

DISS. ETH NO. 22021

NEXT-GENERATION
PROBABILISTIC SEISMICITY FORECASTING

A dissertation submitted to
ETH ZÜRICH

for the degree of
Dr. sc. ETH Zurich

presented by
STEFAN HIEMER

Diplom-Geophysiker, Universität Potsdam, Germany

Born on April 19th 1984
Citizen of Germany

accepted on the recommendation of

Prof. Dr. Stefan Wiemer	ETH Zürich
Dr. Jochen Woessner	ETH Zürich
Prof. Dr. Domenico Giardini	ETH Zürich
Dr. Edward H. Field	USGS

Abstract

The development of probabilistic seismicity forecasts is one of the most important tasks of seismologists at present time. Such forecasts form the basis of probabilistic seismic hazard assessment, a widely used approach to generate ground motion exceedance maps. These hazard maps guide the development of building codes, and in the absence of the ability to deterministically predict earthquakes, good building and infrastructure planning is key to prevent catastrophes.

Probabilistic seismicity forecasts are models that specify the occurrence rate of earthquakes as a function of space, time and magnitude. The models presented in this thesis are time-invariant mainshock occurrence models. Accordingly, the reliable estimation of the spatial and size distribution of seismicity are of crucial importance when constructing such probabilistic forecasts. Thereby we focus on data-driven approaches to infer these distributions, circumventing the need for arbitrarily chosen external parameters and subjective expert decisions.

Kernel estimation has been shown to appropriately transform discrete earthquake locations into spatially continuous probability distributions. However, we show that neglecting the information from fault networks constitutes a considerable shortcoming and thus limits the skill of these current seismicity models. We present a novel earthquake rate forecast that applies the kernel-smoothing method to both past earthquake locations and slip rates on mapped crustal faults applied to Californian and European data. Our model is independent from biases caused by commonly used non-objective seismic zonations, which impose artificial borders of activity that are not expected in nature.

Studying the spatial variability of the seismicity size distribution is of great importance. The b -value of the well-established empirical Gutenberg-Richter model forecasts the rates of hazard-relevant large earthquakes based on the observed rates of abundant small events. We propose a novel automated method to investigate the significance of spatial b -value variations. The method incorporates an objective data-driven partitioning scheme, which is based on penalized likelihood estimates. These well-defined criteria avoid the difficult choice of commonly applied spatial mapping parameters, such as grid spacing or size of mapping radii. We construct a seismicity forecast that includes spatial b -value variations and demonstrate our model's skill and reliability when applied to data from California.

All proposed probabilistic seismicity forecasts were subjected to evaluation methods using state of the art algorithms provided by the "Collaboratory for the Study of Earthquake Predictability" infrastructure. First, we evaluated the statistical agreement between the forecasted and observed rates of target events in terms of number, space

and magnitude. Secondly, we assessed the performance of one forecast relative to another. We find that the forecasts presented in this thesis are reliable and show significant skills with respect to established classical forecasts. These next-generation probabilistic seismicity forecasts can thus provide hazard information that are potentially useful in reducing earthquake losses and enhancing community preparedness and resilience.

Résumé

Le développement de méthodes probabilistes de prévision sismique constitue aujourd'hui une des tâches les plus importantes de la recherche en géophysique. De telles méthodes forment la base de l'évaluation probabiliste du risque sismique, qui est une approche utilisée couramment pour engendrer des cartes de dépassement de mouvement sismique au sol. Ces cartes de risques servent de guides pour la réglementation de construction. De plus, en l'absence de capacité de prédiction de séismes de manière déterministe, une solide construction de bâtiments ainsi qu'une robuste planification des infrastructures est essentielle pour éviter des catastrophes.

Les méthodes de prévision probabiliste de la sismicité sont des modèles qui renvoient la probabilité de survenue de séismes en fonction de l'espace, du temps et de leur magnitude. Les modèles présentés dans cette thèse représentent la sismicité principale indépendamment du temps. L'estimation fiable de la distribution spatiale et des magnitudes est par conséquent d'une importance cruciale pour l'élaboration de ces prévisions probabilistes. C'est pourquoi l'accent est mis sur des approches construites à partir de données pour déduire ces distributions, en contournant l'utilisation de paramètres arbitraires et de décisions d'experts subjectives.

L'estimation par noyau est une méthode appropriée pour transformer les lieux ponctuels de séismes en une distribution spatiale continue de probabilités. Cependant, le fait de négliger des informations provenant de réseaux de failles constitue un déficit considérable et limite l'efficacité de ces modèles de sismicité actuels. Au cours de cette thèse, je présente une nouvelle méthode de prévision sismique probabiliste, qui applique la méthode de l'estimation par noyau non seulement à la sismicité passée, mais aussi aux taux de glissement des cartographies de failles de la croûte enregistrés en Californie et en Europe. Ce modèle a l'avantage d'être indépendant de préjugés causés par l'utilisation courante de zonages sismiques subjectifs, qui imposent des frontières d'activité sismique artificielles car non attendues dans la nature.

L'étude de la variabilité spatiale de la distribution des magnitudes est d'une grande importance. Le paramètre b , issu du célèbre modèle empirique de Gutenberg-Richter, prévoit le taux de grands séismes pertinents sur la base des taux observés sur de nombreux petits séismes. Une nouvelle méthode automatisée d'exploitation de la signification statistique des variations spatiales de b est donc présentée dans cette thèse. La méthode comprend un schéma de partitionnement de données objectives, qui est basé sur des estimations de probabilité pénalisées. Ces critères bien définis permettent d'éviter de faire le choix difficile de paramètres cartographiques spatiaux courants, tels que l'espacement de grille ou la taille de cercles cartographiques. Ainsi, une méthode de prévision probabiliste

de sismicité a été construite et inclut les variations spatiales de b . La démonstration de la compétence et la fiabilité de ce modèle est illustrée par une application à des données en provenance de la Californie.

Tous les modèles de prévision sismique probabiliste proposés ont été soumis aux méthodes d'évaluation utilisant des algorithmes avancés fournis par le «Collaboratory for the Study of Earthquake Predictability». Premièrement, j'ai vérifié l'accord statistique entre les taux sismiques prévus et observés en termes de nombre total, d'espace et de magnitude. Deuxièmement, j'ai également évalué la performance d'une prévision par rapport à l'autre. D'après ces résultats, les prévisions présentées dans cette thèse sont fiables et montrent des améliorations considérables par rapport aux prévisions classiques établies précédemment. La nouvelle génération de prévisions sismiques probabilistes peut ainsi fournir des renseignements sur le risque sismique. De telles informations sont potentiellement utiles dans la réduction des dégâts dus aux séismes et dans l'amélioration de la préparation et de la résilience des populations.