande quantité des points de données est caractérisée par un poids binaire : soit le poids est nul, soit il est maximal (1). Seuls quelques points de données partagent plus équitablement le poids des cinq GMPE. De plus, il semblerait qu'aucun modèle ne soit laissé de côté : chacun prédomine au moins un domaine magnitude - distance.

En appliquant une interpolation du plus proche voisin à partir de ces poids ponctuels, un modèle de poids est obtenu. La Figure 5.23b présente le modèle de poids associé à chaque GMPE pour un pas de 0,1 en magnitude et des distances JB au kilomètre près.

Les GMPE de Drouet et Cotton [2015] (DC15), d'Ameri [2014] (Am14), d'Akkar et al. [2014a] (Ak14) et de Boore et al. [2014] (Bo14) prédominent la quasi-totalité (94%) de la distribution. Chacun de ces GMPE pèsent respectivement 29%, 23%, 22% et 20% sur l'ensemble du domaine magnitude - distance présenté Figure 5.23b. La GMPE de Bindi et al. [2014] contrôle, quant à elle, seulement 6% de l'ensemble domaine magnitude - distance, notamment pour des magnitudes allant de $M_w 7$ à $M_w 7,5$ et des distances JB comprises entre 70 et 100 km.

Même si la prédominance du calcul de PGA semble bien partagée entre les GMPE DC15, Am14, Bo14 et Ak14, leurs domaines magnitude - distance de prédilection respectifs diffèrent. Les GMPE Am14 et Bo14 prédominent principalement sur le domaine central compris entre les magnitudes 4,6 et 6,8 et pour les distances JB allant de 25 et 150 km. Les domaines caractérisés par des valeurs extrêmes de magnitudes ($<4,6$ ou $>6,8$) et de distances JB (<25 km et >150 km) sont de manière générale contrôlés par les trois autres GMPE (DC15, Ak14 et Bi14).

Discussion Martin et al. [2018] utilisent ces cinq GMPE en plus de celle de Cauzzi et al. [2015] (C15, Section 5.2.3) pour une application dans le Sud-Est de la France. À titre de comparaison, ces auteurs fixent, à dire d'experts, les poids suivants : 0,25 pour DC15 et Am14, 0,2 pour C15, 0,15 pour Bo14 et enfin 0,075 pour Bi14 et Ak14. En omettant C15, ces poids deviennent environ 0,31 (DC15 et Am14), 0,19 (Bo14) et 0,09 (Ak14 et Bi14). Hormis Bi14 qui est moins bien représentée, cette hiérarchisation des GMPE est similaire à celle trouvée par notre modèle.

Les GMPE de DC15 et d'Am14 ont été établies pour une application spécifique à la France métropolitaine. Il est donc cohérent de les voir prédominer plus de la moitié (52%) de la distribution des poids. La GMPE européenne de Bi14 a finalement une contribution faible (6%) dans le modèle de poids illustré Figure 5.23b. La GMPE de Bo14 est formulée à l'aide de données de régions actives de par le monde, ne semblant pas de prime abord représenter au mieux la sismicité française. Pourtant, cette GMPE gouverne un cinquième du domaine distance - magnitude étudié Figure 5.23b.

Les formulations des GMPE d'Ak14 et de Bi14 ont nécessités les données de RE-SORCE. Cette remarque soulève un problème [par ex., Mak et al., 2017] : les données utilisées pour tester des GMPE doivent être indépendantes des données utilisées pour

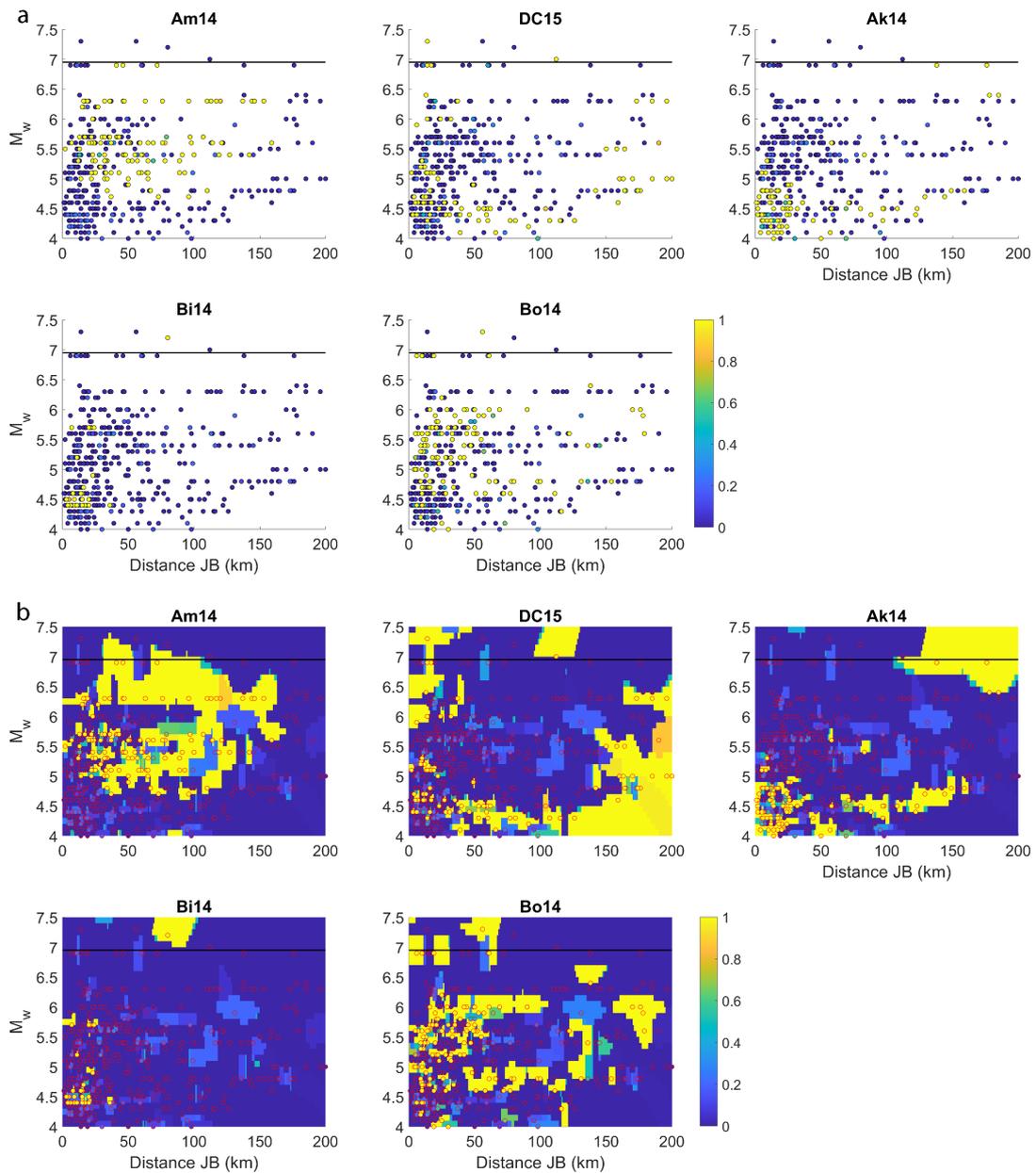


Figure 5.23: POIDS DES GMPE.

a : Chaque point est un enregistrement de RESORCE (point) associé à un poids.
 b : À partir de ces poids ponctuels, une interpolation du plus proche voisin est réalisée afin d'obtenir un poids pour tous les couples magnitude - distance. La ligne correspond à la transition entre les données provenant de l'Europe de l'Ouest (en dessous) et celles provenant de l'ensemble de RESORCE (au dessus).
 Chaque GMPE est mentionnée par son abréviation (voir texte).

fomuler ces GMPE. Tester une GMPE à l'aide de données ayant servi à la formuler ne permet pas d'évaluer l'aspect prédictif de cette GMPE, seulement son aspect explicatif. Néanmoins, Mak et al. [2017] stipulent que tout n'est pas à mettre de côté lorsqu'une analyse comparative de la sorte est réalisée. En effet, ils prennent l'exemple d'une GMPE indépendante des données utilisées pour le test qui aboutit à un poids plus élevée que celui d'une GMPE dépendante de ces données. Dans ce cas, le test constitue un argument fort pour la bonne performance de la première GMPE. Ainsi, nous pouvons conclure que la GMPE de Bo14 est plus performante que celle de Bi14. En utilisant également la base de données RESORCE pour comparer neuf GMPE à l'aide d'une approche Bayésienne, Bertin et al. [2019] estiment que les GMPE de Bi14 et Ak14 sont trois des plus performantes. Dans ce cas de figure, il est *a priori* impossible de conclure car ces deux GMPE ont été établies à l'aide des données RESORCE. Dans le cas de nos résultats, ces deux GMPE n'arrivent pas en tête du classement. Ils contrôlent en effet moins d'un tiers (28%) du domaine distance - magnitude.

5.5.2 Calcul de la PGA

PGA médiane Pour rappel, dans cette thèse aucune distinction n'est faite entre les incertitudes aléatoires et épistémiques. De ce fait, les incertitudes épistémiques liées au choix de la GMPE sont traitées comme des incertitudes aléatoires. L'aléa est donc calculé en faisant la combinaison linéaire (Équation 4.4) des PGA calculées par chaque GMPE (Figure 5.22) multipliées par leurs poids (Figure 5.23b). L'aléa final médian obtenu à partir des GMPE pondérées est présenté Figure 5.24a. Cette distribution de l'aléa calculé à partir des cinq GMPE n'est pas lisse. Or, le passage de dominance d'une GMPE à une autre est généralement à fort gradient : les poids passent rapidement de 1 à 0 ou inversement (Figure 5.23). Ainsi, le passage d'une GMPE à une autre pour estimer l'aléa entre deux pas de distance ou de magnitude peut entraîner une différence non négligeable sur la PGA calculée. De plus, la méthode d'interpolation choisie (plus proche voisin) ne participe pas non plus au lissage de cette distribution.

Variabilité La Figure 5.24b présente, quant à elle, les distributions en magnitude et distance JB de la PGA pondérée médiane et de la PGA pondérée plus ou moins 3σ . Cette large enveloppe prend en compte la variabilité propre à chacun des cinq GMPE et englobe la quasi-totalité des données RESORCE. Sur les 489 enregistrements, seuls dix sont en dehors de cette enveloppe 3σ (Figure 5.24b), soit 2% des données.

Cette variabilité (médiane $\pm\sigma$) de la distribution de la PGA est également présentée Figure 5.25 pour différentes gammes de distances JB et de magnitudes. Le constat d'une distribution non lisse observée Figure 5.24a est de nouveau visible sur cette Figure. Par exemple, la distribution de l'aléa produit par un séisme $M_w 4$ est caractérisée par une diminution marquée entre les distances 20 et 40 km. La distribution de l'aléa à 0 km augmente, quant à elle, pour les faibles magnitudes (4- 4,6).

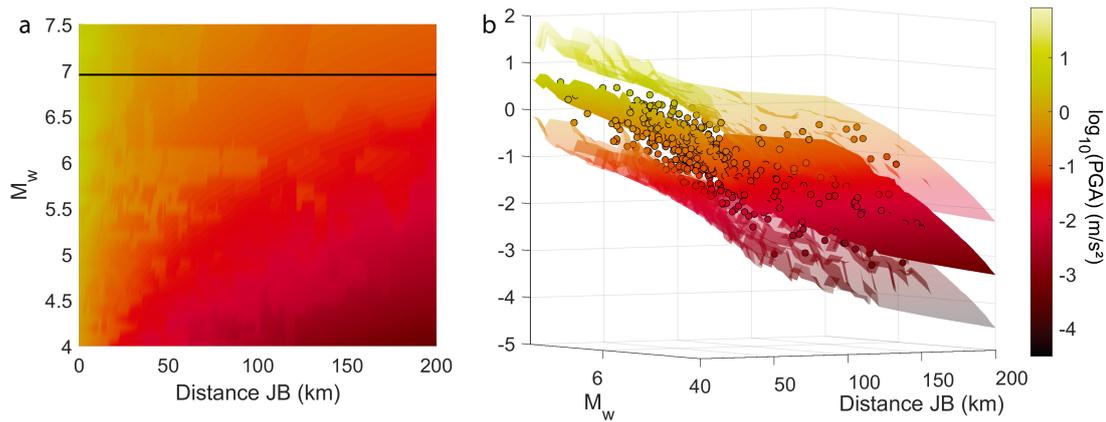


Figure 5.24: DISTRIBUTIONS DE PGA OBTENUES PAR LES GMPE PONDÉRÉES. Distribution du \log_{10} de PGA en fonction de la magnitude et de la distance. La PGA est calculé à l'aide des PGA obtenues par les cinq GMPE et les poids associés à chacun de ces modèles. Le régime de faille décrochant est considéré pour ce calcul.
 a : Distribution médiane de PGA pondérée. La ligne correspond à la transition entre les données provenant de l'Europe de l'Ouest (en dessous) et celles provenant de l'ensemble de RESORCE (au-dessus).
 b : Distributions médiane et plus ou moins 3σ de PGA pondérée. Comme pour la Figure 5.22, chaque point représente un enregistrement RESORCE.

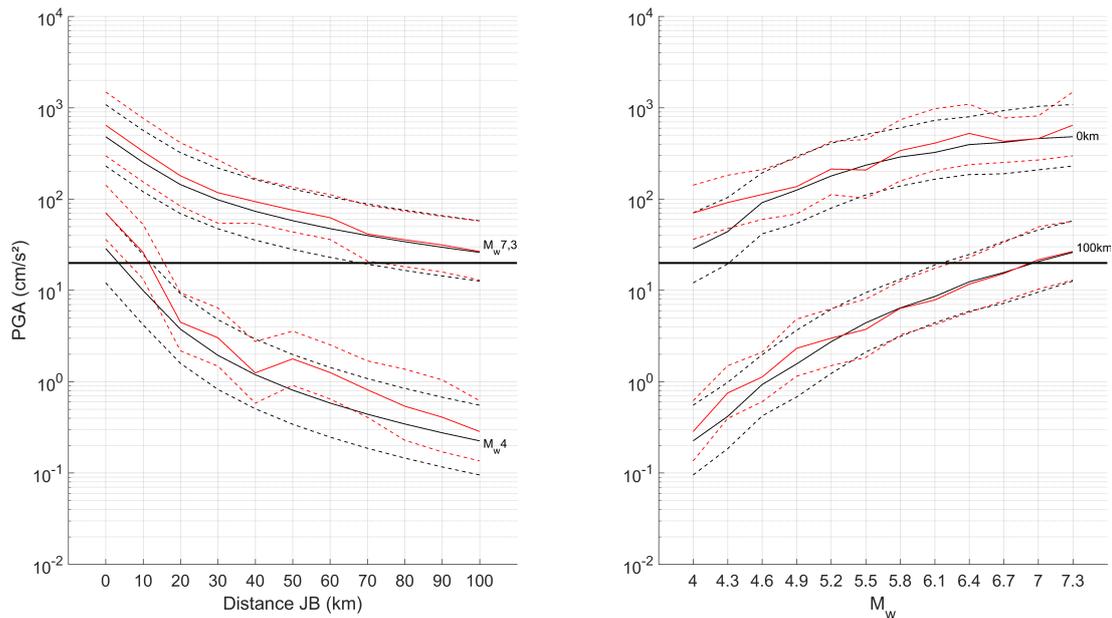


Figure 5.25: VARIABILITÉ DE LA PGA EN FONCTION DE LA MAGNITUDE ET DE LA DISTANCE.

La PGA médiane et sa variabilité à un σ obtenues à l'aide de notre pondération (rouge Figure 5.24) et de celle proposée par Martin et al. [2018] (noir) sont représentées pour différentes valeurs de magnitude et de distance JB. La ligne noire correspond à la valeur 20 cm/s^2 , soit le seuil en dessous duquel l'aléa est considéré comme négligeable dans cette thèse.

Comparaison avec le modèle de pondération de Martin et al. [2018] Notre modèle de pondération est également comparé Figure 5.25 à celui proposé par Martin et al. [2018] (M2018). D’après cette Figure, l’aléa obtenu via notre pondération semble, d’une manière générale, être supérieur à celui obtenu par la pondération de M2018.

La Figure 5.26 présente l’écart normalisé En entre les PGA calculées à partir de la pondération M2018 PGA_{M2018} et celle proposée dans cette thèse (PGA) tel que :

$$En = \frac{PGA_{M2018} - PGA}{PGA}. \quad (5.2)$$

Cet écart normalisé est calculé uniquement pour les couples magnitude-distance produisant un aléa non négligeable comme défini Section 4.2.2 ($PGA \geq 20 \text{ cm/s}^2$). Ainsi, seul le domaine dont la distance est inférieure ou égale à la distance d’emprise définie par l’Équation 4.3 (courbe noire Figure 5.26) est analysé.

Sur l’ensemble de ce domaine, En vaut en moyenne à $-10 \pm 14\%$. Notre modèle de pondération aboutit donc en moyenne à une distribution de PGA plus élevée que celle obtenue avec la pondération M2018. 36, 47 et 17% du domaine visible Figure 5.26 est caractérisé par un En absolu respectivement inférieur ou égale à 10%, compris entre 10 et 25% et supérieur à 25%. L’écart normalisé minimal vaut -61% et est observé Figure 5.26 à faible distance ($< 10 \text{ km}$) et faible magnitude ($\leq 4,4$). Dans ce domaine bien particulier, c’est la GMPE Ak14 qui contrôle le calcul de la PGA (Figure 5.23b). Or, cette GMPE propose les PGA les plus élevées pour ces gammes de distances et de magnitudes. Ainsi, il est cohérent d’observer des En négatifs et conséquents dans ce domaine.

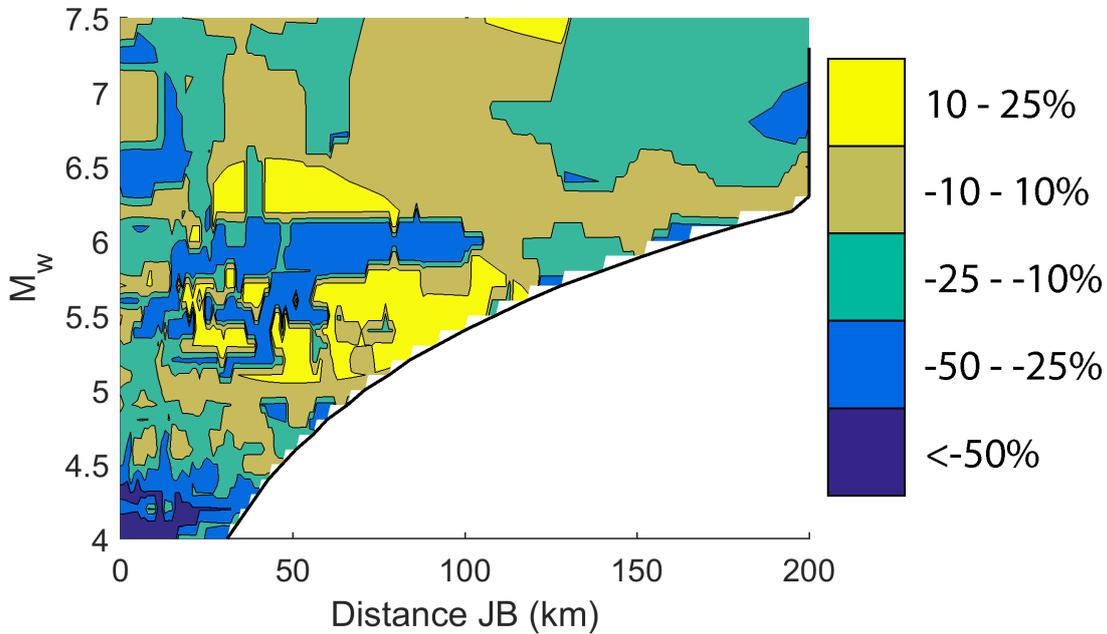


Figure 5.26: COMPARAISON DE DISTRIBUTIONS MÉDIANES DE PGA.

Pour chaque magnitude à 0,1 près et chaque distance JB au kilomètre près, un écart normalisé En (Équation 5.2) entre les PGA calculées via la pondération M2018 et celle proposée dans cette thèse est calculé. Ce calcul est réalisé uniquement aux distances inférieures ou égales à la distance d’emprise (courbe noire, Équation 4.3).

Bien que la tendance générale soit à une surestimation vis-à-vis des PGA de M2018, certains domaines sont caractérisés par un En positif, i.e., une sous-estimation, pouvant aller jusqu'à 24%. C'est le cas notamment de trois domaines. Le premier se situe autour de M_w 5,5 pour des distances comprises entre 20 et 120 km. Le deuxième est compris entre M_w 6,2 et 6,5 pour des distances allant de 25 à 80 km. Dans ces deux domaines, la GMPE Am14 prédomine le calcul de la PGA (Figure 5.23b). Or, cette GMPE propose les PGA les plus faibles pour ces gammes de distances et de magnitudes. Ainsi, il est cohérent d'observer des En positifs et conséquents dans ces domaines. Le troisième domaine, localisé à de fortes distances (105 - 130 km) et magnitudes ($\geq 7,3$), est le domaine de prédilection de la GMPE de DC15. Encore une fois, dans ce strict domaine, cette GMPE produit les plus faibles PGA ce qui tend à augmenter En .

5.6 Estimation de l'aléa sismique en France

Cette Section présente les courbes probabilité annuelle-aléa obtenues à partir des séismes synthétiques présentés Section 5.4 et de l'ensemble des GMPE pondérées détaillé Section 5.5. Le catalogue synthétique dans cette Section a une durée d'un million d'années. L'aléa est calculé France entière à une résolution spatiale de 250 m. Le territoire métropolitain est ainsi décrit par plus de neuf millions de mailles. Afin d'illustrer les résultats, une sélection de six villes situées dans des contextes différents a été réalisée : Nantes, Lourdes, Clermont-Ferrand, Briançon, Nice et Strasbourg. Les valeurs d'aléa pour ces villes sont estimés sur la base d'une emprise de 10×10 km et non 250×250 m qui est infra-communale.

5.6.1 Courbes probabilité annuelle-aléa

5.6.1.1 Courbes moyennes

Résultats Pour les deux régionalisations utilisées, les courbes moyennes sont calculées à partir d'un catalogue synthétique d'un million d'années. La probabilité minimale qu'il est possible d'atteindre est donc égale à 10^{-6} . Néanmoins, les probabilités inférieures à 10^{-4} ne sont pas exploitables car trop peu représentées : moins de 100 réalisations. La Figure 5.27 montre les courbes moyennes probabilité annuelle-aléa obtenues pour ces six villes à l'aide des régionalisations GEOTER (a) et SHARE (b).

Les courbes GEOTER sont plus dispersées que celles de SHARE. En effet, la régionalisation SHARE a tendance à centrer toutes les courbes et de les rassembler en trois groupes : Briançon, Lourdes et Nantes pour les plus élevées, Nice et Strasbourg en intermédiaire et Clermont-Ferrand pour la plus faible. De leur côté, les courbes GEOTER sont plus distantes les unes des autres, ne permettant pas de former des groupes (hormis pour Briançon et Lourdes). Ce constat s'explique par l'aspect plus diffus de la régionalisation SHARE. Par exemple, d'après le Tableau 5.3, des séismes de magnitude 6,5 peuvent être générés sur l'entièreté du territoire tandis que seulement neuf sur les 35 régions GEOTER (Tableau 5.2) peuvent accueillir un tel séisme. Cela fait que la concentration de l'aléa dans les régions les plus actives est plus marquée avec la régionalisation GEOTER, comme l'a déjà illustré la Figure 5.21.

Les valeurs moyennes de PGA à la période de retour 475 ans obtenues par les deux régionalisations sont répertoriées Tableau 5.10 et observables Figure 5.28.

	Clermont	Strasbourg	Nice	Nantes	Briançon	Lourdes
GEOTER	56	67	88	119	147	153
SHARE	66	79	73	110	118	123

Table 5.10: PGA MOYENNES À PÉRIODE DE RETOUR DE 475 ANS.

PGA (cm/s^2) moyennes calculées pour les six villes d'études à l'aide des régionalisations GEOTER et SHARE. Les courbes moyennes probabilité annuelles-aléa associées à ces villes sont visibles Figure 5.27.

De manière générale nos résultats sont plus faibles que l'aléa réglementaire actuel⁶. Martin et al. [2018] constatent que depuis le début du XXI^{ème} siècle les études probabilités [Martin et al., 2002, Martin et Secanell, 2006] tendent vers une diminution de l'estimation de l'aléa en France. Nos résultats suivent la même tendance.

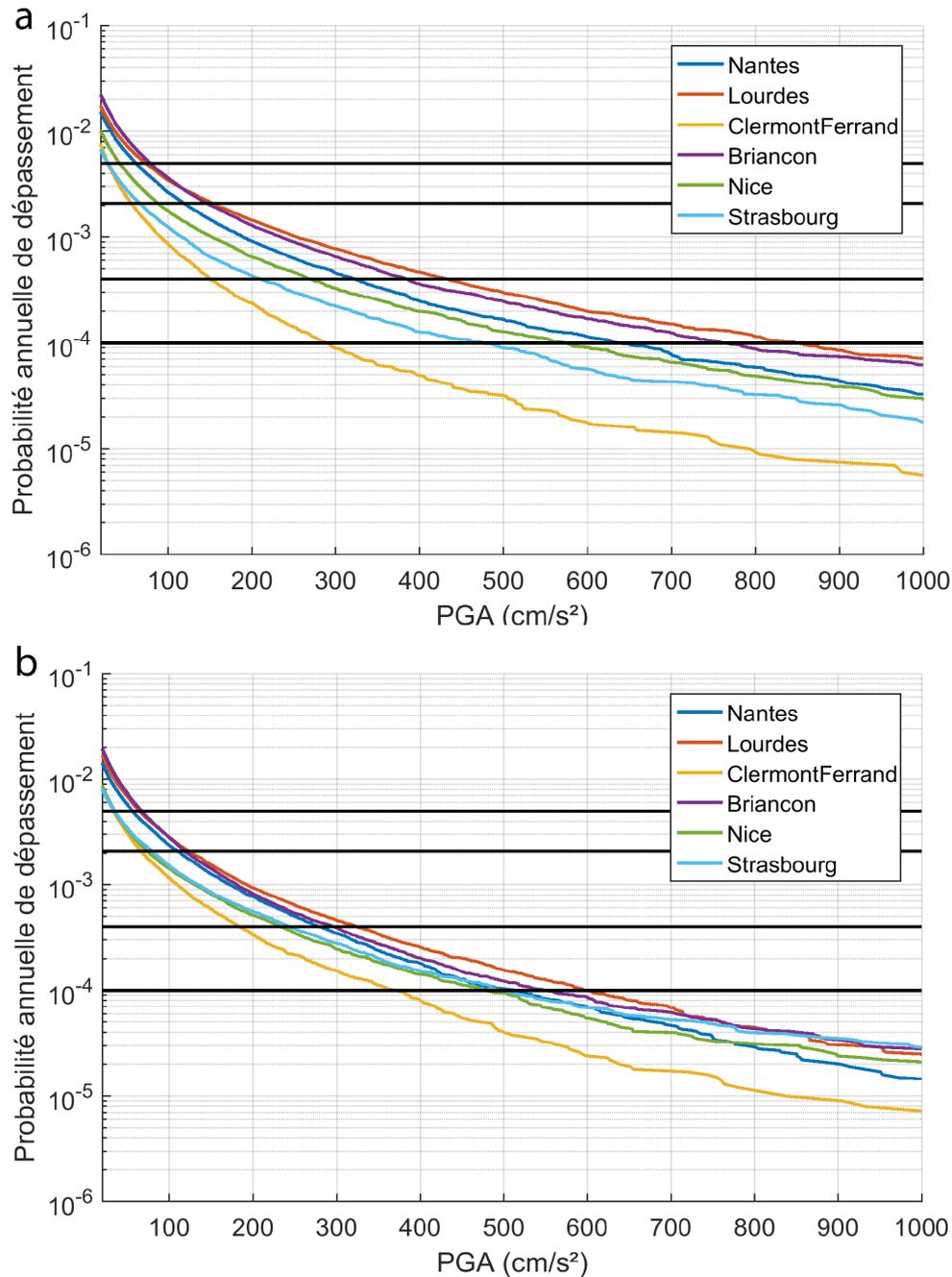


Figure 5.27: COURBES MOYENNES PROBABILITÉ ANNUELLE - ALÉA.

Courbes moyennes obtenues à partir des régionalisations GEOTER (a) et SHARE (b) pour les six villes d'étude.

Les lignes noirs représentent les probabilités annuelles correspondant aux périodes de retour 200, 475, 2 475 et 10 000 ans.

⁶<http://www.planseisme.fr/spip.php?page=carto>

Comparaison avec d'autres études L'aléa à période de retour de 475 ans pour certaines de ces villes a récemment été estimé par Martin et al. [2018], Beauval et al. [2020] et Drouet et al. [2020]. Ces derniers trouvent une valeur médiane de 72 cm/s^2 pour Nice et de 151 cm/s^2 pour Lourdes. Nos PGA médianes (Tableau 5.11) sont très proches de leurs résultats avec respectivement 90 et 159 cm/s^2 d'après les régions GEOTER et respectivement 77 et 125 cm/s^2 d'après les régions SHARE.

Martin et al. [2018] (M2018) estiment la moyenne à environ 44 cm/s^2 pour Nice et 60 cm/s^2 pour un site (OGAG) situé à 12 km de Briançon. Ces résultats représentent environ la moitié de ceux que nous trouvons Tableau 5.10. Les méthodologies employées dans notre étude et dans celle de M2018 sont différentes, ainsi plusieurs raisons peuvent expliquer un tel écart. La plus évidente réside dans le modèle de pondération des GMPE (Section 5.5.1). En effet, à partir des mêmes GMPE (exceptée celle de Cauzzi et al. [2015]), notre modèle de pondération aboutit en général à un aléa plus élevé que celui proposé par le modèle de pondération de M2018 (Figures 5.25 et 5.26). Il est donc cohérent d'observer des résultats de PGA à période de retour de 475 ans plus conséquents.

Les PGA moyennes obtenues dans cette thèse pour les six villes d'étude sont comparées Figure 5.28 avec celles obtenues par Beauval et al. [2020]. Ces résultats moyens sont comparables pour les villes de Clermont-Ferrand, Briançon, Nice et Strasbourg. En revanche, nos PGA moyennes à période de retour de 475 ans calculées pour les villes de Nantes (110 et 119 cm/s^2) et Lourdes (123 et 153 cm/s^2) sont respectivement plus forte et plus faible que celles obtenues par ces auteurs : environ 55 et 300 cm/s^2 respectivement.

En ce qui concerne les zones frontalières, la carte d'aléa sismique d'Italie proposée par l'INGV⁷ estime une PGA médiane à 475 ans comprise entre 98 et 122 cm/s^2 à la frontière proche de Briançon. Ces valeurs sont donc en accord avec une diminution de l'aléa réglementaire actuel. Pour Strasbourg, la comparaison avec la carte d'aléa sismique allemande proposée par le GFZ⁸ [Grünthal et al., 2018] aboutit à la même conclusion. La PGA moyenne à la période de retour 475 ans y est de 70 à 80 cm/s^2 .

À l'échelle européenne, une étude PSHA a été réalisée par Woessner et al. [2015] pour le compte du projet ESHM13⁹. Seul le résultat moyen de la ville de Briançon (environ 142 cm/s^2) est comparable avec les nôtres (118 et 147 cm/s^2 selon régions SHARE et GEOTER). La ville de Nantes est la seule à être caractérisée par une PGA moyenne plus faible : ≈ 49 contre 110 et 119 cm/s^2 dans cette étude. Toutes les autres villes sont caractérisées par une PGA bien plus élevée que nos résultats, étant parfois plus du double. Martin et al. [2018] évoquent plusieurs raisons pouvant expliquer cette surestimation de l'étude européenne vis-à-vis des études nationales.

⁷http://esse1-gis.mi.ingv.it/mps04_eng.jsp

⁸<http://www-app5.gfz-potsdam.de/d-eqhaz16/index.html>

⁹<http://www.efehr.org/en/hazard-data-access/hazard-curves/>

5.6.1.2 Centiles de PGA

Le catalogue synthétique d'un million d'années est divisé en 100 sous-catalogues d'une durée de 10 000 ans chacun. Le calcul des centiles est donc réalisé uniquement pour des périodes de retour inférieures ou égales à 10 000 ans. Dans cette thèse, les deux périodes de retour étudiées sont celles de 475 ans et 2 475 ans. Elles représentent respectivement une probabilité de dépassement de 10% et 2% sur une période d'observation de 50 ans. Les probabilités annuelles de ces deux périodes de retour sont respectivement 0,21% et 0,04%. Pour calculer les centiles, nous appliquons la méthode présentée Section 1.3.3.

Résultats Les centiles des PGA aux périodes de retour étudiées sont visibles Figure 5.28 pour les six villes d'étude. Ces centiles sont calculés à partir des régionalisations GEOTER et SHARE et sont comparés avec les moyennes décrites Section 5.6.1.1 (triangles). Les moyennes sont toujours inférieures ou égales aux médianes. Ce décalage est plus marqué pour la période de retour 2 475 ans.

Le Tableau 5.11 répertorie les PGA médianes (Q50) et les 16^{ème} (Q16) et 84^{ème} (Q84) centiles calculés aux périodes de retour 475 et 2 475 ans pour les six villes d'étude. La même observation que pour les valeurs moyennes peut être faite : la régionalisation SHARE tend à centrer les valeurs de PGA de ces six villes tandis que la régionalisation GEOTER les disperse. En effet, les valeurs Q16, Q50 et Q84 (Tableau 5.11) des villes les plus exposées (Lourdes, Briançon, Nantes et Nice) sont plus faibles avec la régionalisation SHARE. À l'inverse, cette même régionalisation tend à augmenter les valeurs Q16, Q50 et Q84 des villes les moins exposées (Clermont-Ferrand et Strasbourg), et ce quelle que soit la période de retour.

Régions	Centile	Clermont	Strasbourg	Nice	Nantes	Briançon	Lourdes
475 ans							
GEOTER	Q16	49	57	77	102	126	131
	Q50	58	67	90	120	148	159
	Q84	66	83	107	143	170	185
SHARE	Q16	59	65	65	97	108	109
	Q50	68	79	77	113	121	125
	Q84	80	95	89	129	139	141
2 475 ans							
GEOTER	Q16	126	165	224	266	311	339
	Q50	165	229	284	338	412	449
	Q84	208	288	380	440	553	580
SHARE	Q16	155	188	193	230	240	268
	Q50	191	259	247	295	310	333
	Q84	242	339	324	366	387	440

Table 5.11: CENTILES DES PGA AUX PÉRIODES DE RETOUR 475 ET 2 475 ANS. PGA (cm/s^2) médianes (en gras), 16^{ème} et 84^{ème} centiles calculées pour les six villes d'études à l'aide des régionalisations GEOTER et SHARE pour les périodes de retour 475 et 2 475 ans.

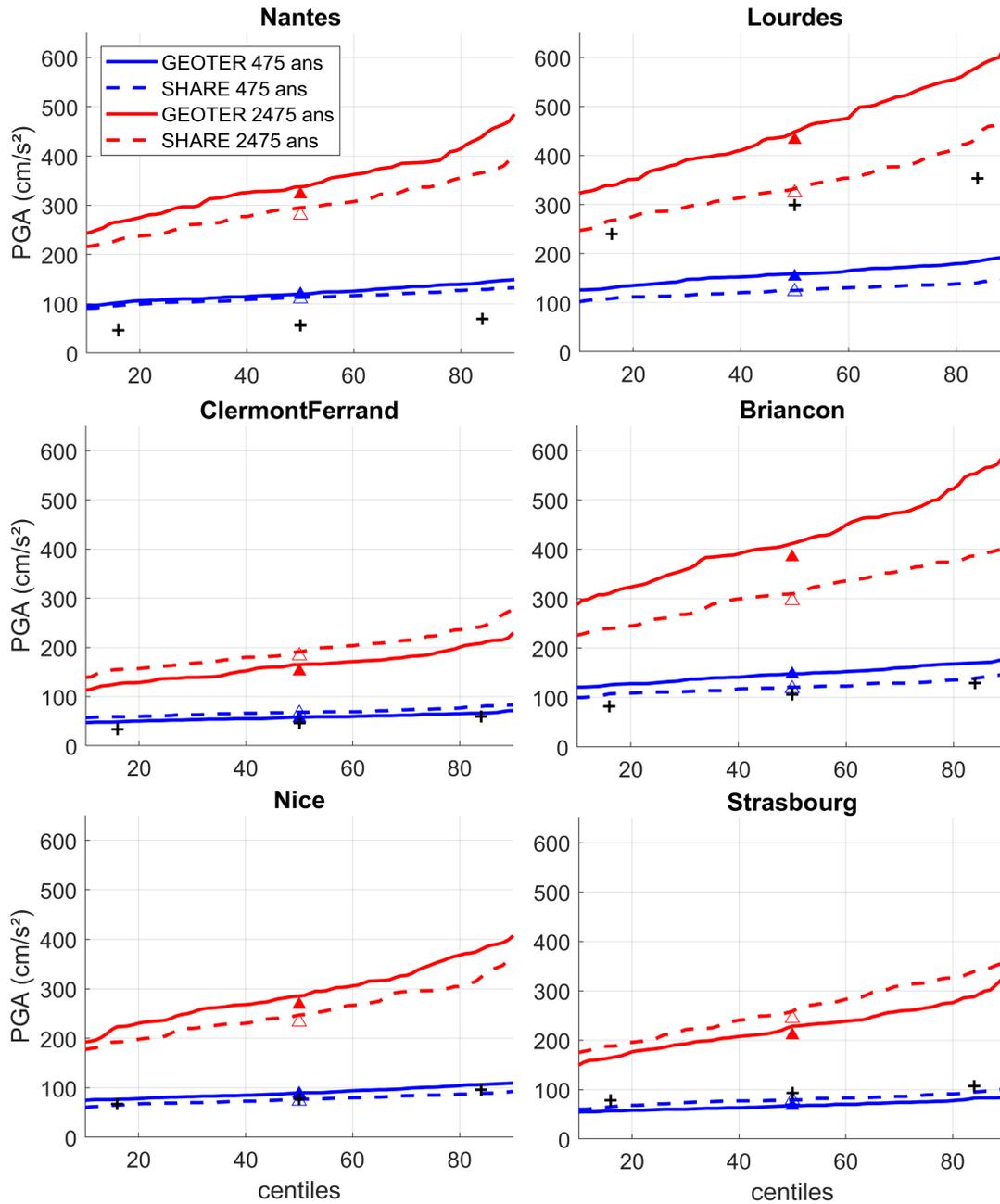


Figure 5.28: CENTILES DES PGA AUX PÉRIODES DE RETOUR 475 ET 2 475 ANS. Centiles des PGA aux périodes de retour 475 (bleu) et 2 475 ans (rouge) obtenus à partir des régionalisations GEOTER (trait plein) et SHARE (tiret). Les triangles représentent les moyennes de PGA à période de retour 475 (bleu) et 2 475 ans obtenus à partir des régionalisations GEOTER (plein) et SHARE (vide) et extraites de la Figure 5.27. Les croix noirs représentent les valeurs moyennes et les 16^{ème} et 84^{ème} centiles obtenues par Beauval et al. [2020].

Comparaison avec l'étude de Beauval et al. [2020] Ces résultats d'aléa de période de retour de 475 ans ont été comparés Figure 5.28 aux moyennes et aux 16^{ème} et 84^{ème} centiles obtenus par Beauval et al. [2020] (B2020). Pour les villes de Clermont-Ferrand et de Nice, les résultats sont similaires. Pour les villes de Briançon et de

Strasbourg, nos résultats SHARE sont comparables avec ceux de B2020, malgré une légère sous-estimation et surestimation respectivement. Pour les deux dernières villes, les résultats sont en revanche bien différents.

En ce qui concerne Lourdes, les résultats de B2020 représentent le double de ce que nous obtenons. Cette ville se trouve au coeur des Pyrénées qui est limité par une magnitude maximale élevée (7- 7,1 Tableaux 5.2 et 5.3). De plus, la densité spatiale dans cette région est également forte (Figure 5.3). Ce constat est donc probablement dû au calcul d'une seule DFM pour la France entière qui aurait tendance à moyenniser l'aléa final dans l'ensemble des régions. Cela conduirait une surestimation de l'aléa dans les régions les moins actives et une sous-estimation dans les régions les plus actives.

En ce qui concerne Nantes, nos résultats sont plus élevés que ceux de B2020. Ce constat fait écho à ce qui a été observé entre nos résultats et ceux du projet ESHM13 [Woessner et al., 2015] ($\approx 49 \text{ cm/s}^2$). Les $M_{max \text{ Région}}$ élevées (6,4 pour GEOTER et 6,5 pour SHARE), conditionnées par le séisme vendéen $M_w 6,3$ de 1799, et la forte densité spatiale observée autour de Nantes participent à la génération de séismes synthétiques de magnitude élevée. Cela entraîne une estimation plus forte de l'aléa pour cette ville.

Incertitudes La métrique *Inc* [Giardini et al., 2004, Douglas et al., 2014] permet de calculer l'incertitude associée à la PGA d'une période de retour donnée via cette expression :

$$Inc = 100 \times \log_{10}(PGA_{Q84}/PGA_{Q16}), \quad (5.3)$$

avec PGA_{Q16} et PGA_{Q84} les 16^{ème} et 84^{ème} centiles de PGA de la période de retour donnée (Tableau 5.11). Le Tableau 5.12 renseigne ces incertitudes calculées pour les six villes d'étude pour les deux périodes de retour analysées dans cette thèse. Ces incertitudes sont comparés avec celles du projet ESHM13 [Woessner et al., 2015]. Ces dernières sont estimées à partir des courbes des 15^{ème} et 85^{ème} centiles disponibles sur internet¹⁰. Les incertitudes déduites de la Figure 14 de l'étude de B2020 sont également reportées dans le Tableau 5.12.

Plusieurs hypothèses peuvent être formulées concernant le comportement des incertitudes de l'aléa [par ex., Douglas et al., 2014] :

- (1) Plus la période de retour est grande, plus l'incertitude devrait augmenter.
- (2) Plus le territoire d'étude est actif, mieux il est connu et donc moins les incertitudes devraient être grandes.
- (3) Plus le territoire d'étude est grand, plus les incertitudes devraient l'être.
- (4) Plus la rigueur est élevée (prise en compte de nombreuses variabilités), plus les incertitudes devraient croître.

Nos résultats sont analysés au regard de ces quatre hypothèses.

Tout d'abord, l'incertitude augmente avec la période de retour comme l'atteste le Tableau 5.12. Ceci est en cohérence avec l'hypothèse (1). Ce n'est en revanche pas le cas pour l'étude ESHM13 [voir Douglas et al., 2014, pour plusieurs raisons possibles à ce constat].

¹⁰<http://www.efehr.org/en/hazard-data-access/hazard-curves/>

		Clermont	Strasbourg	Nice	Nantes	Briançon	Lourdes
Régions	PR	Cette étude					
GEOTER	475	13	16	14	15	13	15
	2 475	22	24	23	22	25	23
SHARE	475	13	16	13	13	11	11
	2 475	19	26	23	20	21	22
Étude	PR	D'autres études					
ESHM13	475	49	34	35	32	20	23
	2 475	40	33	29	26	16	12
B2020	475	24	14	16	18	20	17

Table 5.12: INCERTITUDE ASSOCIÉE AUX PGA AUX PÉRIODES DE RETOUR ÉTUDIÉES. Incertitude Inc (Équation 5.3) de la PGA (cm/s^2) aux périodes de retour (PR) 475 et 2 475 ans pour les régionalisations GEOTER et SHARE. Cette incertitude est également calculée à partir des résultats du projet ESHM13 [Woessner et al., 2015] et de l'étude de B2020 (Figure 5.28).

Deuxièmement, pour une période de retour donnée, la valeur de la métrique Inc varie peu en fonction des villes étudiées, sauf pour le cas SHARE 475 ans (Tableau 5.12) où les incertitudes les plus faibles sont observées à Lourdes et Briançon. Ceci n'est *a priori* pas en cohérence avec l'hypothèse (2). Comme déjà énoncé ci-avant, ce constat trouve probablement son origine dans le fait qu'une seule DFM soit calculée à l'échelle de la France entière. Quelle que soit la région d'étude, l'incertitude associée à la DFM est la même pour chacune de ces régions, tendant vers une homogénéisation de l'incertitude de l'aléa. L'étude ESHM13 semble en cohérence avec cette hypothèse (2) contrairement à l'étude de B2020.

Ensuite, seules les incertitudes SHARE associées à une période de retour de 2 475 ans pour les villes de Briançon et de Lourdes sont supérieures aux incertitudes de l'étude ESHM13 (Tableau 5.12). Toutes nos autres incertitudes sont plus faibles que celles de cette dernière. Or, notre étude est réalisée à l'échelle de la France tandis que l'étude ESHM13 travaille à l'échelle européenne. Sur la base de cette comparaison, l'hypothèse (3) semble donc valide.

Enfin, dans pratiquement tous les cas de figure (périodes de retour et villes), nos incertitudes sont plus faibles que celles observées dans ces deux autres études (Tableau 5.12). Or, B2020 utilisent un arbre logique de 96 branches (par ex. Figure 11 Beauval et al. [2020]) prenant notamment en compte deux catalogues différents, deux algorithmes de *declustering* différents, plusieurs modèles source, plusieurs GMPE ou encore trois magnitudes minimales. De plus, l'arbre logique de l'étude ESHM13 [Woessner et al., 2015] prend en compte, quant à lui, trois modèles source (zonages, failles et lissage), plusieurs magnitudes maximales ou encore plusieurs GMPE. Ainsi, de plus grandes prises en compte de la variabilité sont opérées dans ces deux études comparé à la présente étude. Les résultats de notre étude sont donc *a priori* cohérents avec l'hypothèse (4).

5.6.2 Impact des différents modèles utilisés

Dans cette Section, différents modèles sont comparés afin d'analyser l'impact qu'ils ont sur l'estimation de l'aléa sismique. Cette comparaison se base sur les centiles (Fig-

ure 5.28) obtenus à l'aide de chaque modèle. Pour se faire, un écart normalisé En entre les centiles de PGA calculés par un modèle ($M1$) et ceux obtenus à l'aide d'un autre modèle ($M2$) est calculé à l'image de ce qui est fait Équation 5.2 :

$$En = \frac{M1 - M2}{M2}. \quad (5.4)$$

Les modèles analysés dans cette Section sont la régionalisation, la prise en compte des répliques et le modèle de pondération des GMPE.

Régionalisation Les deux régionalisations utilisées dans cette thèse sont le zonage sismotectonique de GEOTER et la régionalisation tectonique de SHARE. Dans la Section précédente, il a été démontré que le choix de la régionalisation a un impact sur les PGA moyennes (Figure 5.27) et médianes ainsi que sur les centiles (Figure 5.28). La Figure 5.29 présente l'écart normalisé En (Équation 5.4) calculé entre SHARE ($M1$) et GEOTER ($M2$) pour les périodes retour de 475 et 2 475 ans.

Le choix de la régionalisation a un impact sur toutes ces villes, cependant il n'est pas le même partout. En ce qui concerne les valeurs absolues de ces écarts normalisés, la plus petite est respectivement observée à Nantes (4%) et à Nice (7,5%) pour les périodes de retour 475 et 2 475 ans respectivement. L'impact le plus fort du choix de la régionalisation se fait ressentir à Clermont-Ferrand, Strasbourg et Lourdes (24%) à 475 ans et à Briançon (32%) à 2 475 ans. À titre de comparaison, Beauval et al. [2020] estiment qu'à 475 ans de période de retour, l'impact du zonage sismotectonique sur ces six villes s'élève jusqu'à 30% (Clermont-Ferrand). D'après ces mêmes auteurs, Briançon est impactée de l'ordre de 26%. Ces chiffres sont très proches de ce que nous trouvons. Nuançons ces comparaisons par le fait que l'utilisation d'une régionalisation dans cette thèse est différente de l'utilisation classique d'un zonage sismotectonique. De plus, les deux zonages utilisés par B2020 sont différents de ceux utilisés dans cette thèse. En appliqué sur les PGA moyennes (Tableau 5.10) a pour valeurs absolues maximales 19,7% (Briançon) à 475 ans et 25% (Lourdes) à 2 475 ans.

En ce qui concerne le signe de ces écarts normalisés, le même constat que celui fait dans la Section précédente est observé Figure 5.29. Le passage de la régionalisation GEOTER à SHARE entraîne une diminution de l'aléa pour les villes les plus actives (Lourdes, Briançon, Nice et Nantes) et une augmentation de celui-ci pour les villes les moins actives (Clermont-Ferrand et Strasbourg).

Enfin, pour les trois villes les plus actives (Lourdes, Briançon et Nantes), l'augmentation de la période de retour accroît l'impact du choix de la régionalisation. En revanche, le passage de 475 à 2 475 ans ne modifie pas cet impact pour les trois autres villes. Le passage de la régionalisation GEOTER à SHARE entraîne une diminution de la magnitude maximale autorisée (Tableaux 5.2 et 5.3) dans les villes de Lourdes (de 7,1 à 7) et de Briançon (de 7,3 à 6,6). Or, plus la période de retour augmente, plus les fortes magnitudes jouent un rôle important sur l'aléa sismique. L'augmentation de l'impact de la régionalisation lors du passage de 475 à 2 475 ans peut ainsi être expliqué par une limitation plus stricte de la magnitude maximale dans ces deux villes. En ce qui concerne Nantes, la raison de l'accroissement de l'impact reste inconnue.

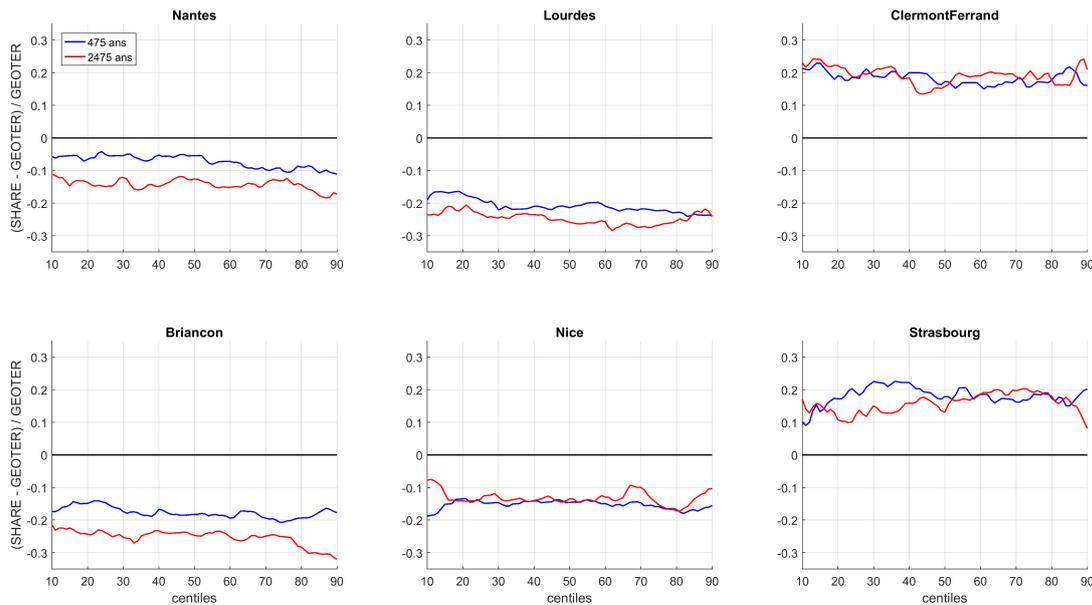


Figure 5.29: IMPACT DE LA RÉGIONALISATION SUR L'ESTIMATION DE L'ALÉA. Écart normalisé (Équation 5.4) entre les centiles des PGA 475 (bleu) et 2475 ans (rouge) obtenus à partir des régionalisations GEOTER et SHARE. La normalisation est réalisée avec les résultats GEOTER.

Prise en compte des répliques Dans ce paragraphe, l'impact de la prise en compte des répliques est analysé. Pour cela, l'écart normalisé En (Équation 5.4) est calculé entre la PGA ne prenant pas en compte les répliques ($M1 = Ms$) et celle les prenant en considération ($M2 = MsAs$). La Figure 5.30 présente la distribution des En appliqués sur les centiles GEOTER pour les six villes et aux périodes de retour 475 et 2475 ans.

Les écarts normalisés absolus sont compris en général entre 0 et 5% pour la plupart des villes. La ville la moins impactée par l'ajout des répliques est Clermont-Ferrand dont En se situe entre 0 et 2,5%. L'impact le plus fort s'élève à 9% (Lourdes, Nice et Strasbourg). En appliqué sur les PGA GEOTER moyennes (Tableau 5.10) a pour valeur absolue maximale 4,6% (Lourdes) à 475 ans et 5,8% (Lourdes) à 2475 ans. Sur les PGA SHARE moyennes, ces valeurs maximales absolues s'élèvent respectivement à 3,3% (Lourdes) et 4,7% (Nice). En moyenne, la régionalisation SHARE diminue légèrement l'impact de la prise en compte des répliques.

Avec en moyenne 33% de répliques $M_L \geq 3$ (Figure 8), Beauval et al. [2006] estiment que ne pas prendre en compte les répliques entraîne une sous-estimation de 10 à 12% de l'aléa à période de retour de 475 ans dans les Pyrénées. Une autre étude [Faenza et al., 2007], réalisée dans le Bassin Rhénan inférieur, a montré que cette sous-estimation est de l'ordre de 8% pour des séismes $M_s \geq 4$. Au regard des six villes étudiées, l'impact sur l'aléa de la non prise en compte des répliques (Figure 5.30), avec en moyenne 23% de répliques $M_w \geq 4$ (Figure 5.17), est comparable avec ce que trouvent ces études.

De leur côté, Iervolino et al. [2018] analysent l'impact de la prise en compte des répliques dans une étude PSHA appliquée à l'Italie. Dans cette étude, ils démontrent que cet impact peut atteindre localement 30%, que ce soit pour une période de retour de 475 ou de 2475 ans. Ils constatent également que les impacts les plus élevés se trouvent principalement dans les régions les plus actives. Une analyse de cet impact à l'échelle

de la France permettrait de comparer ces résultats avec ceux de notre étude. Nous pouvons tout de même constater que Clermont-Ferrand et Lourdes sont respectivement les villes la moins et la plus impactées par la prise en compte des répliques. Or ces deux villes sont respectivement les villes la moins et la plus actives d'après le Tableau 5.10, ce qui va dans le même sens que ce qu'observent Iervolino et al. [2018] en Italie.

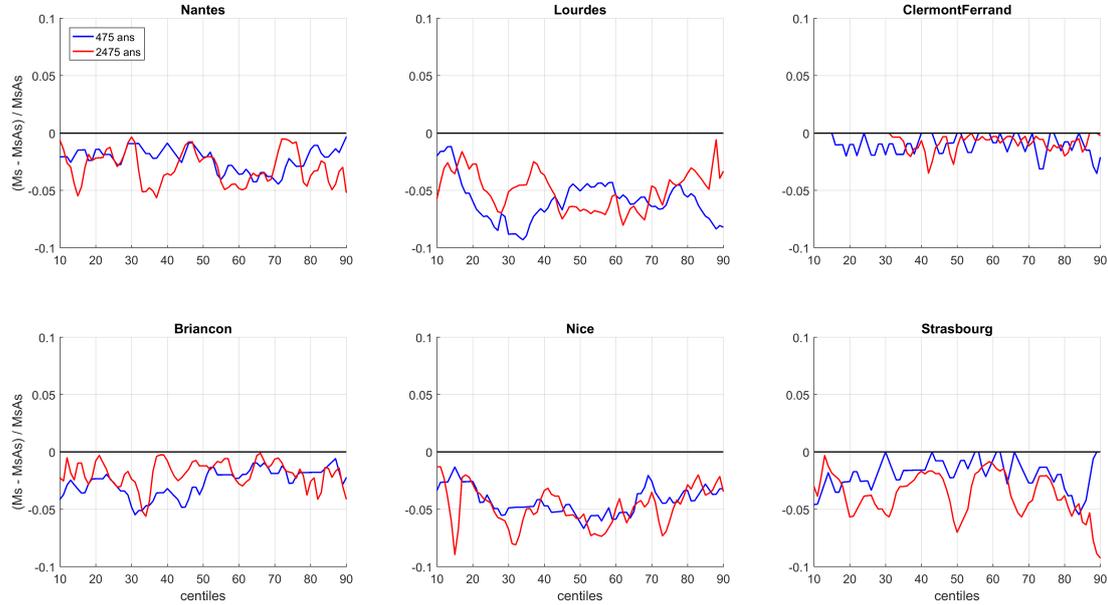


Figure 5.30: IMPACT DE LA PRISE EN COMPTE DES RÉPLIQUES SUR L'ESTIMATION DE L'ALÉA.

Écart normalisé (Équation 5.4) entre les centiles des PGA 475 (bleu) et 2475 ans (rouge) obtenus avec (MsAs) ou sans répliques (Ms). La normalisation est réalisée avec les résultats prenant en compte les répliques. Ces résultats sont obtenus via la régionalisation GEOTER.

Enfin, plus la période de retour augmente, plus les fortes magnitudes contribuent à l'aléa et donc moins les répliques, caractérisées par des magnitudes plus faibles, devraient impacter l'aléa. Or, le passage de 475 à 2475 ans ne change rien à cet impact. Il le fait même augmenter dans le cas de Nantes (de 4,5 à 5,5% maximum), Nice (de 6,5 à 9% maximum) et Strasbourg (de 5 à 9% maximum). Il est difficile de trouver une explication à ce constat. À noter que dans l'étude italienne d'Iervolino et al. [2018], le passage de 475 à 2475 ans ne modifie pas non plus cet impact (Figure 5).

Modèle de pondération des GMPE Le but de ce paragraphe est de comparer via l'écart normalisé En (Équation 5.4) notre modèle de pondération présenté Section 5.5.1 ($M2=Pond$) avec celui proposé par Martin et al. [2018] ($M1=M2018$). La Figure 5.31 présente la distribution de ces En appliqués aux centiles GEOTER pour les six villes et aux périodes de retour 475 et 2475 ans.

L'écart normalisé absolu maximal est observé à Nantes (25%) pour 475 ans et à Nice (17%) pour 2475 ans. Un écart normalisé nul est observé quatre villes sur six à la période de retour 2475 ans tandis qu'à 475 ans, En minimal est égal à 2,3% (Briançon). À titre de comparaison, Delavaud et al. [2012] analysent l'impact du choix du modèle de pondération sur le calcul de l'aléa à l'échelle de l'Europe. Dans cette étude, les écarts

absolus maximaux calculés pour une application aux régions actives peu profondes sont de 7,7% à 475 ans et de 6,5% à 2 475 ans (Tableau 14). La différence entre notre modèle de pondération et celui de M2018 est donc plus forte que celles observées entre les modèles de pondération choisis par Delavaud et al. [2012].

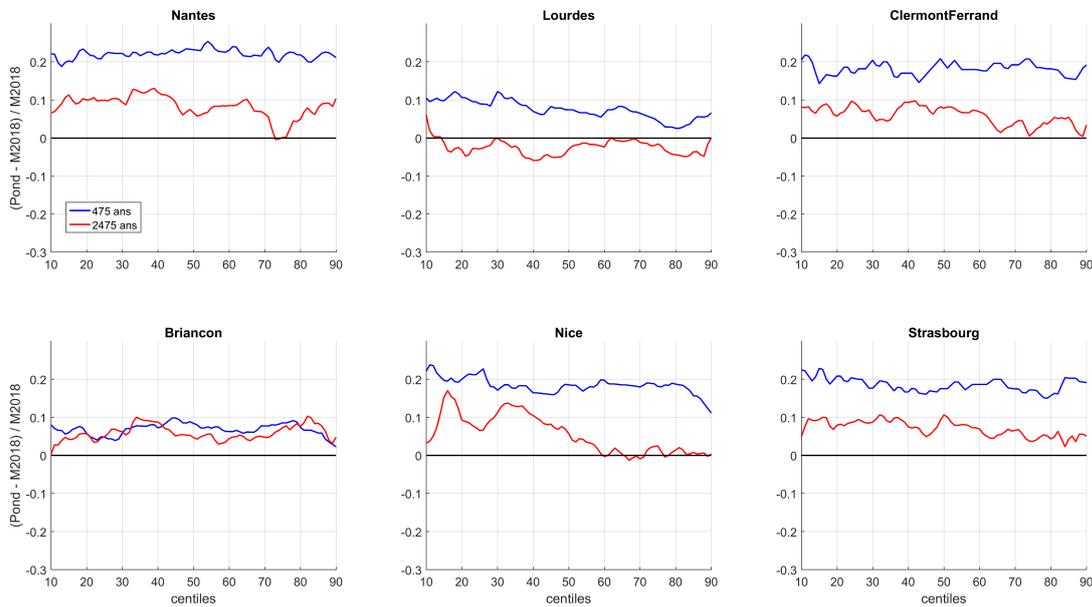


Figure 5.31: IMPACT DE LA MÉTHODE DE PONDÉRATION SUR L'ESTIMATION DE L'ALÉA.

Écart normalisé entre les centiles des PGA 475 (bleu) et 2475 ans (rouge) obtenus avec deux méthodes de pondération des GMPE : celle proposée Section 5.5.1 (Pond) et celle proposée par Martin et al. [2018] (M2018). La normalisation est réalisée avec les résultats de la méthode proposée dans cette thèse. Ces résultats sont obtenus via la régionalisation GEOTER.

À la période de retour 475 ans, l'impact du choix du modèle de pondération est d'environ 10% pour les deux villes considérées comme les plus actives (Lourdes et Briançon) tandis que les autres villes sont caractérisées par un En environ deux fois plus élevé (20%). Cet impact se ferait donc moins ressentir dans les régions les plus actives.

Hormis pour Lourdes à 2 475 ans (jusqu'à -6%), les calculés En sont positifs. Ainsi, de manière générale, notre modèle de pondération produit de l'aléa plus élevé que celui de M2018, comme déjà mentionné 5.5.2 et illustré Figures 5.25 et 5.26. De plus, une diminution marquée de En est observée lors du passage de 475 à 2 475 ans. D'après la Figure 5.25, notre modèle de pondération produit des PGA plus élevées que le modèle de M2018 surtout pour les magnitudes les plus faibles ($M_w \leq 4,5$). Or, plus la période de retour augmente, plus les fortes magnitudes contribuent à l'aléa et donc moins la différence entre notre modèle de prédiction et celui de M2018 est forte.

Enfin, En appliqué sur les PGA GEOTER moyennes (Tableau 5.10) a pour valeur absolue maximale 22,7% (Nantes) à 475 ans et 8,4% (Nantes) à 2 475 ans. Sur les PGA SHARE moyennes, ces valeurs maximales absolues s'élèvent respectivement à 29,5% (Strasbourg) et 9,4% (Nice). À l'inverse de ce qui a été observé pour la prise en compte des répliques, en moyenne, la régionalisation SHARE augmente légèrement l'impact du choix du modèle de prédiction.

5.6.3 Contributions distance - magnitude

Comme énoncé Section 1.3.3, un des avantages de l'approche PSHA des catalogue synthétiques est le calcul naturel des contributions distance - magnitude sur l'aléa sismique. Les Figures 5.32 et 5.33 présentent ces contributions pour les PGA à période de retour de 475 et 2475 ans respectivement. Ces contributions sont obtenues à l'aide de la régionalisation GEOTER. Quelle que soit la période de retour, ces contributions se concentrent en grande majorité à moins de 50 km de distance JB (D_{JB}).

En ce qui concerne la période de retour 475 ans (Figure 5.32), la contribution maximale observée dans toutes les villes, exceptées les deux plus actives (Lourdes et Briançon), se fait à $M_w 4,2$ et une D_{JB} inférieure à 15 km. Ce couple distance - magnitude rassemble pour ces villes (sauf Nice et Lourdes) plus de 10% de la contribution. Pour ce qui est de Briançon, ce couple distance - magnitude contribue également mais moins que deux autres couples : $M_w 5,2 - D_{JB} < 25$ km et $M_w 5,6 - D_{JB} < 25$ km. Ces deux couples sont également ceux qui contribuent le plus à la PGA de période de retour 475 ans pour la ville de Lourdes. Ces deux couples contribuent fortement ($\approx 15\%$ chacun) aussi pour Nantes, la troisième ville la plus active d'après le Tableau 5.10.

Aucune ou très peu de contributions des séismes à magnitude supérieure à 6,5 n'est à noter pour Clermont-Ferrand, Strasbourg et Nantes, villes respectivement limitées par des magnitudes maximales de 5,9, 6,5 et 6,4 (Tableau 5.2). En revanche, les séismes $M_w \geq 6,5$ situés à moins de 35 km contribuent à plus de 20% de l'aléa à Lourdes, localisée dans la région 31 où la magnitude maximale autorisée s'élève à 7,1. Pour Briançon ces séismes $M_w \geq 6,5$ contribuent un peu moins (15%) pour une distance plus élevée ($D_{JB} < 45$ km). Ce constat est lié au fait que Briançon est localisée dans la région 23 limitée par $M_w 6,4$. Cependant, Briançon se trouve à la frontière avec la région 24 elle-même limitée par une magnitude 7,3 (Tableau 5.2). La proximité de cette région 24 impacte donc l'aléa final observé à Briançon. Le même constat, plus marqué, est réalisé à Nice. La région 25 où se trouve Nice est limitée par une faible magnitude maximale ($M_w 5,8$), cependant Nice se situe proche de la même région 24. Cette proximité implique une contribution de plus de 22,5% des séismes $M_w \geq 6,5$ situés à une distance inférieure à 65 km.

En utilisant la régionalisation SHARE, la contribution maximale pour toutes les villes s'observent à $M_w 4,2$ et à $D_{JB} < 15$ km. Cette homogénéisation des contributions est en accord avec ce qui a déjà été dit dans les Sections précédentes concernant la régionalisation SHARE.

En ce qui concerne la période de retour 2475 ans (Figure 5.33), les magnitudes inférieures ou égales à 4,5 contribuent très peu ($< 10\%$) sauf pour les deux villes les moins actives : Clermont-Ferrand (22%) et Strasbourg (13%). Ce constat est cohérent car plus la période de retour augmente, plus les séisme de fortes magnitudes contribuent à l'aléa sismique. Les contributions les plus importantes ($\geq 36\%$) pour les six villes sont associées à des magnitudes comprises entre 5,2 et 5,8 et des distances JB inférieures à 25 km, soit les couples les plus contributeurs des deux villes les plus actives à 475 ans (Figure 5.32).

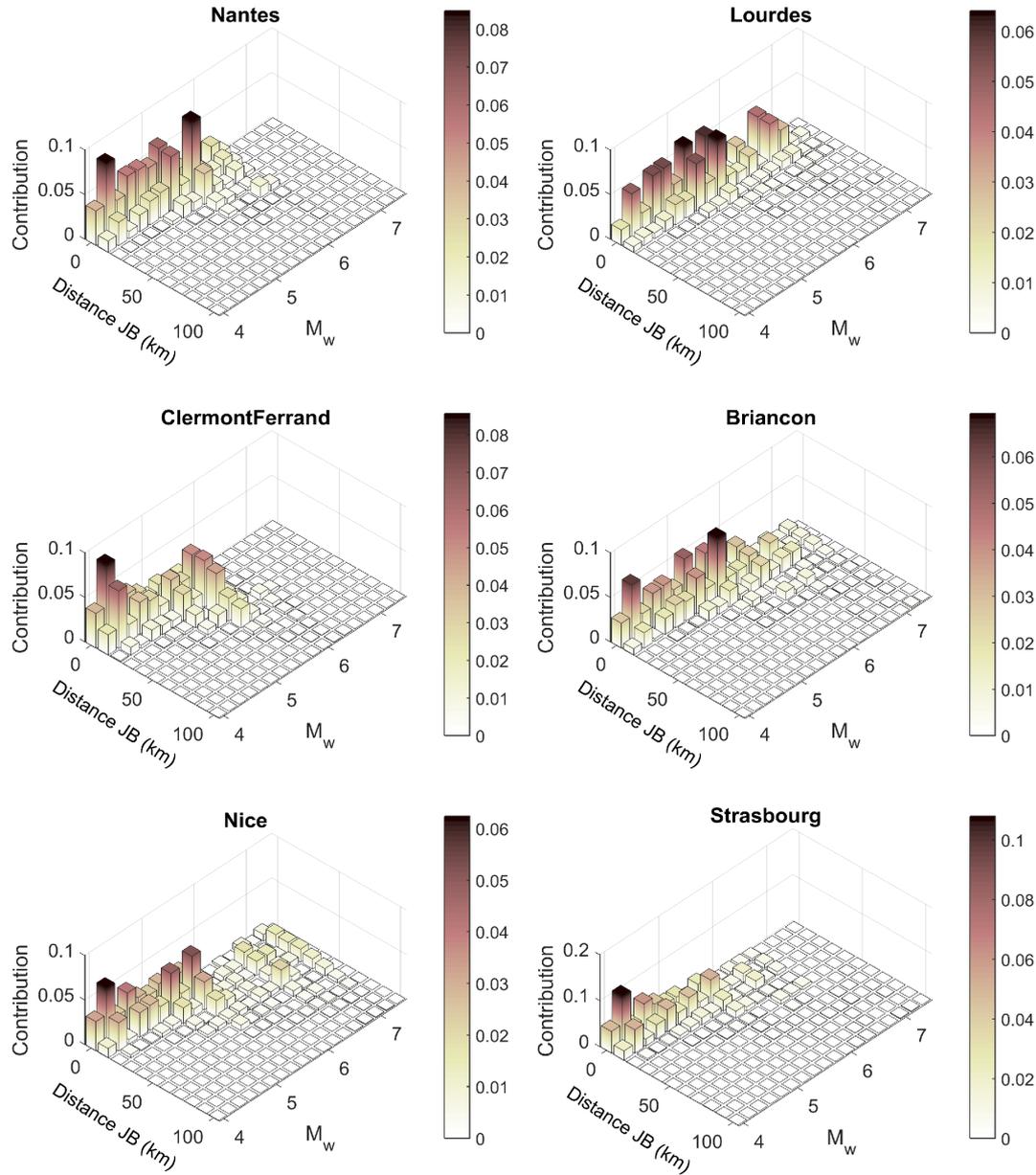


Figure 5.32: CONTRIBUTION DISTANCE - MAGNITUDE DE LA PGA À 475 ANS. Contribution distance-magnitude de la PGA à 475 ans observée dans les six villes d'étude. Ces résultats sont obtenus via la régionalisation GEOTER.

Les trois villes les plus actives (Lourdes, Briançon et Nantes) sont caractérisées par une forte contribution (12- 15%) des séismes $M_w 6,2$ situés à moins de 25 km. Nantes est ensuite limitée par la magnitude maximale de sa région (6,4). En revanche, les contributions des séismes $M_w \geq 6,5$ situés à moins de 25 km s'élèvent à 36% et 25% pour Lourdes et Briançon respectivement, contre 20% et 12,5% à la période de retour de 475 ans. Pour la même raison évoquée à la période de retour de 475 ans, l'aléa autour de la ville de Nice est produit à 30% par des séismes $M_w \geq 6,5$ situés à moins de 45 km.

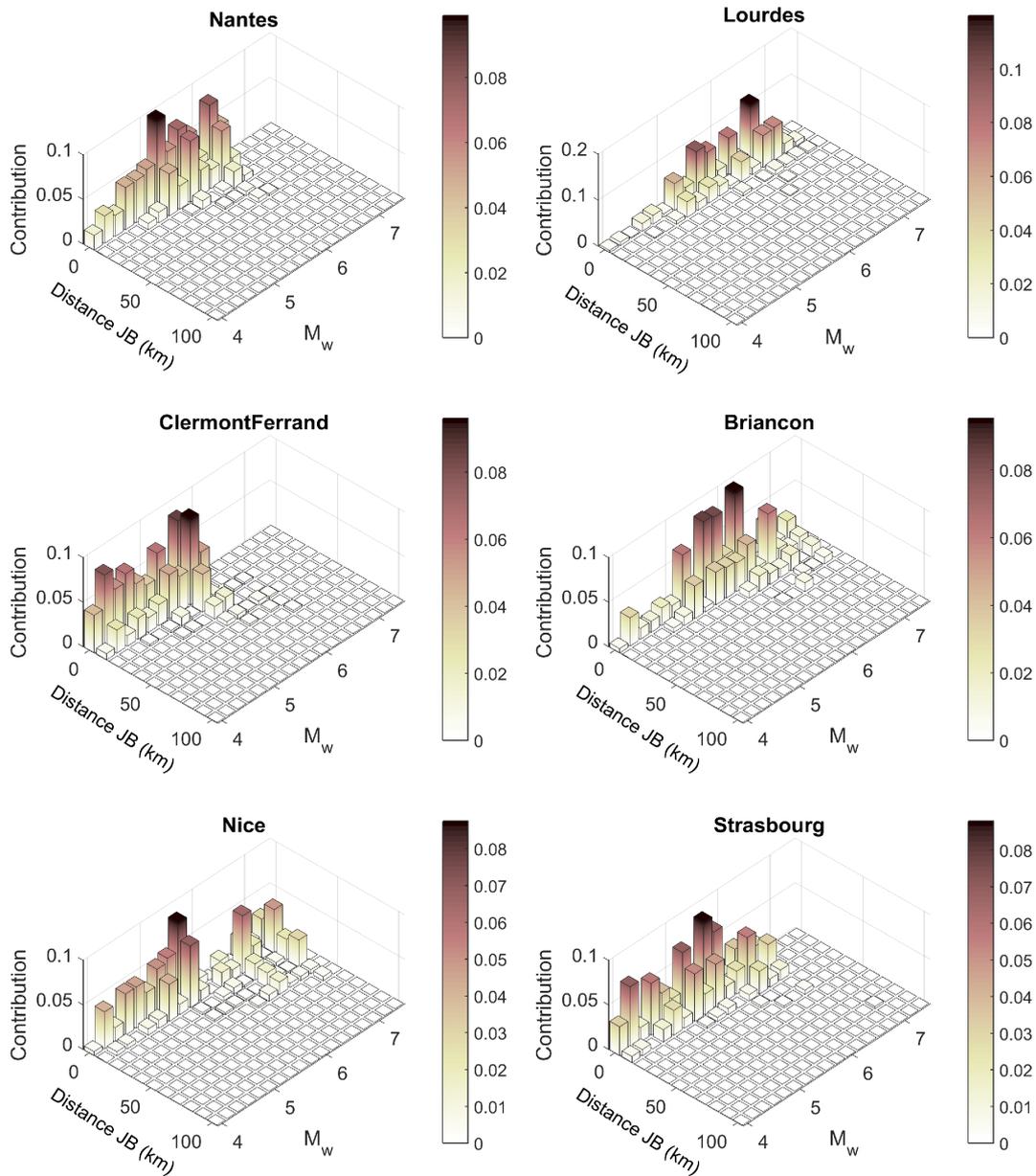


Figure 5.33: CONTRIBUTION DISTANCE- MAGNITUDE DE LA PGA À 2475 ANS. Contribution distance- magnitude de la PGA à 2475 ans observée dans les six villes d'étude. Ces résultats sont obtenus via la régionalisation GEOTER.

Pour toutes les villes, les contributions obtenues à l'aide de la régionalisation SHARE sont encore plus marquées ($\geq 40\%$) autour de M_w 5,2- 5,8 et des distances JB inférieures à 25 km. Encore une fois, cet aspect homogénéisant de la régionalisation SHARE est cohérent avec les constats faits précédemment.

5.7 Conclusion

Dans ce Chapitre, les méthodes proposées dans les Chapitres 2, 3 et 4 ont été appliquées à la France métropolitaine.

L'application de la méthode des temps de latence et de l'implémentation MIMP (Chapitre 2) aux données françaises a permis de calculer les distributions de proportions (DPM) et de fréquences (DFM) de séismes principaux en fonction de la magnitude. Une des conclusions à cette application est qu'en deçà de 40 à 60 données, la méthode des temps de latence semble aboutir à des résultats peu fiables. L'implémentation MIMP permet de remédier à ce manque de fiabilité en produisant une DPM de séismes principaux plus stable. Les proportions de séismes principaux calculées dans cette thèse sont plus faibles que celles des algorithmes de *declustering*. Cela abonde dans le sens de la littérature qui stipule que ces derniers ne détecteraient pas assez de répliques. Cependant, la DPM de séismes principaux obtenue dans cette thèse a un comportement binaire : ≈ 0.7 de séismes principaux à $M_w < 5$ et ≈ 1 au-delà de $M_w 5,2$. Ce constat ne semble pas réaliste et est probablement induit par l'implémentation MIMP appliquée au cas particulier des données françaises métropolitaines.

En ce qui concerne la DFM des séismes principaux, deux lois de puissance ont été observées : l'une de pente moyenne 0,75 entre les magnitudes 4 et 5,5 et l'autre de pente moyenne 1,39 au-delà de $M_w 5,5$. En comparant la valeur de cette pente avec celles obtenues à l'échelle planétaire dans la littérature, nous avons conclu qu'il s'agit d'une véritable observation. De plus, en observant uniquement les données il a été possible de reproduire le plongement asymptotique des fréquences annuelles sans utiliser d'équation; donnant encore plus de crédit à cette DFM atypique. Cette DFM est utilisée comme donnée d'entrée du générateur de séismes synthétiques. Il est à noter qu'une telle distribution est bien différente de ce que la littérature propose et est donc à utiliser avec précaution. Une des conséquences de son utilisation est la diminution des fréquences annuelles des séismes $M_w \geq 5,6$ vis-à-vis d'une approche classique (loi GR). Enfin, le choix du seuil MIMP, seul paramètre de l'implémentation MIMP, impacte peu les DPM et DFM des séismes principaux.

Le générateur de séismes est appliqué à la France métropolitaine avec comme données d'entrée les DPM et DFM des séismes principaux calculées en amont, une densité spatiale des tracés de failles (actives ou non) et deux régionalisations (GEOTER et SHARE). La répartition spatiale de la sismicité synthétique produite par le générateur est cohérente avec la densité spatiale et la régionalisation fournies en entrée du générateur. De plus, sans utiliser de modèles du processus ETAS nécessitant plusieurs paramètres, le générateur parvient à les reproduire. C'est le cas notamment du modèle décrivant le nombre de répliques produites par un séisme principal en fonction de sa magnitude. Le générateur parvient à reproduire la DFM de séismes principaux donnée en entrée, justifiant le bon fonctionnement de ce dernier. En revanche, il ne respecte par totalement la DPM de séismes principaux donnée en entrée. En effet, celle produite par le générateur a une tendance plus proche de celles calculées par les algorithmes de *declustering*.

Les cinq GMPE utilisées dans cette étude sont celles récemment employées dans les travaux de Martin et al. [2018]. L'application du modèle de pondération présenté Chapitre 4 à cet ensemble de GMPE aboutit à une répartition des poids principalement (94%) entre quatre de ces cinq GMPE. Les deux GMPE propres aux données françaises [Drouet et Cotton, 2015, Ameri, 2014] sont au sommet de la hiérarchie avec 29 et 23% du contrôle de l'ensemble de la distribution distance - magnitude.

De manière générale ces poids sont binaires : soit 1, soit 0. Cela est principalement dû à l'analyse résiduelle employée et surtout à la méthode d'interpolation des poids (le plus proche voisin). Cela a pour conséquence de produire une distribution de PGA pondérée non lisse. En effet, de grandes variations de cette PGA peuvent s'observer entre deux magnitudes proches ou entre deux distances proches. Une interpolation linéaire prendrait, par exemple, en compte les variabilités observées entre plusieurs poids calculés sur des points de données proches les uns des autres, rendant la distribution des poids et donc de la PGA plus lisse.

Enfin, de manière générale, la distribution de PGA finale pondérée est plus élevée que celle obtenue avec le modèle de pondération proposé par Martin et al. [2018]. Cette surestimation s'observe notamment pour les plus faibles magnitudes ($M_w \leq 4,5$).

Finalement, l'estimation de l'aléa est réalisé à l'échelle de la France entière sur une maille de 250 mètres de résolution et illustrée ici sur six villes : Lourdes, Briançon, Nantes, Nice, Strasbourg et Clermont-Ferrand. Les résultats de cette estimation sont dans l'ensemble cohérents avec de récentes études françaises [Martin et al., 2018, Drouet et al., 2020, Beauval et al., 2020]. Ces résultats s'inscrivent dans une tendance observée depuis le début du XXI^{ème} siècle, à savoir une diminution de l'estimation de l'aléa sismique en France, déjà mentionnée par Martin et al. [2018] et Drouet et al. [2020]. La plus grande différence entre notre étude et ces récentes études françaises est le fort aléa que nous trouvons pour la ville de Nantes. Cette différence trouve son origine dans la forte densité spatiale de tracés de failles observée autour de Nantes, principalement dû au cisaillement Sud Armoricaïn. Le choix fort de considérer l'ensemble des tracés de failles comme analogue à la sismicité longue durée a donc un impact sur l'estimation de l'aléa sismique.

Des analyses d'impact ont montré que le choix de la régionalisation ainsi que du modèle de pondération des GMPE sont les plus critiques sur l'estimation de l'aléa sismique dans cette étude. Le choix de la régionalisation entre GEOTER et SHARE impacte jusqu'à 24% et 32% les PGA à période de retour 475 et 2 475 ans respectivement. Il a été démontré que l'aspect global de la régionalisation SHARE (seulement six régions contre 35 pour GEOTER) tend à homogénéiser l'aléa sismique sur l'ensemble du territoire français. Cela se traduit par une diminution de l'aléa dans les régions les plus actives (par ex., Lourdes et Briançon) et une augmentation de celui-ci dans les régions les moins actives (par ex., Clermont-Ferrand et Strasbourg). En ce qui concerne le choix du modèle de pondération des GMPE, notre modèle a été comparé avec celui de Martin et al. [2018] (M2018). Comme déjà mentionné, notre modèle aboutit à des PGA plus fortes que celles obtenues avec le modèle de M2018. Au maximum, cette différence peut s'élever à 25% pour la période de retour 475 ans. En revanche, à 2 475 ans de période de retour, cet écart maximal chute à 17%. Cela témoigne que notre modèle de pondération

surestime moins le modèle de M2018 pour des magnitudes plus élevées. Ces conclusions sont toutefois à nuancer étant donné que seulement cinq des six GMPE utilisées par Martin et al. [2018] ont été employées dans cette thèse.

La prise en compte des répliques dans l'estimation de l'aléa sismique a également été analysée afin de quantifier son impact. Ce dernier est faible ($<10\%$) pour l'ensemble des six villes. Nos résultats sont assez proches des travaux de Beauval et al. [2006] et de Faenza et al. [2007]. De plus, l'impact de la prise en compte des répliques semble lié à l'activité sismique de la région étudiée : plus elle est active, plus l'impact est élevé. Ce constat a également été fait par Iervolino et al. [2018] à l'échelle de l'Italie.

Conclusions générales

Contributions principales

L'évaluation de l'aléa sismique dans les régions à sismicité faible à modéré est complexe. Le faible taux de déplacement du sous-sol et le peu de failles actives documentées dans ces régions ne permettent pas d'utiliser les méthodes basées sur le processus de rupture. De même, le peu de données sismiques (séismes passés, mécanismes au foyer, etc.) disponibles limite la fiabilité et la représentativité des méthodes statistiques qui nécessitent beaucoup de données (justesse, paramétrisation, etc.).

Dans cette thèse, je propose une méthodologie d'analyse de l'aléa sismique adaptée au contexte de sismicité faible à modérée et facilement incorporable dans un modèle d'estimation du risque sismique. Cette méthodologie repose sur la méthode PSHA des catalogues synthétiques dont le but est de produire une sismicité synthétique sur une longue période temporelle. La méthodologie globale proposée dans cette thèse regroupe (i) l'estimation du modèle de récurrence des séismes principaux adaptée au contexte de sismicité faible à modérée, (ii) le générateur de séismes et (iii) la stratégie de pondération des modèles de prédiction des mouvements du sol qui dépend de la magnitude du séisme et de la distance séisme-site. Ces trois étapes ainsi que leur application à la France métropolitaine sont les principales contributions de cette thèse, récapitulées ci-après.

Stabilisation du calcul des proportions de séismes principaux

Le calcul des proportions de séismes principaux se base sur la méthode des temps de latence [Molchan, 2004]. Cette méthode perd en justesse lorsqu'elle est conditionnée par une faible quantité de données. Le manque de données sismiques se fait principalement ressentir pour les fortes magnitudes, là où les proportions de séismes principaux sont sensées être fortes.

Le Chapitre 2 propose une solution pour lever ce verrou : l'implémentation MIMP (*Magnitude of Inferred Main shock Proportion*). Cette implémentation se base sur la loi de Bâth qui stipule que plus la magnitude des séismes augmente, plus la proportion de séismes principaux est grande. Le but est de trouver MIMP, la magnitude à partir de laquelle les résultats de la méthode des temps de latence deviennent instables et donc moins dignes de confiance. Toute proportion associée à une magnitude supérieure ou égale à MIMP est alors fixée à 1. Cela permet de stabiliser le calcul de la méthode des temps de latence pour les fortes magnitudes lorsque les données sont peu nombreuses.

Enfin, la prise en compte des incertitudes sur la magnitude du catalogue initial via un processus de Monte Carlo permet d'obtenir une vision complète des proportions de séismes principaux pour chaque magnitude.

L'instabilité de la méthode des temps de latence est principalement observée lorsque moins de 40 à 60 données sont disponibles. Sur les données françaises cela représente une MIMP comprise entre 5,2 et 5,4. Cela a pour conséquence (i) de considérer qu'à partir de ces magnitudes tous les séismes sont des séismes principaux et (ii) d'aboutir à une distribution proportion - magnitude (DPM) de séismes principaux différente de celles proposées par les algorithmes de *declustering*. La DPM obtenue dans cette thèse est plus conservative mais semble moins réaliste que celles proposées par ces derniers.

L'implémentation MIMP appliquée aux données françaises jusqu'au calcul de la distribution fréquence - magnitude (DFM) de séismes principaux a été publiée dans *Journal of Seismology* [Gouache et al., 2021].

Générateur de séismes

Distribution fréquence - magnitude La méthode des temps de latence calcule la DFM pour chaque pas de magnitude. Cela implique que le catalogue analysé doit être le plus complet possible pour toute la gamme de magnitudes étudiées. Calculer la DFM à l'échelle de la France entière permet de conserver le maximum de données et de diminuer les périodes de retour comparé à une analyse réalisée à l'échelle des zones. En considérant que la longue profondeur temporelle qu'offre le catalogue FCAT permet de couvrir exhaustivement les magnitudes les plus élevées, il est possible de calculer à l'aide de la méthode des temps de latence une fréquence représentative pour chaque magnitude directement à partir des séismes de ce catalogue.

La DFM calculée à l'échelle de la France entière est alors caractérisée par une loi de puissance de coefficient égal à 0,75 pour les faibles magnitudes ($M_w < 5,6$) et par une loi de puissance de coefficient égal à 1,39 pour les fortes magnitudes ($M_w \geq 5,6$). L'observation de ces deux lois de puissance est également faite à l'échelle planétaire [par ex., Sachs et al., 2012, Corral et González, 2019]. Le comportement de la DFM en deux lois de puissance est soit imputé à un manque de données dans le catalogue FCAT, auquel cas notre hypothèse d'un catalogue complet est fautive, soit dû à un phénomène physique mettant en jeu la largeur de la zone sismogène [Pacheco et al., 1992], soit dû à une autre explication que nos travaux actuels ne permettent pas de mettre en lumière. Au vu de la cohérence du coefficient de la deuxième pente calculé dans cette thèse (1,39) avec ceux obtenus dans d'autres études à l'échelle planétaire (1,5 d'après Rundle [1989] et 1,39 d'après Pacheco et al. [1992]), l'observation de ces deux lois semble juste. De plus, la seule analyse des données arrive à reproduire le plongement asymptotique des fréquences annuelles pour les plus fortes magnitude sans l'aide de modèle. Nous utilisons donc cette DFM pour simuler les séismes principaux en temps et en magnitude à l'échelle de la France entière. Néanmoins, gardons à l'esprit que cette DFM est bien différente de celle à pente unique obtenue par l'approche classique utilisant la loi GR. Ainsi, en supposant vraie l'observation des deux lois de puissance, les séismes à forte magnitude ($M_w \geq 5,6$) sont moins fréquents que ce qui est observé en utilisant une seule loi de puissance (type loi GR).

Générateur de séismes principaux La DFM décrivant les séismes principaux à l'échelle de la France entière est utilisée pour simuler l'occurrence et la magnitude

des séismes principaux. Cependant, le positionnement de ces séismes dans l'espace doit prendre en compte la variabilité spatiale de la sismicité française. Pour se faire, une régionalisation ainsi qu'une densité spatiale caractérisant la sismicité sont utilisées. Chaque région limite la magnitude maximale autorisée de sorte que le couplage de la densité spatiale et de la régionalisation aboutit à une densité spatiale dépendante de la magnitude. Les séismes sont finalement positionnés dans l'espace en considérant leur magnitude et la densité spatiale. En se basant sur la littérature [par ex., Liu et al., 2011, Stein et al., 2015, Calais et al., 2016], les tracés de failles sont utilisés pour calculer la densité spatiale représentative de la sismicité.

Générateur de répliques Le générateur de répliques donne à chaque séisme principal la possibilité de produire des répliques. La magnitude maximale que les répliques peuvent atteindre est dépendante de la magnitude du séisme principal via la loi de Bâth. Le nombre répliques par tranche de magnitude autorisée est calculé à partir des proportions de répliques, complémentaires des proportions de séismes principaux calculées en amont. Ces proportions de répliques sont assimilées à la probabilité de tirer une réplique en fonction de la magnitude.

Discussions Les générations de séismes principaux et de répliques sont complémentaires car basées sur la même DPM de séismes principaux. En effet, tandis que la génération de séismes principaux utilise la proportion p de séismes principaux, le générateur de répliques dépend de $1 - p$. Ainsi, peu de paramètres sont nécessaires à la production des catalogues synthétiques. Malgré cette faible quantité de paramètres utilisés par le générateur de répliques, le nombre de répliques produites en fonction de la magnitude du séisme principal suit la loi classiquement utilisée dans le modèle ETAS. De plus, la répartition spatiale de la sismicité synthétique est cohérente avec la densité spatiale et les régionalisations fournies en entrée du générateur de séismes principaux.

En revanche, le calcul de la DFM de séismes principaux à l'échelle de la France entière homogénéise le comportement temporel de la sismicité synthétique sur l'ensemble du territoire. Concrètement, les DFM synthétiques de chaque région partagent les mêmes pentes. La prise en compte de la variabilité spatiale de la sismicité à l'aide d'une densité dépendante de la magnitude ne permet pas de pallier cet aspect.

Enfin, les choix méthodologiques faits dans cette thèse peuvent perturber la production des séismes synthétiques. En effet, le générateur aboutit à une DPM différente de celle donnée en entrée.

Pondération des GMPE en fonction de la magnitude et de la distance

Les modèles de prédiction des mouvements du sol (GMPE) sont nombreux et variés. Leur application se fait généralement sous la forme d'un ensemble de GMPE pondérées. La pondération peut être guidée par des données à l'aide d'analyses comparatives, par dires d'experts ou les deux [par ex., Delavaud et al., 2012]. Dans tous les cas, un seul poids est attribué à chacune des GMPE. Il paraît pourtant judicieux que cette pondération prenne en compte un certain nombre de paramètres contrôlant le calcul de l'aléa tels que la distance séismes-site, la magnitude et le mécanisme au foyer du séisme ou

encore le type de sol au site.

Le Chapitre 4 présente une méthode de pondération de GMPE pré-sélectionnées. Les poids dépendent ici de la distance séisme-site et de la magnitude du séisme grâce à l'utilisation de données d'accélération (PGA). Ces valeurs de PGA enregistrées servent de référence et sont comparées aux PGA calculées par chacune des GMPE pré-sélectionnées. Le poids de chaque GMPE est calculé en chaque point de données dans l'espace distance-magnitude à l'aide d'une approche résiduelle. Une interpolation du plus proche voisin est ensuite appliquée entre les points de données afin d'obtenir un modèle de pondération en fonction de l'ensemble des distances et des magnitudes considérées. Pour toute distance et magnitude comprise dans les intervalles de validité de l'étude, ce modèle ressort donc les poids associés à chaque GMPE.

La pondération des cinq modèles de prédiction des mouvements du sol choisis pour l'application à la France métropolitaine donne la part belle à quatre d'entre eux. Parmi ces quatre modèles, on retrouve en première place les deux GMPE formulées avec des données françaises [Drouet et Cotton, 2015, Ameri, 2014], puis une GMPE établie avec les données européennes de RESORCE [Akkar et al., 2014a] et enfin une GMPE basée sur des données provenant de régions actives de par le monde [Boore et al., 2014]. La dernière GMPE [Bindi et al., 2014], établie avec les données européennes de RESORCE, est peu représentée (6%) dans ce modèle de pondération. Notre modèle de pondération aboutit à une PGA en général plus élevée que celle obtenue par la pondération, à dire d'experts, des mêmes GMPE proposée par Martin et al. [2018]. Ces PGA plus élevées s'observent principalement pour les magnitudes les plus faibles ($M_w \leq 5$).

Estimation de l'aléa sismique français

Finalement, l'ensemble des contributions apportées par cette thèse sont appliquées à la France métropolitaine afin d'estimer l'aléa sismique français. L'aléa produit par les séismes synthétiques est calculé au rocher à l'échelle de six villes d'étude à l'aide des GMPE pondérées. Cet aléa est plus faible que l'aléa réglementaire actuel¹¹, ce qui est en accord avec l'observation faite par Martin et al. [2018] et par Drouet et al. [2020] : l'aléa français est, semble-t-il, revu à la baisse depuis le début du XXI^{ème} siècle. De plus, nos résultats sont proches de ceux des dernières récentes études françaises [par ex., Martin et al., 2018, Beauval et al., 2020, Drouet et al., 2020]. La sous-estimation des fréquences annuelles de séismes principaux $M_w \geq 5,6$ de notre modèle de récurrence est contrebalancée par la surestimation de notre modèle de pondération des GMPE.

Enfin, les choix de la régionalisation et du modèle de pondération des GMPE sont ceux qui impactent le plus l'aléa sismique final avec des sur- ou sous-estimations pouvant atteindre 30%. La non prise en compte des répliques tend, quant à elle, à sous-estimer l'aléa sismique jusqu'à 9%, comme le suggèrent d'autres études à ce sujet [par ex., Beauval et al., 2006, Faenza et al., 2007].

¹¹<http://www.planseisme.fr/spip.php?page=carto>

Perspectives

Le générateur et les modèles développés dans cette thèse montrent des perspectives d'améliorations techniques intéressantes. Par exemple, une recherche d'exhaustivité automatique pourrait être implémentée afin d'estimer cette dernière pour chacun des N_{cb} catalogues initiaux bruités. De plus, une approche d'enrichissement des données d'accélération par des données synthétiques produites par des simulations de propagation d'ondes sismiques est envisageable. Cette thèse a initié des travaux sur cette approche qui n'ont abouti qu'à des résultats préliminaires. Ensuite, comme stipulé par Mak et al. [2017], dans l'idéal nous ne devrions pas utiliser la base de données RESORCE pour tester les GMPE d'Akkar et al. [2014a] et de Bindi et al. [2014] étant donné que ces dernières ont été élaborées à partir de ces données. D'autres bases de données peuvent être envisagées comme par exemple ESM [Lanzano et al., 2018]. Enfin, l'analyse de la nouvelle base de données de 1700 mécanismes au foyer en France (FMHex20, Mazzotti et al. [2021]) permettrait de définir les paramètres de rupture des séismes de manière dépendante. Des corrélations azimut- pendage et pendage-régime de faille pourraient être envisagées afin de rendre le générateur plus réaliste.

Au-delà de ces améliorations techniques, je liste ci-dessous les principales perspectives aux travaux réalisés pendant ma thèse.

Vers une meilleure prise en compte des incertitudes

L'approche proposée dans cette thèse prend en compte l'incertitude associée à la magnitude lors du processus de Monte Carlo (Section 2.3) et celle associée aux modèles de prédiction lors du calcul de l'aléa sismique (Section 4.2). L'incertitude due au choix de la régionalisation est exprimée Figure 5.27 et comprend également le choix des paramètres décrivant chacune des régions.

Dans cette thèse, la variabilité des magnitudes maximales du territoire et des régions n'est pas considérée. Ce choix est justifié par l'analyse de périodes de retour peu élevées (475 et 2475 ans) qui sont moins impactées par les magnitudes maximales (Figures 5.32 et 5.33). Néanmoins, même si cet impact sur l'estimation de l'aléa sismique est moindre il serait intéressant de l'évaluer. Des approches telles que celles proposées par Kijko [2004] ou Zentner [2018] pourraient être appliquées pour définir cette variabilité dans le but de la prendre en compte lors du calcul de la DFM et de la génération des séismes.

Un autre point d'amélioration concerne la stratégie de pondération des modèles de prédiction des mouvements du sol. Les incertitudes liées à la magnitude et à la distance ne sont pas prises en compte dans le processus de pondération. Même si la base de données RESORCE ne donne pas d'informations sur ces incertitudes, des catalogues instrumentaux fournissent des incertitudes moyennes sur les magnitudes ainsi que des informations concernant l'incertitude sur la distance (incertitude sur la profondeur et la localisation, nombre de stations, etc.). En appliquant un processus de Monte Carlo à la méthode de pondération, il serait possible de prendre en compte ces incertitudes afin d'atteindre des magnitudes et des distances non atteintes (Section 5.5.1), comme par exemple les magnitudes comprises entre 6,5 et 6,8.

Toujours en ce qui concerne la stratégie de pondération des GMPE, des tests de sensi-

bilité sur les pas de magnitude et de distance utilisés pour l'interpolation (Section 4.3.2) peuvent aussi être réalisés afin de déterminer leur influence sur cette dernière. Par exemple, ces pas peuvent être choisis à partir des incertitudes moyennes de la distance et de la magnitude ou encore en fonction du nombre de données disponibles.

Vers une application à large échelle

Nous avons fait l'hypothèse que la distribution fréquence - magnitude peut être définie de façon unique à l'échelle de la France métropolitaine. Une telle hypothèse se justifie par les aspects pratiques suivants :

- l'augmentation du nombre de données analysées impliquant une diminution des périodes de retour. Sous couvert d'un catalogue sismique complet, cela permet de calculer directement les fréquences annuelles des séismes pour chaque pas de magnitude sans extrapolation.
- la cohérence des DFM obtenues dans chaque région avec la DFM du territoire entier : la somme des DFM de chaque région aboutit à celle observée à l'échelle du territoire.

Dès lors, nous considérons que la variabilité de la sismicité est principalement liée à la localisation des séismes. La généralisation de cette hypothèse peut être envisagée à d'autres territoires stables ou bien à une plus large échelle. Toutefois, pour étudier globalement la sismicité un prérequis nécessaire est l'harmonisation des catalogues de sismicité. Certains travaux comme les projets européens SHARE ou SIGMA, tentent d'harmoniser la sismicité (temps, magnitude, espace, etc.) de différents territoires, rendant ainsi possible l'étude du modèle de récurrence à de telles échelles.

Nuançons tout de même cette hypothèse par le fait que le comportement temporel de la sismicité ne saurait être dépendant de la spatialité. Or, dans cette thèse un seul comportement temporel est décrit pour l'entièreté du territoire. Une des conséquences est la moyennisation de l'aléa sismique final aux périodes de retour 475 et 2475 ans. À plus grande échelle (par ex., Europe), il pourrait donc être envisagé de calculer plusieurs DFM au sein de macro-régions. Cela a par exemple été fait pour la France par [Drouet et al., 2020] dans le but de calculer des valeurs préalables de b . Ces auteurs utilisent six macro-zones et calculent des valeurs de b variant de 0,79 à 1,01.

Sur le même principe, l'idée de base de notre stratégie de pondération des GMPE est de prendre en compte la magnitude et la distance car ces paramètres sont la base de toute GMPE. Néanmoins, la prise en compte de ces deux paramètres, en plus du type de sol défini au rocher, tend à diminuer significativement le nombre de données disponibles pour chaque couple distance-magnitude. Or, d'après notre modèle de pondération, la GMPE de [Boore et al., 2014] domine un cinquième de la distribution distance-magnitude alors qu'elle a été établie à l'aide de données de régions actives de par le monde (principalement Californie, Tawaï, Chine et Japon). Elle est même prédominante par rapport à celle de Bindi et al. [2014] qui a pourtant été formulée à l'aide de la base de données RESORCE. Une telle observation défend l'idée d'utiliser une base de données à l'échelle mondiale, en complément ou non de RESORCE, afin d'augmenter le

nombre de données. La base de données NGA-West2 [Ancheta et al., 2014], contenant principalement des données californiennes, taiwanaises, japonaises et chinoises, pourrait par exemple être utilisée à cet effet.

Liste des Figures

1	Impact des séismes sur la société de 1970 à 2019	2
2	Sinistralité sismique en France métropolitaine de 1982 à 2019 . . .	3
3	Occurrence spatio-temporelle des séismes interplaques	4
4	Occurrence spatio-temporelle des séismes intraplaques	7
5	Répartition mondiale des lois et modèles décrivant la sismicité . .	8
6	Résumé des enseignements dispensés au cours de la thèse	12
1.1	Les quatre étapes principales du PSHA	20
1.2	Calcul de la probabilité annuelle de dépassement via la méthode des catalogues synthétiques	22
1.3	Variabilité des courbes probabilité-aléa	24
1.4	Exemples de zonages sismotectoniques français	26
1.5	Comparaison des modèles source zonage et lissage	27
1.6	Loi GR tronquée appliquée à la sismicité française	29
1.7	Loi GR appliquée à l'échelle mondiale	30
1.8	Comparaison de différents algorithmes de <i>declustering</i>	32
1.9	Distribution PGA- magnitude et modèles associés	36
2.1	Distributions de proportions de séismes principaux calculée et at- tendue	41
2.2	Représentation graphique de l'implémentation MIMP	44
3.1	Schéma général du tirage spatial	52
3.2	Calcul de ΔM à l'aide du ratio R des moments sismiques	56
3.3	Exemples de tirage des répliques d'un séisme principal $M_{sp6,8}$. .	58
3.4	Distribution spatiale de répliques autour du séisme principal . . .	60
4.1	Comparaison de PGA calculée pour deux magnitudes différentes .	67
4.2	Création du plan de rupture	68
4.3	Domaine distance - magnitude utilisé dans cette thèse	71
4.4	Distribution de la PGA en fonction de la magnitude et de la distance	73
5.1	Évolution spatio-temporelle du réseau de sismographes large bande	79
5.2	Complémentarité des catalogues instrumental et historique	81
5.3	Densité spatiale utilisée dans cette thèse	83
5.4	Régionalisation du territoire français	84
5.5	Densités de probabilité des orientations de failles	87
5.6	Exemple de tirage spatial d'un séisme de magnitude 6,5	88

5.7	Répartitions des données d'accélération contenues dans RESORCE	90
5.8	Distribution magnitude- distance des enregistrements utilisés	91
5.9	Nombre cumulé de séismes dans le catalogue d'étude	92
5.10	Proportions de séismes principaux en France	93
5.11	Comparaison de proportions de séismes principaux	96
5.12	Fréquences annuelles de séismes principaux en France	97
5.13	Densité de probabilité des fréquences de séismes principaux	98
5.14	Limitation de la largeur du plan de rupture	99
5.15	DFM obtenues via différentes approches	101
5.16	Impact du seuil MIMP sur la proportion de séismes principaux	102
5.17	Comparaison des DPM d'entrée et de sortie du générateur	104
5.18	Nombre de répliques par séisme principal	105
5.19	Comparaison des DFM d'entrée et de sortie du générateur	106
5.20	Moment sismique cumulé par les sismicités passée et générée	108
5.21	Densités spatiales en entrée et en sortie du générateur	110
5.22	Comparaison de PGA des données et des GMPE	111
5.23	Poids des GMPE	113
5.24	Distributions de PGA obtenues par les GMPE pondérées	115
5.25	Variabilité de la PGA en fonction de la magnitude et de la distance	115
5.26	Comparaison de distributions médianes de PGA	116
5.27	Courbes moyennes probabilité annuelle- aléa	119
5.28	Centiles des PGA aux périodes de retour 475 et 2475 ans	122
5.29	Impact de la régionalisation sur l'estimation de l'aléa	126
5.30	Impact de la prise en compte des répliques sur l'estimation de l'aléa	127
5.31	Impact de la méthode de pondération sur l'estimation de l'aléa	128
5.32	Contribution distance- magnitude de la PGA à 475 ans	130
5.33	Contribution distance- magnitude de la PGA à 2475 ans	131

Liste des Tables

4.1	Coefficients et intervalles de validité de l'Équation 4.1	68
4.2	Spécificités des deux sources possibles dans cette thèse	70
5.1	Magnitudes et années de coupure appliquées dans cette thèse . . .	81
5.2	Magnitude maximale et paramètres de faille des régions GEOTER	86
5.3	Magnitude maximale et paramètres de faille des régions SHARE .	88
5.4	Domaines de validité et type de distance des GMPE utilisées . . .	89
5.5	Proportions de séismes principaux observées dans FCAT	94
5.6	Coefficients b des deux lois de puissance observées	98
5.7	b obtenus à partir de trois DFM différentes	100
5.8	Valeurs b calculées pour différents seuils MIMP	103
5.9	Périodes de retour moyennes en fonction de la magnitude	107
5.10	PGA moyennes à période de retour de 475 ans	118
5.11	Centiles des PGA aux périodes de retour 475 et 2475 ans	121
5.12	Incertitude associée aux PGA aux périodes de retour étudiées . .	124

Liste des Algorithmes

1	Calcul des proportions de séismes principaux	42
2	Recherche de MIMP	43
3	Correction des proportions de séismes principaux	44
4	Générateur de séismes principaux	54
5	Générateur de répliques	61

Bibliographie

- N. A. Abrahamson. State Of The Practice Of Seismic Hazard Evaluation. In *ISRM International Symposium*, 2000. (Cité p. 69)
- S. Akkar, M. A. Sandıkkaya, et J. J. Bommer. Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering*, 12(1): 359–387, 2014a. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-013-9461-4. (Cité p. ix, 36, 67, 89, 112, 138, 139)
- S. Akkar, M. A. Sandıkkaya, M. Şenyurt, A. Azari Sisi, B. Ö. Ay, P. Traversa, J. Douglas, F. Cotton, L. Luzi, B. Hernandez, et S. Godey. Reference database for seismic ground-motion in Europe (RESORCE). *Bulletin of Earthquake Engineering*, 12(1): 311–339, 2014b. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-013-9506-8. (Cité p. 72, 73, 89, 90)
- G. Ameri. Empirical ground motion model adapted to the French context. *Deliverable SIGMA: SIGMA-2014-D2-131*, 2014. (Cité p. ix, 72, 73, 89, 112, 133, 138)
- G. Ameri, D. Baumont, C. Gomes, K. L. Dortz, B. Le, C. Martin, et R. Secanell. On the choice of maximum earthquake magnitude for seismic hazard assessment in metropolitan France – insight from the Bayesian approach. In *9ème Colloque National AFPS*, p. 13, 2015. (Cité p. 47, 52)
- T. D. Ancheta, R. B. Darragh, J. P. Stewart, E. Seyhan, W. J. Silva, B. S.-J. Chiou, K. E. Wooddell, R. W. Graves, A. R. Kottke, et D. M. Boore. NGA-West2 database. *Earthquake Spectra*, 30(3): 989–1005, 2014. (Cité p. 89, 141)
- ASN. Règle fondamentale de sûreté. Technical Report 2001-01, Autorité de Sûreté Nucléaire, Paris, France, 2001. (Cité p. 18)
- K. Assatourians et G. M. Atkinson. EqHaz: An Open-Source Probabilistic Seismic-Hazard Code Based on the Monte Carlo Simulation Approach. *Seismological Research Letters*, 84(3): 516–524, 2013. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220120102. (Cité p. 23, 24)
- T. E. Azak, D. Kalafat, K. Şeşetyan, et M. B. Demircioğlu. Effects of seismic declustering on seismic hazard assessment: A sensitivity study using the Turkish earthquake catalogue. *Bulletin of Earthquake Engineering*, 16(8): 3339–3366, 2017. doi: 10.1007/s10518-017-0174-y. (Cité p. 32)

- S. Baize, E. M. Cushing, F. Lemeille, et H. Jomard. Updated seismotectonic zoning scheme of Metropolitan France, with reference to geologic and seismotectonic data. *Bulletin de la Societe Geologique de France*, 184(3): 225–259, 2013. ISSN 0037-9409. doi: 10.2113/gssgfbull.184.3.225. (Cit e p. 25, 26)
- M. B ath. Lateral inhomogeneities of the upper mantle. *Tectonophysics*, 2(6): 483–514, 1965. ISSN 00401951. doi: 10.1016/0040-1951(65)90003-X. (Cit e p. 41, 43, 55)
- D. Baumont, K. Manchuel, P. Traversa, C. Durouchoux, E. Nayman, et G. Ameri. Intensity predictive attenuation models calibrated in Mw for metropolitan France. *Bulletin of Earthquake Engineering*, 16(6): 2285–2310, 2018. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-018-0344-6. (Cit e p. 15, 80)
- C. Beauval, S. Hainzl, et F. Scherbaum. Probabilistic seismic hazard estimation in low-seismicity regions considering non-Poissonian seismic occurrence. *Geophysical Journal International*, 164(3): 543–550, 2006. (Cit e p. 23, 28, 33, 35, 95, 126, 134, 138)
- C. Beauval, P.-Y. Bard, S. Hainzl, et P. Gueguen. Can strong-motion observations be used to constrain probabilistic seismic-hazard estimates? *Bulletin of the Seismological Society of America*, 98(2): 509–520, 2008. (Cit e p. 37)
- C. Beauval, H. Tasan, A. Laurendeau, E. Delavaud, F. Cotton, P. Gu eguen, et N. Kuehn. On the testing of ground-motion prediction equations against small-magnitude data. *Bulletin of the Seismological Society of America*, 102(5): 1994–2007, 2012. (Cit e p. 37)
- C. Beauval, P.-Y. Bard, et L. Danciu. The influence of source-and ground-motion model choices on probabilistic seismic hazard levels at 6 sites in France. *Bulletin of Earthquake Engineering*, 18(10): 4551–4580, 2020. doi: 10.1007/s10518-020-00879-z. (Cit e p. x, 25, 29, 30, 32, 81, 82, 120, 122, 124, 125, 133, 138)
- M. Bertin, S. Marin, C. Millet, et C. Berge-Thierry. Using Bayesian model averaging to improve ground motion predictions. *Geophysical Journal International*, 220(2): 1368–1378, 2019. ISSN 0956-540X, 1365-246X. doi: 10.1093/gji/ggz486. (Cit e p. 75, 114)
- G. Bertrand, O. Bellier, L. Bollinger, M. Cushing, C. Durouchoux, F. Hollender, B. Meyer, P. Sabourault, A. Schlupp, et M. S ebrier¹⁰. N eoPal: Base de donn ees nationale des d eformations n eotectoniques et des pal eos ismes. In *7 eme Colloque National AFPS*, 2007. (Cit e p. 16)
- L. Bevere, M. Gloor, et A. Sobel. Natural catastrophes in times of economic accumulation and climate change. 2/2020, Swiss Re, Zurich, Switzerland, 2020. (Cit e p. 2)
- D. Bindi, L. Luzi, F. Pacor, F. Sabetta, et M. Massa. Towards a new reference ground motion prediction equation for Italy: Update of the Sabetta–Pugliese (1996). *Bulletin of Earthquake Engineering*, 7(3): 591–608, 2009. (Cit e p. 37)

- D. Bindi, M. Massa, L. Luzi, G. Ameri, F. Pacor, R. Puglia, et P. Augliera. Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5 %-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. *Bulletin of Earthquake Engineering*, 12(1): 391–430, 2014. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-013-9525-5. (Cité p. ix, 89, 112, 138, 139, 140)
- J. Bommer, R. Spence, M. Erdik, S. Tabuchi, N. Aydinoglu, E. Booth, D. del Re, et O. Peterken. Development of an earthquake loss model for Turkish catastrophe insurance. *Journal of Seismology*, 6(3): 431–446, 2002. ISSN 1573-157X. doi: 10.1023/A:1020095711419. (Cité p. 23, 28)
- J. J. Bommer. Deterministic vs. probabilistic seismic hazard assessment: An exaggerated and obstructive dichotomy. *Journal of Earthquake Engineering*, 6(spec01): 43–73, 2002. (Cité p. 6)
- J. J. Bommer, N. A. Abrahamson, F. O. Strasser, A. Pecker, P.-Y. Bard, H. Bungum, F. Cotton, D. Fäh, F. Sabetta, F. Scherbaum, et J. Studer. The challenge of defining upper bounds on earthquake ground motions. *Seismological Research Letters*, 75: 82–95, 2004. doi: 10.1785/gssrl.75.1.82. (Cité p. 69)
- J. J. Bommer, F. Scherbaum, H. Bungum, F. Cotton, F. Sabetta, et N. A. Abrahamson. On the Use of Logic Trees for Ground-Motion Prediction Equations in Seismic-Hazard Analysis. *Bulletin of the Seismological Society of America*, 95(2): 377–389, 2005. doi: 10.1785/0120040073. (Cité p. 24)
- J. J. Bommer, H. Crowley, et R. Pinho. A risk-mitigation approach to the management of induced seismicity. *Journal of Seismology*, 19(2): 623–646, 2015. ISSN 1383-4649, 1573-157X. doi: 10.1007/s10950-015-9478-z. (Cité p. 23)
- D. M. Boore, J. P. Stewart, E. Seyhan, et G. M. Atkinson. NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthquake Spectra*, 30(3): 1057–1085, 2014. doi: 10.1193/070113EQS184M. (Cité p. ix, 89, 112, 138, 140)
- S. J. Bourne, S. J. Oates, J. J. Bommer, B. Dost, J. van Elk, et D. Doornhof. A Monte Carlo Method for Probabilistic Hazard Assessment of Induced Seismicity due to Conventional Natural Gas Production. *Bulletin of the Seismological Society of America*, 105(3): 1721–1738, 2015. ISSN 0037-1106, 1943-3573. doi: 10.1785/0120140302. (Cité p. 23)
- M. Burkhard et G. Grünthal. Seismic source zone characterization for the seismic hazard assessment project PEGASOS by the Expert Group 2 (EG1b). *Swiss Journal of Geosciences*, 102(1): 149–188, 2009. ISSN 1661-8726, 1661-8734. doi: 10.1007/s00015-009-1307-3. (Cité p. 17, 95, 96)
- E. Calais, T. Camelbeeck, S. Stein, M. Liu, et T. J. Craig. A new paradigm for large earthquakes in stable continental plate interiors: LARGE EARTHQUAKES IN SCRs. *Geophysical Research Letters*, 43(20): 10,621–10,637, 2016. ISSN 00948276. doi: 10.1002/2016GL070815. (Cité p. 4, 6, 7, 82, 137)

- M. Cara, Y. Cansi, A. Schlupp, P. Arroucau, N. Béthoux, E. Beucler, S. Bruno, M. Calvet, S. Chevrot, et A. Deboissy. SI-Hex: A new catalogue of instrumental seismicity for metropolitan France. *Bulletin de la Société Géologique de France*, 186(1): 3–19, 2015. doi: 10.2113/gssgfbull.186.1.3. (Cit e p. vii, 79)
- C. Cauzzi, E. Faccioli, M. Vanini, et A. Bianchini. Updated predictive equations for broadband (0.01–10 s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records. *Bulletin of Earthquake Engineering*, 13(6): 1587–1612, 2015. doi: 10.1007/s10518-014-9685-y. (Cit e p. ix, 89, 112, 120)
- T. Chartier. *Mod elisation de l’activit e sismique des failles pour le calcul probabiliste du risque sismique*. PhD thesis, Universit e PSL, Paris, France, 2019. (Cit e p. 20)
- A. Christophersen et E. G. Smith. Foreshock rates from aftershock abundance. *Bulletin of the Seismological Society of America*, 98(5): 2133–2148, 2008. (Cit e p. 31)
- A. Christophersen, M. C. Gerstenberger, D. A. Rhoades, et M. W. Stirling. Quantifying the effect of declustering on probabilistic seismic hazard. In *9th Pacific Conference on Earthquake Engineeing*, p. 8, 2011. (Cit e p. 32)
- R. Console et M. Murru. A simple and testable model for earthquake clustering. *Journal of Geophysical Research: Solid Earth*, 106(B5): 8699–8711, 2001. ISSN 2156-2202. doi: 10.1029/2000JB900269. (Cit e p. 35)
- R. Console, A. M. Lombardi, M. Murru, et D. Rhoades. B ath’s law and the self-similarity of earthquakes: B ATH’S LAW. *Journal of Geophysical Research: Solid Earth*, 108(B2), 2003. ISSN 01480227. doi: 10.1029/2001JB001651. (Cit e p. 55, 56)
- K. J. Coppersmith et R. R. Youngs. Capturing uncertainty in probabilistic seismic hazard assessments within intraplate tectonic environments. In *Proceedings of the Third US National Conference on Earthquake Engineering*, vol. 1, p. 301–312, 1986. (Cit e p. 19)
- A. Cornell. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5): 1583–1606, 1968. (Cit e p. 5, 19, 20, 21, 25, 40, 75)
- C. A. Cornell et E. H. Vanmarcke. The major influences on seismic risk. In *Proceedings of the Fourth World Conference on Earthquake Engineering*, vol. 1, p. 69–83, 1969. (Cit e p. 29)
-  . Corral. Long-Term Clustering, Scaling, and Universality in the Temporal Occurrence of Earthquakes. *Physical Review Letters*, 92(10): 108501, 2004. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.92.108501. (Cit e p. 32, 33)
-  . Corral et  . Gonz alez. Power Law Size Distributions in Geoscience Revisited. *Earth and Space Science*, p. 2018EA000479, 2019. ISSN 2333-5084, 2333-5084. doi: 10.1029/2018EA000479. (Cit e p. 31, 99, 136)
- F. Cotton, F. Scherbaum, J. J. Bommer, et H. Bungum. Criteria for selecting and adjusting ground-motion models for specific target regions: Application to central Europe and rock sites. *Journal of Seismology*, 10(2): 137–156, 2006. (Cit e p. 37)

- F. Cotton, G. Pousse, F. Bonilla, et F. Scherbaum. On the discrepancy of recent European ground-motion observations and predictions from empirical models: Analysis of KiK-net accelerometric data and point-sources stochastic simulations. *Bulletin of the Seismological Society of America*, 98(5): 2244–2261, 2008. (Cit e p. 37)
- H. Crowley et J. J. Bommer. Modelling Seismic Hazard in Earthquake Loss Models with Spatially Distributed Exposure. *Bulletin of Earthquake Engineering*, 4(3): 249–273, 2006. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-006-9009-y. (Cit e p. 28)
- E. M. Cushing, O. Bellier, S. Nechtschein, M. S ebrier, A. Lomax, P. Volant, P. Dervin, P. Guignard, et L. Bove. A multidisciplinary study of a slow-slipping fault for seismic hazard assessment: The example of the Middle Durance Fault (SE France). *Geophysical Journal International*, 172(3): 1163–1178, 2008. (Cit e p. 28)
- E. Delavaud, F. Cotton, S. Akkar, F. Scherbaum, L. Danciu, C. Beauval, S. Drouet, J. Douglas, R. Basili, M. A. Sandikkaya, M. Segou, E. Faccioli, et N. Theodoulidis. Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe. *Journal of Seismology*, 16(3): 451–473, 2012. ISSN 1383-4649, 1573-157X. doi: 10.1007/s10950-012-9281-z. (Cit e p. 66, 127, 128, 137)
- D. Diderot et J. L. R. D’Alembert. *Encyclop die ou Dictionnaire raisonn  des sciences, des arts et des m tiers*. 1751. (Cit e p. 1)
- J. Douglas. Ground motion prediction equations 1964-2019. Technical report, Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, United Kingdom, 2019. (Cit e p. 37)
- J. Douglas, T. Ulrich, D. Bertil, et J. Rey. Comparison of the Ranges of Uncertainty Captured in Different Seismic-Hazard Studies. *Seismological Research Letters*, 85(5): 977–985, 2014. ISSN 0895-0695. doi: 10.1785/0220140084. (Cit e p. ix, 123)
- S. Drouet et F. Cotton. Regional stochastic GMPEs in low-seismicity areas: Scaling and aleatory variability analysis—Application to the French Alps. *Bulletin of the Seismological Society of America*, 105(4): 1883–1902, 2015. doi: 10.1785/0120140240. (Cit e p. ix, 36, 66, 72, 73, 89, 112, 133, 138)
- S. Drouet, G. Ameri, K. Le Dortz, R. Secanell, et G. Senfaute. A probabilistic seismic hazard map for the metropolitan France. *Bulletin of Earthquake Engineering*, 18(5): 1865–1898, 2020. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-020-00790-7. (Cit e p. 25, 26, 27, 32, 47, 53, 85, 120, 133, 138, 140)
- J. E. Ebel et A. L. Kafka. A Monte Carlo approach to seismic hazard analysis. *Bulletin of the Seismological Society of America*, 89(4): 854–866, 1999. (Cit e p. 23, 51)
- N. EN. Eurocodes 8—Calcul des structures pour leur r sistance aux s ismes—Partie 1: R gles g n rales, actions sismiques et r gles pour les b timents. Technical Report 0335-3931, ISSN, 2005. (Cit e p.)
- L. Faenza, S. Hainzl, F. Scherbaum, et C. Beauval. Statistical analysis of time-dependent earthquake occurrence and its impact on hazard in the low seismicity region Lower

- Rhine Embayment. *Geophysical Journal International*, 171(2): 797–806, 2007. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2007.03564.x. (Cité p. 34, 126, 134, 138)
- N. Feuillet, I. Manighetti, et P. Tapponnier. Extension active perpendiculaire à la subduction dans l’arc des Petites Antilles (Guadeloupe, Antilles françaises). *Comptes Rendus de l’Académie des Sciences-Series IIA-Earth and Planetary Science*, 333(9): 583–590, 2001. (Cité p. 28)
- N. Feuillet, I. Manighetti, P. Tapponnier, et E. Jacques. Arc parallel extension and localization of volcanic complexes in Guadeloupe, Lesser Antilles: EXTENSIONAL ACTIVE FAULTING IN GUADELOUPE. *Journal of Geophysical Research: Solid Earth*, 107(B12): ETG 3–1–ETG 3–29, 2002. ISSN 01480227. doi: 10.1029/2001JB000308. (Cité p. 28)
- A. D. Frankel, C. Mueller, T. Barnhard, D. Perkins, E. Leyendecker, N. Dickman, S. Hanson, et M. Hopper. *National Seismic-Hazard Maps: Documentation June 1996*. US Geological Survey Reston, VA, 1996. (Cité p. 27)
- N. A. Galina, V. V. Bykova, R. N. Vakarchuk, et R. E. Tatevosian. Effect of Earthquake Catalog Declustering on Seismic Hazard Assessment. *Seismic Instruments*, 55(1): 59–69, 2019. doi: 10.3103/S0747923919010079. (Cité p. 32)
- J. Gardner et L. Knopoff. Is sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America*, 64(5): 1363–1367, 1974. (Cité p. 17, 59, 82)
- R. J. Geller. Earthquake prediction: A critical review. *Geophysical Journal International*, 131(3): 425–450, 1997. (Cité p. 5)
- D. Giardini, S. Wiemer, D. Fäh, et N. Deichmann. *Seismic Hazard Assessment of Switzerland, 2004*. ETH Zurich, Switzerland, swiss seismological service edition, 2004. (Cité p. ix, 123)
- K. Goda et R. De Risi. Probabilistic Tsunami Loss Estimation Methodology: Stochastic Earthquake Scenario Approach. *Earthquake Spectra*, 33(4): 1301–1323, 2017. ISSN 8755-2930, 1944-8201. doi: 10.1193/012617eqs019m. (Cité p. 23)
- C. Gouache et P. Tinard. Estimation de l’aléa sismique en France. In *11ème Journée CCR Cat*, 2020. (Cité p. 12)
- C. Gouache, F. Bonneau, et P. Tinard. Representative seismic zone-dependent catalogue for France. In *RING Meeting 2018*, 2018. (Cité p. 12)
- C. Gouache, F. Bonneau, et P. Tinard. Regional and exhaustive French seismic catalogues. In *20th Annual Conference IAMG*, 2019a. (Cité p. 12)
- C. Gouache, F. Bonneau, et P. Tinard. Stochastic estimation of French annual main shock frequencies. In *RING Meeting 2019*, 2019b. (Cité p. 12)
- C. Gouache, P. Cupillard, et P. Tinard. Calibration de lois d’atténuation par simulation de propagation d’ondes sismiques. In *10ème Colloque national AFPS*, 2019c. (Cité p. 12)

- C. Gouache, P. Tinard, et J.-M. Montel. Génération stochastique de catalogues sismiques pour le territoire français. In *10ème Colloque national AFPS*, 2019d. (Cité p. 12)
- C. Gouache, P. Cupillard, et P. Tinard. Enriching the RESORCE database by synthetic waveforms to weight Ground-Motion Prediction Equations for seismic hazard assessment. In *RING Meeting 2020*, 2020a. (Cité p. 12)
- C. Gouache, P. Tinard, F. Bonneau, et J.-M. Montel. Stochastic generator of earthquakes in French mainland. In *EGU General Assembly 2020*, 2020b. doi: 10.5194/egusphere-egu2020-5150. (Cité p. 12)
- C. Gouache, F. Bonneau, P. Tinard, et J. M. Montel. Estimation of main shock frequency–magnitude distributions by adapting the inter-event time method for low-to-moderate seismicity areas: Application to French mainland. *Journal of Seismology*, 2021. ISSN 1573-157X. doi: 10.1007/s10950-021-10001-8. (Cité p. 12, 92, 136)
- G. Grünthal. The up-dated earthquake catalogue for the German democratic Republic and adjacent areas—statistical data characteristics and conclusions for hazard assessment. In *Proc. 3rd Int. Symp. on the Analysis of Seismicity and Seismic Risk*, 1985. (Cité p. x, 17, 31, 59, 95, 96, 106, 107)
- G. Grünthal et R. Wahlström. The European-Mediterranean earthquake catalogue (EMEC) for the last millennium. *Journal of seismology*, 16(3): 535–570, 2012. (Cité p.)
- G. Grünthal, D. Stromeyer, C. Bosse, F. Cotton, et D. Bindi. The probabilistic seismic hazard assessment of Germany—version 2016, considering the range of epistemic uncertainties and aleatory variability. *Bulletin of Earthquake Engineering*, 16(10): 4339–4395, 2018. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-018-0315-y. (Cité p. 62, 69, 120)
- Z. Guo et Y. Ogata. Statistical relations between the parameters of aftershocks in time, space, and magnitude. *Journal of Geophysical Research: Solid Earth*, 102(B2): 2857–2873, 1997. ISSN 01480227. doi: 10.1029/96JB02946. (Cité p. 105)
- B. Gutenberg et C. F. Richter. Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4): 185–188, 1944. (Cité p. viii, 5, 16, 28)
- C. Guyonnet-Benaize. *Modélisation 3D Multi-Échelle Des Structures Géologiques de La Région de La Faille de La Moyenne Durance (SE France)*. PhD Thesis, Aix-Marseille 1, 2011. (Cité p. 28)
- C. Guyonnet-Benaize, J. Lamarche, F. Hollender, S. Viseur, P. Münch, et J. Borgomano. Three-dimensional structural modeling of an active fault zone based on complex outcrop and subsurface data: The Middle Durance Fault Zone inherited from polyphase Meso-Cenozoic tectonics (southeastern France). *Tectonics*, 34(2): 265–289, 2015. (Cité p. 28)

- S. Hainzl, F. Scherbaum, et C. Beauval. Estimating Background Activity Based on Interevent-Time Distribution. *Bulletin of the Seismological Society of America*, 96(1): 313–320, 2006. ISSN 0037-1106. doi: 10.1785/0120050053. (Cité p. 33, 40, 95)
- S. Hainzl, G. B. Brietzke, et G. Zöller. Quantitative earthquake forecasts resulting from static stress triggering. *Journal of Geophysical Research*, 115(B11): B11311, 2010. ISSN 0148-0227. doi: 10.1029/2010JB007473. (Cité p. 35, 59)
- S.-W. Han et Y.-S. Choi. Seismic hazard analysis in low and moderate seismic region-Korean peninsula. *Structural Safety*, 30(6): 543–558, 2008. (Cité p. 23, 28)
- T. C. Hanks et H. Kanamori. A moment magnitude scale. *Journal of Geophysical Research: Solid Earth*, 84(B5): 2348–2350, 1979. doi: 10.1029/JB084iB05p02348. (Cité p. 45, 55)
- O. Heidbach, A. Barth, B. Müller, J. Reinecker, O. Stephansson, M. Tingay, et A. Zang. WSM quality ranking scheme, database description and analysis guidelines for stress indicator. Technical Report 16-01, 2016. (Cité p. 85)
- A. Helmstetter. Is Earthquake Triggering Driven by Small Earthquakes? *Physical Review Letters*, 91(5): 058501, 2003. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.91.058501. (Cité p. 35, 105)
- A. Helmstetter et D. Sornette. Bâth’s law derived from the Gutenberg-Richter law and from aftershock properties. *Geophysical Research Letters*, 30(20), 2003a. ISSN 1944-8007. doi: 10.1029/2003GL018186. (Cité p. 55, 56)
- A. Helmstetter et D. Sornette. Importance of direct and indirect triggered seismicity in the ETAS model of seismicity. *Geophysical Research Letters*, 30(11), 2003b. ISSN 1944-8007. doi: 10.1029/2003GL017670. (Cité p. 35)
- I. Iervolino, E. Chioccarelli, et M. Giorgio. Aftershocks’ Effect on Structural Design Actions in Italy. *Bulletin of the Seismological Society of America*, 108(4): 2209–2220, 2018. ISSN 0037-1106. doi: 10.1785/0120170339. (Cité p. 126, 127, 134)
- K. Jacob, J. Armbruster, N. Barstow, et S. Horton. Probabilistic ground motion estimates for New York: Comparison with design ground motions in national and local codes. In *Proceedings of Fifth US National Conference on Earthquake Engineering, Chicago*, vol. 3, p. 119–128, 1994. (Cité p. 27)
- H. Jomard, E. M. Cushing, L. Palumbo, S. Baize, C. David, et T. Chartier. Transposing an active fault database into a seismic hazard fault model for nuclear facilities – Part 1: Building a database of potentially active faults (BDFa) for metropolitan France. *Natural Hazards and Earth System Sciences*, 17(9): 1573–1584, 2017. ISSN 1684-9981. doi: 10.5194/nhess-17-1573-2017. (Cité p. 28)
- Y. Y. Kagan. *Earthquakes: Models, Statistics, Testable Forecasts*. American Geophysical Union, Washington, DC, 2013. (Cité p. 5)

- Y. Y. Kagan et L. Knopoff. Stochastic synthesis of earthquake catalogs. *Journal of Geophysical Research: Solid Earth*, 86(B4): 2853–2862, 1981. (Cité p. 34)
- H. Kanamori et E. E. Brodsky. The physics of earthquakes. *Reports on Progress in Physics*, 67(8): 1429–1496, 2004. ISSN 0034-4885, 1361-6633. doi: 10.1088/0034-4885/67/8/R03. (Cité p. 4)
- A. Kijko. Estimation of the maximum earthquake magnitude, m_{max} . *Pure and Applied Geophysics*, 161(8): 1655–1681, 2004. (Cité p. 47, 139)
- S. L. Kramer. *Geotechnical Earthquake Engineering*. Prentice Hall, 1996. (Cité p. 5, 18)
- E. L. Krinitzky. How to obtain earthquake ground motions for engineering design. *Engineering Geology*, 65(1): 1–16, 2002. ISSN 0013-7952. doi: 10.1016/S0013-7952(01)00098-9. (Cité p. 18)
- R. B. Kulkarni, R. R. Youngs, et K. J. Coppersmith. Assessment of confidence intervals for results of seismic hazard analysis. In *Proceedings of the Eighth World Conference on Earthquake Engineering*, vol. 1, p. 263–270, 1984. (Cité p. 19)
- G. Lanzano, L. Luzi, E. Russo, C. Felicetta, M. C. D’Amico, S. Sgobba, et F. Pacor. Engineering Strong Motion Database (ESM) flatfile, 2018. (Cité p. 139)
- K. Le Dortz, P. Combes, et D. Carbon. An alternative seismotectonic zonation for probabilistic and deterministic seismic hazard assessment for Metropolitan France. In *10ème Colloque National AFPS*, 2019. (Cité p. 25, 26)
- M. Leonard. Self-Consistent Earthquake Fault-Scaling Relations: Update and Extension to Stable Continental Strike-Slip Faults. *Bulletin of the Seismological Society of America*, 104(6): 2953–2965, 2014. ISSN 0037-1106. doi: 10.1785/0120140087. (Cité p. 29, 59, 60, 67, 68, 70)
- M. Liu, S. Stein, et H. Wang. 2000 years of migrating earthquakes in North China: How earthquakes in midcontinents differ from those at plate boundaries. *Lithosphere*, 3(2): 128–132, 2011. ISSN 1941-8264. doi: 10.1130/L129.1. (Cité p. 6, 82, 137)
- C. Louarn, F. Bonneau, et C. Gouache. A web application to analyze seismic catalog. In *RING Meeting 2019*, 2019. (Cité p. 12)
- B. Luen et P. B. Stark. Poisson tests of declustered catalogues: Poisson tests. *Geophysical Journal International*, 189(1): 691–700, 2012. ISSN 0956540X. doi: 10.1111/j.1365-246X.2012.05400.x. (Cité p. 31, 32)
- S. Mak, F. Cotton, et D. Schorlemmer. *Measuring the Performance of Ground-Motion Models: The Importance of Being Independent*. Seismological Society of America, 2017. (Cité p. 66, 112, 114, 139)
- K. C. Makropoulos et P. W. Burton. Seismic risk of circum-pacific earthquakes i. strain energy release. *pure and applied geophysics*, 121(2): 247–267, 1983. (Cité p. 47)

- K. Manchuel, P. Traversa, D. Baumont, M. Cara, E. Nayman, et C. Durouchoux. The French seismic CATalogue (FCAT-17). *Bulletin of Earthquake Engineering*, 16: 2227–2251, 2017. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-017-0236-1. (Cit   p. vii, 29, 45, 46, 80)
- D. Marsan et O. Lengline. Extending earthquakes’ reach through cascading. *Science*, 319(5866): 1076–1079, 2008. doi: 10.1126/science.1148783. (Cit   p. 17)
- C. Martin et R. Secanell. D  veloppement d’un modele probabiliste d’al  a sismique cal   sur le retour d’exp  rience. Phase 2: Calculs et cartographie suivant l’arbre logique d  fini par le groupe «zonage». Technical Report GTR/CEA/0306-294, GEOTER, 2006. (Cit   p. 119)
- C. Martin, P. Combes, R. Secanell, G. Lignon, A. Fioravanti, D. Carbon, O. Monge, et B. Grellet. R  vision du zonage sismique de la France. Technical Report GTR/MATE/0701-150, 2002. (Cit   p. 25, 26, 84, 119)
- C. Martin, G. Ameri, D. Baumont, D. Carbon, G. Senfaute, J.-M. Thiry, E. Faccioli, et J. Savy. Probabilistic seismic hazard assessment for South-Eastern France. *Bulletin of Earthquake Engineering*, 16(6): 2477–2511, 2018. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-017-0249-9. (Cit   p. x, 28, 89, 91, 112, 115, 116, 119, 120, 127, 128, 133, 134, 138)
- S. Mazzotti, H. Jomard, et F. Masson. Processes and deformation rates generating seismicity in metropolitan France and conterminous Western Europe. *Bulletin de la Soci  t   G  ologique de France*, 191(1): 19, 2020. ISSN 1777-5817. doi: 10.1051/bsgf/2020019. (Cit   p. 6, 82)
- S. Mazzotti, C. Aubagnac, L. Bollinger, K. C. Oscanoa, B. Delouis, D. D. Paco, C. Doubre, M. Godano, H. Jomard, C. Larroque, A. Laurendeau, F. Masson, M. Sylvander, et A. Trilla. FMHex20: An earthquake focal mechanism database for seismotectonic analyses in metropolitan France and bordering regions. *BSGF - Earth Sciences Bulletin*, 192: 10, 2021. ISSN 1777-5817. doi: 10.1051/bsgf/2020049. (Cit   p. 139)
- R. K. McGuire. Probabilistic seismic hazard analysis and design earthquakes: Closing the loop. *Bulletin of the Seismological Society of America*, 85(5): 1275–1284, 1995. (Cit   p. 23)
- A. Mignan et J. Woessner. Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis*, p. 1–45, 2012. doi: 10.5078/corssa-00180805. (Cit   p. 16)
- K. R. Milner, B. E. Shaw, C. A. Goulet, K. B. Richards-Dinger, S. Callaghan, T. H. Jordan, J. H. Dieterich, et E. H. Field. Toward Physics-Based Nonergodic PSHA: A Prototype Fully Deterministic Seismic Hazard Model for Southern California. *Bulletin of the Seismological Society of America*, 2021. doi: 10.1785/0120200216. (Cit   p. 23, 28)

- L. Mizrahi, S. Nandan, et S. Wiemer. The Effect of Declustering on the Size Distribution of Mainshocks. *Seismological Research Letters*, 2020. (Cité p. 102)
- G. Molchan. Interevent time distribution in seismicity: A theoretical approach. *Pure and Applied Geophysics*, 162(6-7): 1135–1150, 2004. ISSN 0033-4553, 1420-9136. doi: 10.1007/s00024-004-2664-5. (Cité p. 32, 33, 40, 94, 96, 135)
- P. Mouroux, B. Le Brun, S. Depinois, E. Bertrand, et P. Masure. Projet européen RISK-UE: application à la ville de Nice. Technical Report BRGM/RP-53202-FR, 2004. (Cité p. 3)
- R. M. Musson. On the use of Monte Carlo simulations for seismic hazard assessment. In *6th U.S. National Conference on Earthquake Engineering*, p. 12, 1998. (Cité p. 21, 22, 34)
- R. M. W. Musson. Determination of Design Earthquakes in Seismic Hazard Analysis Through Monte Carlo Simulation. *Journal of Earthquake Engineering*, 3(4): 463–474, 1999. ISSN 1363-2469. doi: 10.1080/13632469909350355. (Cité p. 21, 23, 34, 40, 51)
- R. M. W. Musson. The Effect of Magnitude Uncertainty on Earthquake Activity Rates. *Bulletin of the Seismological Society of America*, 102(6): 2771–2775, 2012. ISSN 0037-1106. doi: 10.1785/0120110224. (Cité p. 22, 23)
- H. N. Nguyen, P. Vernant, S. Mazzotti, G. Khazaradze, et E. Asensio. 3-D GPS velocity field and its implications on the present-day post-orogenic deformation of the Western Alps and Pyrenees. *Solid Earth*, 7(5): 1349–1363, 2016. ISSN 1869-9510. doi: 10.5194/se-7-1349-2016. (Cité p. 6)
- O. Nicolis, M. Chiodi, et G. Adelfio. Windowed ETAS models with application to the Chilean seismic catalogs. *Spatial Statistics*, 14: 151–165, 2015. ISSN 22116753. doi: 10.1016/j.spasta.2015.05.006. (Cité p. 105)
- C. I. Nieves, J. J. Bommer, H. Crowley, J. van Elk, M. Ntinalexis, et M. Sangirardi. A database of damaging small-to-medium magnitude earthquakes. *Journal of Seismology*, 24(2): 263–292, 2020. ISSN 1383-4649, 1573-157X. doi: 10.1007/s10950-019-09897-0. (Cité p. 37)
- J.-M. Nocquet. Present-day kinematics of the Mediterranean: A comprehensive overview of GPS results. *Tectonophysics*, 579: 220–242, 2012. ISSN 0040-1951. doi: 10.1016/j.tecto.2012.03.037. (Cité p. 6)
- Y. Ogata. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of the American Statistical Association*, 83(401): 9–27, 1988. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.1988.10478560. (Cité p. ix, 34)
- Y. Ogata. Space-Time Point-Process Models for Earthquake Occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2): 379–402, 1998. ISSN 1572-9052. doi: 10.1023/A:1003403601725. (Cité p. 35)
- F. Omori. *On the After-Shocks of Earthquakes*, vol. 7 of *Seismological Journal of Japan*. The University of Tokyo, 1894. (Cité p. ix, 5, 35)

- J. F. Pacheco, C. H. Scholz, et L. R. Sykes. Changes in frequency–size relationship from small to large earthquakes. *Nature*, 355(6355): 71–73, 1992. (Cit e p. 99, 100, 136)
- P. Reasenber. Second-order moment of central California seismicity, 1969-1982. *Journal of Geophysical Research: Solid Earth*, 90(B7): 5479–5495, 1985. ISSN 01480227. doi: 10.1029/JB090iB07p05479. (Cit e p. 17, 31, 32, 33, 59, 82, 95)
- L. Reiter. *Earthquake Hazard Analysis: Issues and Insights*, vol. 22. Columbia University Press, New York, 1990. (Cit e p. 5, 18)
- J. Rey et P. Tinard. Estimation financi re du risque sismique   l’ chelle d partementale:   l’interface entre sismologie et r assurance, travaux communs CCR-BRGM (2014-2019). In *9 me Colloque de l’AFPS : AFPS15 : "Anticiper, limiter, g rer les effets des s ismes dans les territoires"*, p. 12, 2015. (Cit e p. 3)
- J. Rey et P. Tinard. Partenariat CCR-BRGM 2014-2019. Travaux men s en 2016 dans le cadre du programme commun sur le risque sismique. Technical Report BRGM/RP-66473-FR, BRGM, 2016. (Cit e p. 3)
- J. Rey et P. Tinard. Partenariat CCR-BRGM 2014-2019. Travaux men s en 2017 dans le cadre du programme commun sur le risque sismique. Technical Report BRGM/RP-68082-FR, BRGM, 2017. (Cit e p. 3)
- C. F. Richter. *Elementaty Seismology*. Number BOOK. W.F. Freeman and Company, San Francisco, and Bailey Bros. & Swinfen Ltd., London, 1958. (Cit e p. 41, 43, 55, 56)
- J.-F. Ritz, S. Baize, M. Ferry, C. Larroque, L. Audin, B. Delouis, et E. Mathot. Surface rupture and shallow fault reactivation during the 2019 Mw 4.9 Le Teil earthquake, France. *Communications Earth & Environment*, 1(1): 10, 2020. ISSN 2662-4435. doi: 10.1038/s43247-020-0012-z. (Cit e p. 6, 82)
- R. Romeo et A. Prestininzi. Probabilistic versus deterministic seismic hazard analysis: An integrated approach for siting problems. *Soil Dynamics and Earthquake Engineering*, 20(1-4): 75–84, 2000. (Cit e p. 23)
- W. Rosenhauer. Methodological aspects encountered in the Lower Rhine area seismic hazard analysis. In *Seismicity and Seismic Risk in the Offshore North Sea Area*, p. 385–396, 1983. (Cit e p. 23)
- J.-J. Rousseau. Lettre   Monsieur de Voltaire sur ses deux po mes sur « la Loi naturelle » et sur « le D sastre de Lisbonne », 1756. (Cit e p. 1)
- J. B. Rundle. Derivation of the complete Gutenberg-Richter magnitude-frequency relation using the principle of scale invariance. *Journal of Geophysical Research: Solid Earth*, 94(B9): 12337–12342, 1989. (Cit e p. 99, 136)
- M. K. Sachs, M. R. Yoder, D. L. Turcotte, J. B. Rundle, et B. D. Malamud. Black swans, power laws, and dragon-kings: Earthquakes, volcanic eruptions, landslides, wildfires, floods, and SOC models. *The European Physical Journal Special Topics*, 205(1):

- 167–182, 2012. ISSN 1951-6355, 1951-6401. doi: 10.1140/epjst/e2012-01569-3. (Cité p. 30, 31, 99, 136)
- F. Scherbaum, F. Cotton, et P. Smit. On the use of response spectral-reference data for the selection and ranking of ground-motion models for seismic-hazard analysis in regions of moderate seismicity: The case of rock motion. *Bulletin of the seismological society of America*, 94(6): 2164–2185, 2004. (Cité p. 75)
- F. Scherbaum, E. Delavaud, et C. Riggelsen. Model selection in seismic hazard analysis: An information-theoretic perspective. *Bulletin of the Seismological Society of America*, 99(6): 3234–3247, 2009. (Cité p. 75)
- D. Schorlemmer, S. Wiemer, et M. Wyss. Variations in earthquake-size distribution across different stress regimes. *Nature*, 437(7058): 539, 2005. (Cité p. 99)
- P. Shahbazi et B. Mansouri. Grid Source Event-Based Seismic Hazard Assessment of Iran. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, p. 1–11, 2019. (Cité p. 23, 28)
- P. Shearer. Southern California Hypocenter Relocation with Waveform Cross-Correlation, Part 2: Results Using Source-Specific Station Terms and Cluster Analysis. *Bulletin of the Seismological Society of America*, 95(3): 904–915, 2005. ISSN 0037-1106. doi: 10.1785/0120040168. (Cité p. 59)
- V. Sokolov, H. M. Zahran, S. Youssef, M. El-Hadidy, et W. W. Alraddadi. Probabilistic seismic hazard assessment for Saudi Arabia using spatially smoothed seismicity and analysis of hazard uncertainty. *Bulletin of Earthquake Engineering*, 15(7): 2695–2735, 2016. doi: 10.1007/s10518-016-0075-5. (Cité p. 32)
- D. Sornette, L. Knopoff, Y. Y. Kagan, et C. Vanneste. Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes. *Journal of Geophysical Research: Solid Earth*, 101(B6): 13883–13893, 1996. (Cité p. 31, 99)
- S. Stein, M. Liu, T. Camelbeeck, M. Merino, A. Landgraf, E. Hintersberger, et S. Kübler. Challenges in assessing seismic hazard in intraplate Europe. *Geological Society, London, Special Publications*, 432(1): 13–28, 2015. ISSN 0305-8719, 2041-4927. doi: 10.1144/SP432.7. (Cité p. 4, 6, 7, 28, 82, 137)
- M. Stucchi, A. Rovida, A. A. Gomez Capera, P. Alexandre, T. Camelbeeck, M. B. Demircioglu, P. Gasperini, V. Kouskouna, R. M. W. Musson, M. Radulian, K. Sesetyan, S. Vilanova, D. Baumont, H. Bungum, D. Fäh, W. Lenhardt, K. Makropoulos, J. M. Martinez Solares, O. Scotti, M. Živčić, P. Albini, J. Batllo, C. Papaioannou, R. Tatevossian, M. Locati, C. Meletti, D. Viganò, et D. Giardini. The SHARE European Earthquake Catalogue (SHEEC) 1000–1899. *Journal of Seismology*, 17(2): 523–544, 2013. ISSN 1383-4649, 1573-157X. doi: 10.1007/s10950-012-9335-2. (Cité p. vii, 15, 29)
- M. Terrier, C. Carnec, C. Champion, P. Combes, H. Fabriol, J. Gourry, C. Innocent, B. Lebrun, et D. Raucoules. FAILLES ACTIVES ET EVALUATION DE L’ALEA SISMIQUE : Prise en compte des failles actives dans l’aménagement du territoire

- aux Antilles (Martinique et Guadeloupe). Partie 3 : Des méthodes et des outils pour l'identification des failles actives. Technical Report BRGM/RP-51564-FR, BRGM, 2002. (Cité p. 28)
- M. Terrier, D. Bertil, et J. Rohmer. Méthode d'identification des failles actives en domaine de déformation lente. Technical Report BRGM/RP-68 553-FR, BRGM, 2018. (Cité p. 28)
- S. Tinti et F. Mulargia. Completeness analysis of a seismic catalog. In *Annales Geophysicae (1983)*, vol. 3, p. 407–414, 1985. (Cité p. 81)
- X. Tong, B. Smith-Konter, et D. T. Sandwell. Is there a discrepancy between geological and geodetic slip rates along the San Andreas Fault System? *Journal of Geophysical Research: Solid Earth*, 119(3): 2518–2538, 2014. ISSN 2169-9356. doi: 10.1002/2013JB010765. (Cité p. 5)
- P. Traversa, D. Baumont, K. Manchuel, E. Nayman, et C. Durouchoux. Exploration tree approach to estimate historical earthquakes Mw and depth, test cases from the French past seismicity. *Bulletin of Earthquake Engineering*, 16: 2169–2193, 2017. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-017-0178-7. (Cité p. 15, 45, 80)
- R. A. Uhrhammer. Characteristics of northern and central California seismicity. *Earthquake Notes*, 57(1): 21, 1986. (Cité p. 17)
- T. Utsu. Magnitudes of earthquakes and occurrence of their aftershocks. *Zisin, Ser. 2*, 10: 35–45, 1957. (Cité p. ix, 5, 35)
- T. Utsu. Aftershocks and Earthquake Statistics(1) : Some Parameters Which Characterize an Aftershock Sequence and Their Interrelations. *Journal of the Faculty of Science, Hokkaido University. Series 7, Geophysics*, 3(3): 129–195, 1970. (Cité p. 35, 105)
- T. Utsu, Y. Ogata, R. S, et Matsu'ura. The Centenary of the Omori Formula for a Decay Law of Aftershock Activity. *Journal of Physics of the Earth*, 43(1): 1–33, 1995. ISSN 1884-2305, 0022-3743. doi: 10.4294/jpe1952.43.1. (Cité p. 35)
- T. van Stiphout, J. Zhuang, et D. Marsan. Seismicity Declustering. *Community Online Resource for Statistical Seismicity Analysis*, 2012. doi: 10.5078/corssa52382934. (Cité p. 31, 32, 95)
- G. Vannucci, P. Imprescia, et P. Gasperini. Earthquake Mechanisms of the Mediterranean Area (EMMA) version 3: An improved tool for characterizing the tectonic deformation styles in the Mediterranean. In *EGU 2009*, p. 1, 2009. (Cité p. 85)
- D. Vere-Jones. How to educate yourself as a statistical seismologist. *Community Online Resource for Statistical Seismicity Analysis*, p. 1–14, 2010. doi: 10.5078/CORSSA-17728079. (Cité p. 5)
- D. L. Wells et K. J. Coppersmith. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the seismological Society of America*, 84(4): 974–1002, 1994. (Cité p. 29, 59)

- J. Woessner, D. Laurentiu, D. Giardini, H. Crowley, F. Cotton, G. Grünthal, G. Valensise, R. Arvidsson, R. Basili, M. B. Demircioglu, S. Hiemer, C. Meletti, R. W. Musson, A. N. Rovida, K. Sesetyan, et M. Stucchi. The 2013 European Seismic Hazard Model: Key components and results. *Bulletin of Earthquake Engineering*, 13(12): 3553–3596, 2015. ISSN 1570-761X, 1573-1456. doi: 10.1007/s10518-015-9795-1. (Cité p. 28, 53, 69, 84, 120, 123, 124)
- G. Woo. Kernel Estimation Methods for Seismic Hazard Area Source Modeling. *Bulletin of the Seismological Society of America*, 86(2): 353–362, 1996. (Cité p. 27)
- C. B. Worden, D. J. Wald, T. I. Allen, K. Lin, D. Garcia, et G. Cua. A Revised Ground-Motion and Intensity Interpolation Scheme for ShakeMap. *Bulletin of the Seismological Society of America*, 100(6): 3083–3096, 2010. ISSN 0037-1106. doi: 10.1785/0120100101. (Cité p. 70)
- O. Zakharova, S. Hainzl, et C. Bach. Seismic moment ratio of aftershocks with respect to main shocks: SEISMIC MOMENT RATIO OF AFTERSHOCKS. *Journal of Geophysical Research: Solid Earth*, 118(11): 5856–5864, 2013. ISSN 21699313. doi: 10.1002/2013JB010191. (Cité p. 55, 56, 58)
- I. Zentner. Bayesian estimation of the maximum magnitude M_{\max} based on the statistics of extremes. Technical Report SIGMA-2-2018-D5-004, SIGMA, 2018. (Cité p. 47, 52, 139)
- J. X. Zhao, J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H. K. Thio, P. G. Somerville, Y. Fukushima, et Y. Fukushima. Attenuation Relations of Strong Ground Motion in Japan Using Site Classification Based on Predominant Period. *Bulletin of the Seismological Society of America*, 96(3): 898–913, 2006. ISSN 0037-1106. doi: 10.1785/0120050122. (Cité p. 37)
- J. Zhuang, Y. Ogata, et D. Vere-Jones. Stochastic Declustering of Space-Time Earthquake Occurrences. *Journal of the American Statistical Association*, 97(458): 369–380, 2002. ISSN 0162-1459, 1537-274X. doi: 10.1198/016214502760046925. (Cité p. 17, 32)

Générateur stochastique de séismes en contexte de sismicité faible à modérée : des données à l'aléa.

Cas de la France métropolitaine.

Résumé : La France métropolitaine est éloignée des principales limites de plaques tectoniques. De ce fait, l'origine de la sismicité est plus difficile à comprendre que dans les régions actives localisées le long des limites de plaques. De plus, les données sismiques disponibles (séismes, accumulation des contraintes dans le sol, failles, mécanismes au foyer, etc.) y sont moins nombreuses. Ces observations sont deux des principaux freins rencontrés lorsque l'on souhaite caractériser l'aléa sismique sur un territoire à sismicité faible à modérée.

L'approche proposée dans cette thèse s'appuie sur un générateur stochastique produisant des séismes synthétiques en combinant des observations de terrain et des connaissances théoriques sur la sismicité du territoire d'étude. Ce générateur de séismes se divise en trois parties : (i) le tirage temporel des séismes principaux, suivi du (ii) tirage de leur localisation conditionné par la magnitude et enfin de (iii) la production de répliques associées aux séismes principaux.

Le tirage temporel est réalisé à l'aide d'une distribution fréquence-magnitude (DFM) calculée sur l'intégralité du territoire d'étude via la méthode non paramétrique des temps de latence. Ce calcul à l'échelle du territoire entier permet d'augmenter le nombre de données, de diminuer les périodes de retour et donc d'estimer les fréquences de chaque

magnitude directement à partir des données. Une attention particulière a été portée à la méthode des temps de latence afin de pallier la faible quantité de données observées notamment pour les fortes magnitudes. Le tirage spatial des séismes principaux nécessite, quant à lui, une régionalisation et une densité spatiale caractérisant la sismicité. La régionalisation permet de limiter spatialement la magnitude maximale autorisée des séismes synthétiques. La localisation d'un séisme synthétique d'une magnitude donnée est tirée dans la densité spatiale uniquement au sein des régions autorisant cette magnitude. Les répliques sont ensuite générées autour des séismes principaux à partir de la loi de Bâth et de la distribution proportion-magnitude des répliques.

L'aléa associé à chaque séisme (principal et réplique) généré est calculé à l'aide de modèles de prédiction des mouvements du sol pondérés. La pondération est dépendante de la magnitude du séisme et de la distance séisme-site. Cette pondération s'effectue en utilisant la base de données européenne d'accélération du sol RESORCE. Enfin, l'observation en un site donné des aléas produits par les séismes synthétiques sur un million d'années permet d'estimer les probabilités annuelles de dépassement de seuils d'aléa donnés.

Mots-clés : sismicité faible à modérée, générateur de séismes, stochastique, aléa, France

Stochastic earthquake generator for low-to-moderate seismicity areas: from data to hazard. Application to French mainland.

Abstract: French mainland seismicity is considered as low to moderate due to its remoteness from tectonic plate boundaries. A first consequence is that the origins of its seismicity are harder to understand than in active regions close to tectonic plate boundaries. Another consequence is the lack of available data (earthquakes recorded but also strain rates, active faults...). These two observations make difficult to estimate seismic hazard in low-to-moderate seismic areas.

The proposed approach is to generate synthetic earthquakes by combining observation and theoretical knowledge on the seismicity of the studied territory. This generator is based on a three-step scheme: (i) the temporal draw of main shocks, (ii) their spatial draw conditioned by magnitude and finally (iii) the generation of aftershocks they produce.

The temporal step needs a recurrence rate. Past seismicity of the whole studied area is analysed thanks to the non-parametric inter event time method in order to obtain this wished recurrence rate. Computing the recurrence rate at the whole territory scale allows to keep the maximum quantity of data, reduce the return periods and so estimate main shock frequencies directly from observed data for

each magnitude. An implementation has been developed to overcome the accuracy fall of the inter event time method observed when data are sparse. The spatial step needs a regionalization and a spatial density representing seismicity. The regionalization allows maximum magnitude limitation in space: each region is characterized by an allowed maximal magnitude. Location of a synthetic earthquake with a given magnitude is drawn in the spatial density only within regions that allow this magnitude. Aftershocks are generated around main shocks thanks to the Bâth law and the proportion-magnitude distribution of aftershocks.

The seismic hazard produced by each of the generated earthquakes (main shocks and aftershocks) is computed thanks to a set of weighted Ground-Motion Prediction Equations. The weights are obtained as function of magnitude and distance thanks to The European ground-motion database RESORCE. Finally, from direct observation of the seismic hazard produced by synthetic earthquakes over one million years, annual probabilities of exceedance can be calculated with ease.

Keywords: low-to-moderate seismicity, generator of earthquakes, stochastic, hazard, France