ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY
ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
DEPARTMENT OF STATISTICS

Fault-specific insurance pricing, reserving and CAT bond design for seismic risk assessment - The case of Greece

By
Emmanouil Louloudis

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of PhD in Statistics

Athens, Greece
May, 2023

The implementation of the doctoral thesis was co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the Act “Enhancing Human Resources Research Potential by undertaking a Doctoral Research” Sub-action 2: IKY Scholarship Programme for PhD candidates in the Greek Universities.

Ευρωπαϊκή Ένωση
European Social Fund

Operational Programme
Human Resources Development, Education and Lifelong Learning
Co-financed by Greece and the European Union

ΕΣΠΑ
2014-2020
ανάπτυξη - εργασία - αλληλεγγύη

Τιμολόγηση ασφαλίστρων, αποθεματοποίηση και σχεδιασμός ομολόγου καταστροφής για τον κίνδυνο του σεισμού με χρήση ενεργών ρηγμάτων - Η περίπτωση της Ελλάδας

Εμμανουήλ Λουλούδης

ΔΙΑΤΡΙΒΗ
Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Διδακτορικού Διπλώματος στη Στατιστική

Αθήνα
Μάιος, 2023

Η υλοποίηση της διδακτορικής διατριβής συγχρηματοδοτήθηκε από την Ελλάδα και την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) μέσω του Επιχειρησιακού Προγράμματος «Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Δια Βίου Μάθηση», 2014-2020, στο πλαίσιο της Πράξης «Ενίσχυση του ανθρώπινου δυναμικού μέσω της υλοποίησης διδακτορικής έρευνας Υποδράση 2: Πρόγραμμα χορήγησης υποτροφιών ΙΚΥ σε υποψηφίους διδάκτορες των ΑΕΙ της Ελλάδας.

Copyright

Supervisor
Prof. Alexandros Zimbidis

Thesis committee
Prof. Athanasios Yannacopoulos
Prof. Andrianos Tsekrekos

External examination committee
Prof. Stelios Psarakis
Prof. Epaminondas Kyriakidis
Prof. Alberto Adrego Pinto
Prof. Diogo Pinheiro

DEDICATION

This thesis is dedicated to my grandmother, Konstantina.

ACKNOWLEDGEMENTS

I would like to thank the following persons, without whom I would not have been able to complete this journey. First of all, I would like to deeply thank my supervisor Prof. Alexandros Zimbidis. His guidance and overall insights in this field have made this research an inspiring experience for me. I feel honor for being one of his PhD students and I am very grateful that he totally stood by my side all these years supporting me. I would also like to thank Prof. Athanasios Yannacopoulos for his advice and support during these years and for inspiring me to start this PhD journey after supervising my MSc thesis in the field of risk management. I would also like to express my special thanks to Prof. Andrianos Tsekrekos for providing me all his advice, comments and corrections. I also thank all the members of my PhD Committee for spending time on my research work providing their useful comments. Many thanks to Dr. Georgios Deligiannakis for helping me collecting the data for faults and responding to my geological questions. I special thank Hellas Direct for providing me a research award for the earthquake pricing model and especially its co-founders Mr. Emilios Markou and Mr. Alexis Pantazis, but also Mr. Yiannis Parizas who was at that time the head of pricing. As well, I would like to express my gratitude to Prof. Ioannis Ntzoufras for providing me with valuable advice during my PhD studies. Last but not least, I thank from the deepest of my heart, my family for their extreme patience and support during all my studies.

ABSTRACT

The main objective of this dissertation is the stochastic modelling and quantification of the earthquake risk in the insurance context. It serves as a basis for making the most important actuarial decisions of the engaged institutions, including the premium rating, the assessment of the Solvency Capital Requirement (SCR) and the design of the respective Catastrophe (CAT) bond. Generally, the earthquake models developed in the insurance industry (Poisson processes at most) are based on historical catalogues, which provide information relevant to a few hundred years. Contrary, the simulation mechanism in this thesis considers the information derived by the geometry of faults covering a period up to 15 thousand years in the past in order to get reliable and robust actuarial estimates. Moreover, Voronoi polygons or the ETAS model are used for the modelling of historical catalogues instead of typical Poisson processes and continuous fragility curves instead of discrete and more uncertain damage probability matrices. While similar models in the industry are defined in terms of region pricing, the coordinate precision achieved in this work is useful for an insurance company to have complete information about the composition of its portfolio of buildings and avoid or handle adverse selection. Furthermore, the shortage of equity capital in insurance and reinsurance companies makes them unable to compensate the large-scale claims caused by extreme catastrophic events. Therefore, the industry employs catastrophe (CAT) bonds, that transfer this risk to investors in capital markets. In the present thesis, CAT bond designing and pricing is processed with respect to the proposed fault-specific earthquake model involving various statistical and machine learning discounting methods and credit risk. CAT bonds could be issued in order to mitigate the earthquake effects.

ΠΕΡΙΛΗΨΗ

Κύριος σκοπός της διατριβής είναι η στοχαστική μοντελοποίηση και ποσοτικοποίηση του σεισμικού κινδύνου στο πλαίσιο της ασφάλισης του συγκεκριμένου κινδύνου. Θέτει τη βάση για τις πιο ορθές αναλογιστικές αποφάσεις των εμπλεκόμενων μερών, όπως η τιμολόγηση ασφαλίστρων, ο υπολογισμός του απαιτούμενου κεφαλαίου φερεγγυότητας και ο σχεδιασμός του αντίστοιχου ομολόγου καταστροφής. Τα σεισμικά μοντέλα που χρησιμοποιούνται ευρέως στην ασφαλιστική αγορά (κατά κύριο λόγο διαδικασίες Poisson) είναι βασισμένα σε ιστορικούς καταλόγους, οι οποίοι παρέχουν σχετική πληροφορία κάποιον εκατοντάδων ετών. Αντίθετα, ο μηχανισμός προσομοίωσης που παρουσιάζεται στη διατριβή βασίζεται στη γεωμετρία των ρηγμάτων, η οποία καλύπτει πληροφορία έως και 15 χιλιάδων ετών στο παρελθόν ώστε να εξαχθούν αξιόπιστα αναλογιστικά μεγέθη. Επιπρόσθετα, πολύγωνα Voronoi ή το επιδημικό μοντέλο ETAS χρησιμοποιούνται για την μοντελοποίηση των ιστορικών καταλόγων αντί τυπικών διαδικασιών Poisson και συνεχείς καμπύλες τρωτότητας αντί διακριτών και πιο αβέβαιων πινάκων πιθανοτήτων ζημίας. Καθώς παρόμοια μοντέλα της αγοράς είναι κατασκευασμένα ώστε να παράγουν τιμολόγηση ανά περιοχές, στην εργασία αυτή έχει επιτευχθεί ακρίβεια ανά συντεταγμένη χρήσιμη για μια ασφαλιστική εταιρεία για πλήρη γνώση του χαρτοφυλακίου κτιρίων της ώστε να μπορεί η ασφαλιστική εταιρεία να αποφύγει ή να διαχειριστεί καταστάσεις αντιεπιλογής. Επιπρόσθετα, οι μεγάλης κλίμακας καταστροφικές αποζημιώσεις του σεισμού καθιστούν τις αντασφαλιστικές εταιρείες ανίκανες να διαχειριστούν μόνες τους τα έξοδα αυτά. Για τον λόγο αυτό, η ασφαλιστική αγορά δημιούργησε τα ομόλογα καταστροφής ώστε να μεταφέρεται ο εν λόγω κίνδυνος στους επενδυτές της αγοράς κεφαλαίων. Στην παρούσα διατριβή, διεξάγεται ο σχεδιασμός και η τιμολόγηση του σχετικού ομολόγου καταστροφής συναρτήσει του προτεινόμενου σεισμικού μοντέλου χρησιμοποιώντας είτε αμιγώς στατιστικές προεξοφλητικές μεθόδους είτε σε συνδυασμό με μηχανικής μάθησης λαμβάνοντας υπόψη και τον πιστωτικό κίνδυνο κάθε εκδότη. Τα ομόλογα αυτά μπορούν να εκδοθούν για την αποτελεσματική

αντιμετώπιση των οικονομικών συνεπειών ισχυρών σεισμών.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this introductory chapter, we discuss some preliminary notation with respect to the Solvency II Directive regulating the reserving process of insurance institutions and the main type of deformations causing earthquakes, the so-called faults. First, we discuss the necessity of a proper estimation of the technical provisions and capital requirements according to Solvency II. Then, the reason of using simulations and faults in a seismic risk project (insurance or CAT bond pricing) is discussed as means of dealing with the lack of sufficient historical data (Deligiannakis et al., 2018; Deligiannakis et al., 2021; Louloudis et al. 2022) and their economic impacts. Finally, we present the structure of the present thesis referring briefly to the innovative methods used and the results induced.

1.1 Solvency II Directive

The European Directive Solvency II for insurance companies was introduced in 2009 and was fully adopted in 2016 to replace the former directive Solvency I which was active by 2002. Solvency II is essentially the insurance analogue of the Basel II Directive for international financial regulations. Basel II was published in 2004, following a series of amendments, concerning the capital adequacy of banking institutions and the establishment of a stable framework for risk management in the banking sector. Compared to its predecessor, Solvency II is structured in such a way so as to create lower surpluses while maintaining large-sized reserves for capital needs as follows in Figure 1.1 below. Generally, Solvency I (at least regarding life insurance) and Solvency II stipulate that, within the framework of the internal market, consumers will be provided with a more diverse range of insurance products and will benefit from the increased competition amongst insurance undertakings. Both Solvency I (with respect to both life and non-life insurance) and Solvency II specify that it is in the best interest of insurance consumers to have access to the full range of possible insurance services so that they can choose the insurance service which is best suited to their needs (Loguinova, 2019).

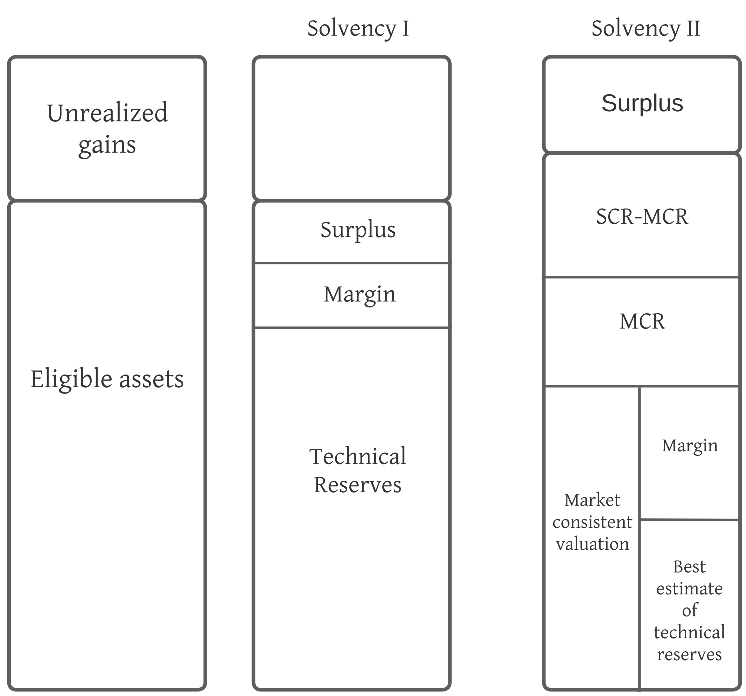

Figure 1.1: Comparison between Solvency I and Solvency II

As the book valuation may differ from the real one if the company is a bankruptcy candidate and has several liens against its assets, assets and liabilities are mainly priced according to the value of the stock market in Solvency II, i.e., the market value. The market technical reserves are currently (under Solvency II) separated into the best estimates of technical reserves as the expected present value of future cash flows and the risk margin characterizing their variation, which was not the case with Solvency I. The risk margin is designed to ensure that the value of technical provisions is sufficient for another insurer to take over and meet the insurance obligations. According to Solvency I (see Figure 1.1), the net surplus of an insurance company was part of its eligible assets allowing the firm to realize additional gains while exposing it to risk. In contrast, the amount of surplus is significantly reduced under Solvency II and it is used for reserving purposes against unexpected losses. Therefore, the standards for capital requirements do not depend primarily on the amount of premiums as before, but rather on the magnitude of risks. The previous Directive (Solvency I) was not so risk sensitive due to several facts including the following:

- Some important risks were not properly considered, such as credit risk.
- It was based on past data.
- It did not adequately recognize the risk mitigation techniques, such as reinsurance, derivatives, and securitization.
- It did not allow for a deduction from the requirement of diversification effect, between the lines of business of an insurer and legal entities of a group, and therefore, it did not encourage insurers to had properly managed their risks and their capital.
- It did not take into consideration qualitative requirements, such as quality of governance and the risk management framework; thus, did not incentivise an insurer to pursue best practices of risk management.

The assets of the insurer's balance sheet as described in Solvency II (Figure 1.2) consist of investments (bonds, shares, property etc.) and receivables from customers and reinsurers (Starita & Malafronte, 2014). The liabilities on the other hand consist of the technical provisions (best estimates plus the risk margin) for premiums and claims and the own funds, i.e., the assets subtracting the debt. The minimum capital requirement (MCR) and the solvency capital requirement (SCR) are calculated as portions of the unexpected liabilities (own funds).

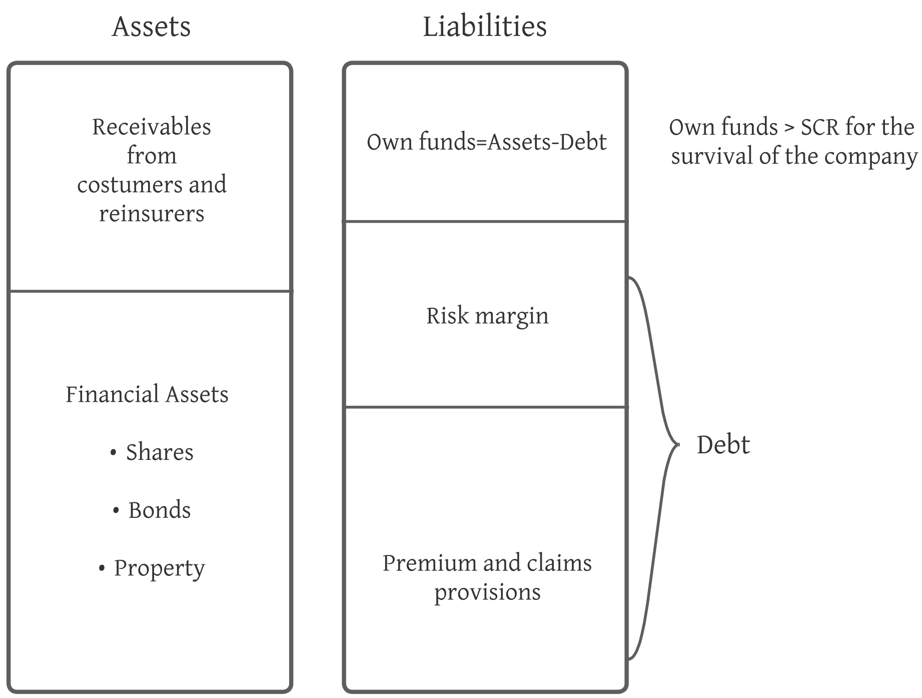

Figure 1.2: Description of the Assets and Liabilities according to Solvency II

The SCR (unexpected losses covered by own funds) is calculated as the 99.5% quantile of the distribution of losses (Figure 1.3) subtracting the expected losses (technical provisions), i.e., an insurer is expected to exceed its own funds once in

200 years. Similarly, the MCR is calculated as the 85% quantile of the distribution of losses subtracting the expected losses. As long as the SCR is sufficient to meet all the needs of policyholders and beneficiaries of insurance obligations, the MCR can be regarded as a sublevel of the SCR, since its breach is likely to lead to supervisory intervention.

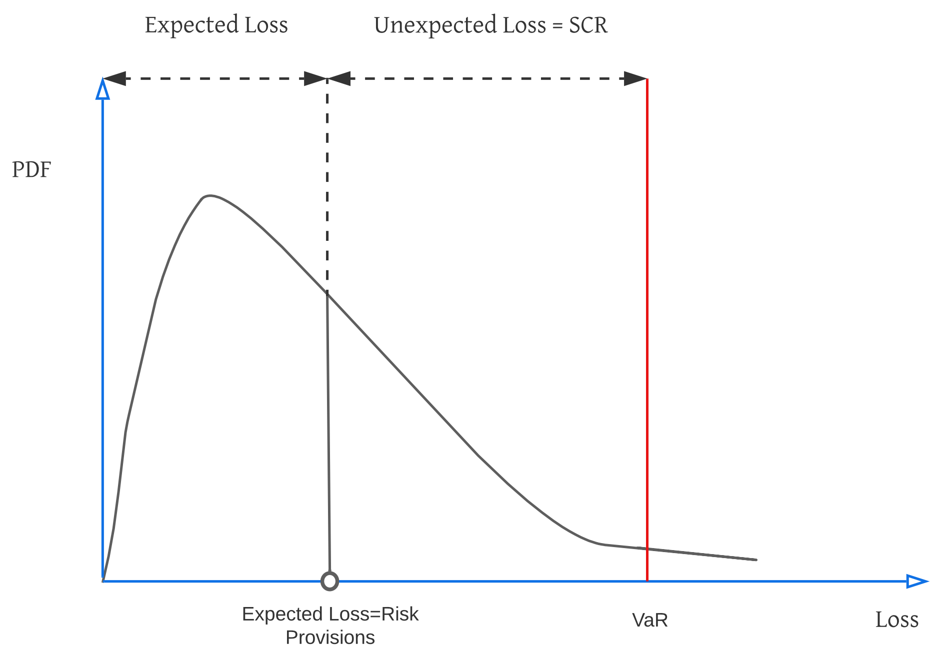

Figure 1.3: Distribution of the loss random variable and its relationship with reserving

According to the standard formula (SF), the calculated SCR covers the following quantifiable risks: the underwriting risk for life, non-life, and health business, the credit risk, the market risk, the risk linked to intangible assets, and the operational risk. According to a modular approach, each risk is divided into a set of sub risks, and the process of aggregating the SCR for each risk is based on the correlation between each risk component and the others. An interesting work by Cifuentes (2016) investigates how operational risk is correlated to the other 5 main risks of pillar 1. A more detailed identification of risks is presented in Figure 1.4.

The overall capital requirement SCR of an insurance company according to the SF of Solvency II as defined in Pillar I follows:

$$SCR = BSCR + SCR_{op},$$

where SCR_{op} is the SCR due to operational risk and

$$BSCR = \sqrt{\sum_{i,j} \mathbf{Cor}_{i,j} SCR_i SCR_j} + SCR_{int},$$

where SCR_{int} is the SCR due to intangibles.

The Cor matrix is described below in Table 1.1.

Table 1.1: Correlation matrix for BSCR estimation

Cor	**Market**	**Default**	**Life**	**Health**	**Non-Life**
Market	1	0.25	0.25	0.25	0.25
Default	0.25	1	0.25	0.25	0.25
Life	0.25	0.25	1	0.25	0
Health	0.25	0.25	0.25	1	0
Non-Life	0.25	0.5	0	0	1

The objectives of the Solvency II project represent an articulated structure based on objectives introduced below. The objectives of the first order are general: improving the protection of policyholders and beneficiaries through the integration of the European Insurance Market, enhancing the international competitiveness of European insurers and reinsurers, and promoting better regulation as pursued by the European Union.

The objectives of the second order are specific: from the insurer's perspective, making the risk management of insurers and reinsurers in the EU more effective,

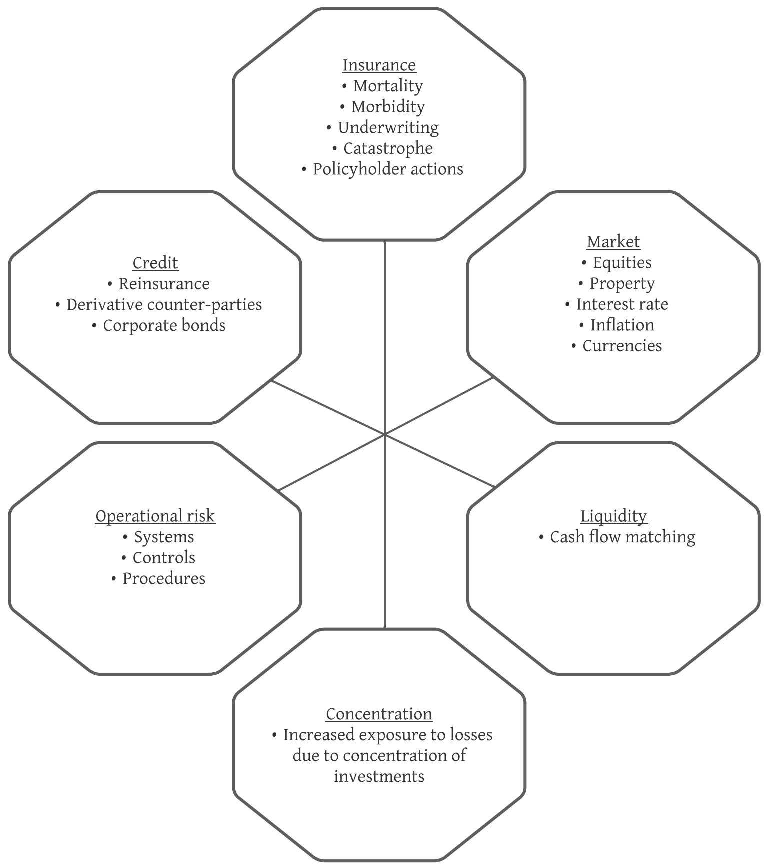

Figure 1.4: Map of identified risks according to Pillar I of Solvency II

and giving them better tools to allocate capital to risk; from the supervisor's perspective, advancing supervisory convergence and cooperation as well as promoting international convergence and encouraging cross-sectorial consistency. All of these efforts lead to an increase in transparency. The third order seeks to achieve several operative objectives: from the insurer's point of view, harmonizing the most important component of solvency, that is, the calculation of technical provisions, also to introduce risk-sensitive and proportionate capital requirements for small insurers; from the supervisor's point of view, to ensure efficient supervision of insurance groups and financial conglomerates by harmonizing supervisory powers, methods and tools as well as supervisory reporting.

Generally speaking, the three pillars of Solvency II are organized as follows (Heep-Altiner et al. 2018). The first pillar is composed of quantitative requirements, which include the solvency capital requirement (SCR), the minimum capital requirement (MCR), and the rules to calculate the technical provisions and the portfolio of assets that cover technical provisions. Insurance companies can use internal models, completely or partially to calculate the SCR, which models must be approved by the supervisory authorities. However, not all type of risks can be adequately assessed through solely quantitative measures. The second pillar of Solvency II aims at identifying businesses that have economic, organizational or other characteristics that would be likely to lead to a higher risk profile. Therefore, the second pillar represents qualitative requirements related to the oversight of the risk profile and the capital necessary to satisfy solvency needs by the regulator and by the insurer itself through its risk management policy and its own risk solvency assessment. The second pillar also includes the asset liabilities management (ALM) from a solvency point of view. Finally, the third pillar is based on market discipline requirements to communicate the quantitative and qualitative requirements through supervisory reporting and public disclosure (transparency). The insurance part of this thesis is concentrated on the first pillar since quantitative requirements are investigated.

1.2 Natural hazards in Greece and capital requirements

The Hellenic Association of Insurance Companies has provided a report composed in 2018 including all natural hazard maps of Greece[1] i.e., earthquake, flood, hail,

[1] http://www.eaee.gr/cms/sites/default/files/cat-hazard_ maps.pdf

wind/storm, landslide, wildfire, lighting and tsunami. There are certain predefined regions where risk associated with natural hazards is estimated and plotted on geographical maps. These regions are called CRESTA zones. A total of 137 countries are divided in CRESTA zones and there are two resolutions available: High Resolution (HR) and Low Resolution (LR). These zones are derived from postal and administrative boundary data rather than being peril dependent. HR is used for risk modelling and data exchange, while LR is used for risk analyses and reporting. HR CRESTA zones in Greece consist of the 1188 different 5-digit zip codes and LR CRESTA zones consist of the 70 different 2-digit zip codes. Different organizations produce different hazard maps with respect to their subjective models. For example, the seismic hazard map provided by the Nathan Tool of Munich Re is the following (Figure 1.5).

The legend of this map refers to the maximum intensity with an exceedance probability of 10% in 50 years. There are of course similar maps for all perils, but we only focus on the seismic hazard in this thesis.

In the present thesis, we focus on the calculation of a submodule of $SCR_{non-life}$, the SCR due to earthquakes, namely SCR_{EQ}. According to the SF published in the Official Journal of the European Union[2], SCR_{EQ} should be estimated as follows[3].

$$SCR_{EQ} = \sqrt{\sum_{r,s} \mathbf{CorEQ_{r,s}} SCR_{EQ,r} SCR_{EQ,s} + SCR^2_{EQ,other}},$$

where the sum includes all possible combinations (r,s) of the regions set out in Annex VI of Solvency II Directive that are subsets of the portfolio (e.g., portfolio of countries) under investigation and $SCR_{EQ,r}$ and $SCR_{EQ,s}$ denote the capital

[2]DIRECTIVE 2009/138/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 November 2009

[3]https://www.eiopa.europa.eu/rulebook/solvency-ii/article-6931_ en?fbclid=IwAR0EFkuPpragkp3QA_ ebxJUfzlL_ vd7GT7sEQUV7XT1LDbv3HKe9X6xY6Gk

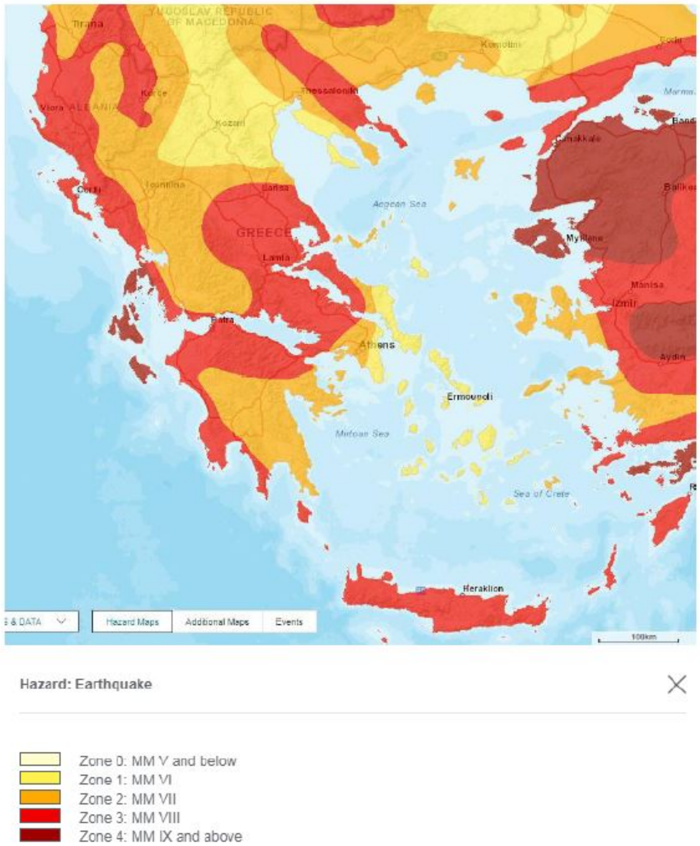

Figure 1.5: Seismic hazard map provided by the Nathan Tool of Munich Re

requirements for earthquake risk in region r and s respectively. $\mathbf{CorEQ_{r,s}}$ denotes the correlation coefficient for earthquake risk for region r and region s as set out in Annex VI. Finally, $SCR_{EQ,other}$ denotes the capital requirement for earthquake risk in regions other than those set out in Annex XIII.

For all regions (countries) set out in Annex VI, the capital requirement for earthquake risk in a particular region r i.e., $SCR_{EQ,r}$ shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from an instantaneous loss of an amount that, without deduction of the amounts recoverable from reinsurance contracts and special purpose vehicles, is equal to the following amount:

$$SCR_{EQ,r} = \sqrt{\sum_{i,j} \mathbf{CorEQ}_{r,i,j} WSI_{EQ,r,i} WSI_{EQ,r,j}}$$

where the sum includes all possible combinations of risk zones (i, j) belonging to region r as set out in Annex IX, $\mathbf{CorEQ}_{r,i,j}$ denotes the correlation coefficient for earthquake risk in risk zones i and j of region r set out in Annex XXIII, $WSI_{EQ,r,i}$ is the weighted sum insured for earthquake risk in risk zone i of region r set out in Annex IX described as:

$$WSI_{EQ,r,i} = Q_{EQ,r} W_{EQ,r,i} SI_{EQ,r,i},$$

$Q_{EQ,r}$ denotes the earthquake risk factor for region r as set out in Annex VI, $W_{EQ,r,i}$ denotes the risk weight for earthquake risk in risk zone i of region r set out in Annex X and $SI_{EQ,r,i}$ denotes the sum insured for earthquake risk in earthquake zone i of region r. It is defined as:

$$SI_{EQ,r,i} = SI(\text{property}; r, i) + SI(\text{on-shore property}; r, i).$$

SI(property;r,i) denotes the sum insured of the insurance or reinsurance undertaking for lines of business 7 and 19 as set out in Annex I in relation to contracts that cover earthquake risk and where the risk is situated in risk zone i of region r and SI(on-shore property;r,i) denotes the sum insured of the insurance or reinsurance undertaking for lines of business 6 and 18 as set out in Annex I in relation to contracts that cover onshore property damage by earthquake and where the risk is situated in risk zone i of region r.

The capital requirement for earthquake risk in regions other than those set out in Annex XIII shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from an instantaneous loss in relation to each insurance and reinsurance contract that covers one or both of the following insurance or reinsurance obligations: (a) obligations of lines of business 7 or 19 as set out in Annex I that cover earthquake risk, where the risk is not situated in one of the regions set out in Annex XIII, (b) obligations of lines of business 6 or 18 as set out in Annex I in relation to onshore property damage by earthquake, where the risk is not situated in one of the regions set out in Annex XIII.

Finally, we estimate:

$$SCR_{EQ,other} = 1.2(0.5DIV_{EQ} + 0.5)P_{EQ},$$

where DIV_{EQ} is calculated in accordance with Annex III, but based on the premiums in relation to the obligations referred to in points (a) and (b) of paragraph 5 and restricted to the regions 5 to 18 set out in Annex III and P_{EQ} is an estimate of the premiums to be earned by insurance and reinsurance undertakings for each contract that covers the obligations referred to in points (a) and (b) of paragraph 5 of the Directive during the following 12 months: for this purpose premiums shall be gross, without deduction of premiums for reinsurance contracts.

However, due to lack of transparency of the derivation of the SF rendering it a "black-box", a self-constructed purely internal model for SCR_{EQ} is proposed in this thesis.

1.3 Preliminaries of earthquake generation

Earth consists of several rigid plates that evolve in space and time. The interaction between these plates explains earthquakes (Marshak, 1997). Faults are formed in response to pushes and pulls associated with the forces that arise from the movement of tectonic plates or as a consequence of differential buoyancy between parts of the lithosphere. A fault in a broad sense is defined as a surface or zone across which there has been measurable sliding parallel to the surface. The various types of faults based on the direction of their blocks move are presented in Figure 1.6.

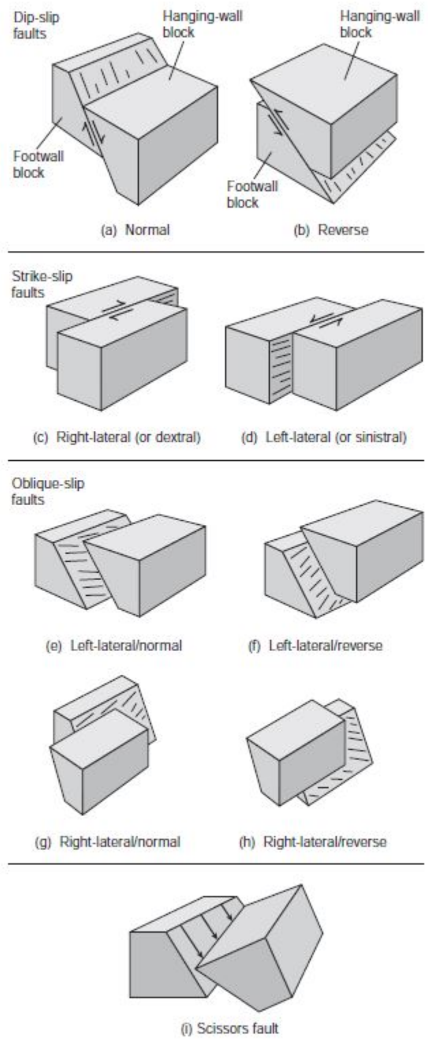

Figure 1.6: Different types of faults according to their geometry as described in Marshak (1997)

A normal fault is a dip-slip fault on which the hanging wall has slipped down relative to the footwall. When someone thinks about a typical fault, the most usual case is the normal fault. The angle that a planar geologic surface (in this study, a fault) is inclined from the horizontal is called the dip angle[4]. Moreover, the intersection between a given plane and the horizontal surface is called strike.

Description of the fault steepness

- Horizontal faults: dip about 0^o.
- Sub-horizontal faults: $0^o <$ dip $< 10^o$.
- Shallowly dipping faults: $10^o <$ dip $< 30^o$.
- Moderately dipping faults: $30^o <$ dip $< 60^o$.
- Steeply dipping fault: $60^o <$ dip $< 80^o$.
- Vertical faults: 90^o.

When fault movement occurs, one fault block slides relative to the other, which is described by the net slip. Net slip therefore is the vector between two before stuck together points and rake is the angle they created after the event. This also becomes more obvious in Figure 1.7 below.

The slip direction on a dip-slip fault is approximately parallel to the dip of the fault (i.e., has a rake between 80° and 90°). The slip direction on an oblique-slip fault has a rake that is not parallel to the strike or dip of the fault. In the field, faults with a slip direction between 10° and 80° are generally called oblique-slip. The slip direction on a strike-slip fault is approximately parallel to the fault strike (i.e., the line representing slip direction has a rake [pitch] in the fault plane of less than 10°). Strike-slip faults are generally steeply dipping to vertical.

The boundary between the slipped and unslipped region at the end of a fault is the tip line of the fault. In the field, faults are mapped at the ground surface either as a result of the fault crossing the ground during its movement or because the fault has been exposed by erosion. An underground fault that does not intersect

[4]https://earthquake.usgs.gov/learn/glossary/?term=dip

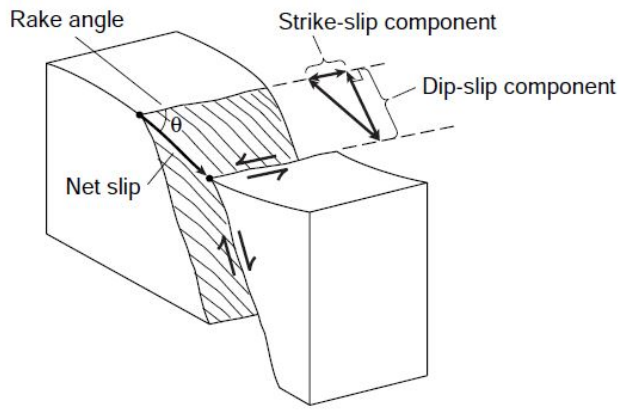

Figure 1.7: Graphical representation of geometrical definitions for a sliding normal fault

the ground surface, but dies out in the subsurface, is known as a blind fault. Blind faults have no surface ruptures by definition. However, the traces of the fault may be inferred from differential ground movement. Ground-based survey equipment and GPS (global positioning system) equipment are both capable of detecting this movement. The surface area above rupture produces a 2-D polygon, which is the main object of study in a fault-specific hazard analysis.

Most earthquakes are caused by release of elastic strain (distortion of a body in response to an applied force) accompanying sudden displacement on faults (Yeats, 2012). An earthquake's hypocenter, or the focal point where it originated, is located beneath the Earth. On the other hand, its epicenter, the point directly above the hypocenter, is located on Earth's surface. This release of strain energy results in ground shaking that is recorded on seismograms, providing information about the earthquake process and the materials through which seismic waves pass. Increasing tectonic forces continuously shear and compress the earth's crust (Towhata, 2008). Over time, the strain energy in rock is accumulated as the stress in rock increases. The crust is eventually broken by mechanical means, which releases elastic energy. This results in an earthquake. The subsequent accumulation of strain energy paves the way for an earthquake in the future. This cycle is explained by time-dependent inter-event time probability distribution functions as we describe in section 2.

In reality, the whole fault does not rupture in a single event (Towhata, 2008). The first breakage occurs at a place where the factor of safety is minimum. A further rupture occurs due to the working stress being transferred to other areas of the fault. This leads to the rupture propagating along the fault. Each rupture has a considerable impact on the ground surface. The combined effects of all the impacts produce the sensation of an earthquake with duration. The rock rupture is reported to be associated with a generation of electric current. This may be related to many precursors. Measurement of electric current in Greece is making a success to some extent in prediction of earthquakes. This method is called the VAN method by the initials of its founders (Varotsos et al. 1981a). The VAN method is more suitable for short-term earthquake prediction, while the present thesis is concentrated to financial and insurance products associated with seismic risk where the investigated hazard is at least annual but can also reach decade(s) of years. Therefore, long-term stochastic models are needed.

There are two types of elastic seismic waves that generate earthquake shaking: body and surface waves (Elnashai & Di Sarno, 2008). The feeling of shaking is generally caused by the combination of all these waves. An analytical description of body ("P-" and "S-") and surface ("Love" and "Rayleigh") waves can be found in Kumar (2008). There is no evidence that all plate boundaries are associated with strong earthquakes. San Andreas Fault in California, located along the boundary between North American and Pacific Ocean Plates, undergoes continuous deformation and hence does not rupture suddenly. In contrast, in the past there have been many large earthquakes that occurred inside tectonic plates. Therefore, there is virtually no earthquake-risk-free region in the world. The level of this risk is thus determined by factors such as local population (life) and building (non-life) density (Towhata, 2008).

As large-magnitude earthquakes are infrequent events causing lack of completeness in their historical catalogues, the geometry of fault sources can contribute to our missing knowledge. Moreover, for the same exact reason there is lack of known losses or claims from these events. Therefore, seismic projects should be based on simulations, ideally stochastic in nature.

1.4 Thesis structure

The structure of the present thesis is further organized as follows. In the second chapter, the statistical-geomechanical part of the seismic phenomenon is analyzed. More specifically, some inter-event time distributions are introduced and the spatial inference of the two different (area and fault) seismic sources is presented. The characteristics of the area sources are developed with respect to Voronoi polygons, while the characteristics of fault sources are developed by their geometrical properties. We also describe the ground motion prediction equations relating the intensity arrived on a geographical point after an earthquake of a certain moment magnitude (energy released). Finally, the very popular Epidemic Type Aftershock Sequence (ETAS) model is described. In the third chapter, the loss random variable based on fragility curves is introduced and the algorithms used for the insurance pricing and SCR evaluation are provided. The fourth chapter presents a potential structure of the respective CAT bond that could be issued as an alternative reinsurance method potentially by the Greek government to deal with catastrophic events such as that of Athens, 1999. More specifically, the trigger parameters are tuned with respect to the recurrence of faults and their consequences to structures as described by the model of chapter 2. Both loss and nature triggering parameters are considered to yield the probabilities of default. In the absence of a Special Purpose Vehicle (SPV), the CAT bond pricing is also evaluated under the existence of credit risk by issuers of different creditworthiness. Discounting methods based on short rate models and yield curves are examined concluding to different prices. Finally, the fifth chapter includes the results and further discussion and research for this topic.

Chapter 2

Earthquake statistical modelling

In this chapter, the theoretical background needed for the seismic risk assessment is provided and explained. Two main types of sources characterize earthquake events, namely area and fault sources (McGuire, 2004). An area source is a region within which future seismicity is assumed to have characteristics and locations of energy release that are constant over time and space. It is possible to define the geometry of area sources using historical seismicity alone in the simplest case. Alternatively, fault sources are faults or zones that have been identified as the origins of earthquakes by tectonic movements as already discussed in chapter 1. In contrast to area sources, fault sources exhibit a cycle behavior, where stress is accumulated and released over an expected time period. This return period, however, is characterized by a large degree of uncertainty. Unlike historical catalogs, which provide information up to a few hundred years in the past, the geometry of fault sources provides information for seismicity over a thousand years in the past. Therefore, the co-integration of area and fault sources is essential in developing a successful and realistic model.

The first purpose of this thesis is the simulation of earthquakes over space and time. The events over time are modelled by exponential inter-event times (hence events are modelled in terms of the Poisson process) for area sources and fault sources without historical activity. Fault sources associated with a historical event in the past are modeled by a time-dependent log-normal model as in Papanikolaou et al. (2013). There is a detailed description of the inter-event time distributions characterizing seismic events due to faults in Console et al. (2017). The most significant of them are outlined below in a brief summary.

1. Time-independent exponential inter-event times (Poisson Process) Let λ_s be the intensity of the Poisson process Z_t. Then, the distribution of the i.i.d inter-event times $T_1, ..., T_n \sim$ exp (λ_s) and $\lambda_s = 1/E(T)$, while $E(Z_t) = \lambda_s t$. Thus, the probability distribution function (PDF) of the inter-event times denoted by $f_T(t)$ is:

   $$f_T(t) = \lambda_s e^{-\lambda_s t}$$

   .

   The likelihood of the inter-event times given that an amount of time s has passed, is characterized by:

   $$P(T_1 > s + t | T_1 > s) = P(T > t),$$

   thus independent of time s. This memoryless property does not hold for the other distributions presented below. These are renewal processes and generate earthquakes with respect to the time-dependent conditional probabilities $P(T_1 > s + t | T_1 > s)$. Thus, the inter-event time for a next fault activation is a truncated distribution of the below stated PDFs having the knowledge that time s has passed since the latest fault decompression.

2. Time-dependent Log-Normal inter-event times

   $$f(t) = \frac{1}{\sigma_s \sqrt{2\pi x}} \exp^{-\frac{(lnt-\mu_S)^2}{2\sigma_S^2}}$$

   ,where μ_s and σ_s are the mean and standard deviation of the logarithm of the identically distributed inter-event times.

3. Time-dependent Gamma inter-event times

$$f(t) = \frac{1}{\beta_s \Gamma(\gamma_s)} \left(\frac{t}{\beta_s} \right)^{\gamma_s - 1} \exp^{-\frac{t}{\beta_s}},$$

where Γ is the Gamma function and γ_s and β_s are respectively the shape and the scale parameters of this PDF.

4. Time-dependent Weibull inter-event times

$$f(t) = \frac{\gamma_s}{\mu_s} \left(\frac{x}{\mu_s} \right)^{\gamma_s - 1} \exp^{-(\frac{x}{\mu})^{\gamma_s}}$$

5. Time-dependent Double-Exponential inter-event times

$$f(t) = \frac{1}{2b_s} \exp^{-|x - \frac{\mu}{b_s}|},$$

where b_s and μ_s are the shape and scale parameters, respectively.

6. Time-dependent Brownian Passage Time (BPT) inter-event times

$$f(t) = \left(\frac{E(T)}{w\pi C_v^2 x^3} \right)^{1/2} \exp \left(-\frac{(t - E(T))^2}{2C_v^2 E(T) t} \right),$$

where E(T) is the mean value of the inter-event time and C_v is the coefficient of variation or aperiodicity, defined as

$$C_v = \frac{\sigma(T)}{E(T)}.$$

The use of a more advanced epidemic-type branching process for modeling earthquake recurrence is also analysed in the end of this chapter.

The second purpose of this work is the generation of earthquake magnitudes processed by the use of the inverse simulation method. Regarding area sources, this simulation is dependent on a significant parameter known as the b-value, which is discussed in section 2.1.1.

2.1 Statistical analysis of area sources

2.1.1 Analysis of seismicity

The most frequent law regarding seismicity is the frequency–magnitude distribution by Gutenberg and Richter (1944). The Gutenberg–Richter (GR) law is applied to regions to describe their seismicity behaviour (rate of earthquakes arrived and proportionality of the different magnitudes generated). According to the GR law:

$$\Lambda(m_1) = 10^{a-bm_1}, \tag{2.1}$$

where $\Lambda(m_1)$ is the rate of events having magnitude greater than m_1. The rate of all earthquakes with positive magnitude is 10^a, while b is the unknown parameter of the probability density function (PDF) of their magnitude. These unknown parameter values (a referred as the α-value and b referred as the b-value) vary from region to region. According to the GR law, the rate of earthquakes of magnitudes m_0 up to m_1 is described as follows:

$$\Lambda(m_1; m_0) = 10^{a-bm_0} - 10^{a-bm_1}. \tag{2.2}$$

2.1.2 Inverse simulation method for truncated Gutenberg–Richter (GR) distribution

In this subsection, the computational method for generating values from the truncated Gutenberg–Richter distribution describing the magnitudes in area sources is presented. This theoretical tool is necessary for the simulation performed in the algorithms that follow. According to the inverse simulation method if $U \sim U(0, 1)$, then $P(F_X^{-1}(U) \leq x) = F_X(x)$, where F_X denotes the cumulative distribution function (CDF) of the random variable X. This provides a way to generate values from the truncated GR distribution by generating values from a standard uniform distribution and applying the inverse CDF. The CDF of the truncated GR is obtained by equation (2.2) that yields equation (2.3):

$$F(m) = \frac{\Lambda(m; m_{min})}{\Lambda(m_{max}; m_{min})} = \frac{1 - \exp(-B(m - m_{min}))}{1 - \exp(-B(m_{max} - m_{min}))}, \tag{2.3}$$

where $B = b \ln 10$ and $m_{min} \leq m \leq m_{max}$. The inverse function of F is then computed as follows:

$$F^{-1}(u) = \frac{ln[u \exp(-B(m_{max} - m_{min})) - u + 1] - Bm_{min}}{B} \tag{2.4}$$

2.1.3 Spatiotemporal analysis using Voronoi polygons

As it is evident, these parameters (α-value and b-value of equation (2.2)) are required to be estimated. The method used for estimating these parameters is described and analyzed in this section. We used the historical catalogue of the National Observatory of Athens from 1968 to 2019 on magnitudes greater or equal to 4 M_w. A sufficient number of events corresponding to magnitudes of $4M_w$ and higher have been recorded throughout the past 50 years. This historical catalogue contains 7,051 events with magnitude greater than 4 M_w since 1968. The annual minimum observed magnitude in the historic catalogues downloaded from the National Observatory of Athens indicates that there is a sufficient completeness above this threshold, from 1968 and after something that is very closely justified by Papazachos et al. (2000) proposing $4.5M_w$ as a completeness threshold.

First, one needs to decluster the seismic historical catalogue. Declustering is a technique performed to separate background events from foreshocks and aftershocks. Background events are considered to be independent and sufficiently modelled in terms of a Poisson process. In this work, the declustering technique by Gardner and Knopoff (1972 and 1974) is used. This declustering technique is used as it keeps in general the largest events that are useful for this risk analysis. Many declustering techniques can be found in the literature (Talbi et al. 2013) but they have not been rigorously tested (Kagan, 2013). Most of them are based on the Epidemic Type Aftershock Sequence model (ETAS), see Jalilian (2019) and Ogata (1988, 1998). The ETAS model however identifies a significant proportion of large events as aftershocks and removes them from background events leading to an underestimation of the premium rating. Nevertheless, a method in section 2.4 is also proposed that overcomes the above difficulty.

The events are ordered in decreasing magnitude to perform the declustering. Starting from the first event of the ordered catalogue, space–time windows are measured around each event in the catalogue. The size, S, and duration, T, of each window vary depending on the magnitude, M, of the potential mainshock. The largest event in each window is identified as a mainshock, while the others (foreshocks and aftershocks) are identified and removed. The space–time windows are the predicted values of the following regressions (2.5) and (2.6) performed on the data of Table 2.1 following Gardner and Knopoff (1974):

$$\ln(T) = a_1 M + b_1 \tag{2.5}$$

$$\ln(S) = a_2 M + b_2 \tag{2.6}$$

The result of the declustering is shown in Figure 2.1. The statistics of these regressions are:

$$\hat{b_1} = -0.2683, \hat{a_1} = 0.998$$

with an $R^2 = 0.9491$ and

$$\hat{b_2} = 2.2634, \hat{a_2} = 0.2851$$

with an $R^2 = 0.9997$.

The area under analysis is enclosed by the rectangle with longitude degrees that belong to the interval [19, 29] and latitude degrees that belong to the interval [34,

Table 2.1: Gardner-Knopoff windows

M	S (km)	T (days)
2.5	19.5	6
3	22.5	11.5
3.5	26	22
4	30	42
4.5	35	83
5	40	155
5.5	47	290
6	54	510
6.5	61	790
7	70	915
7.5	81	960
8	94	985

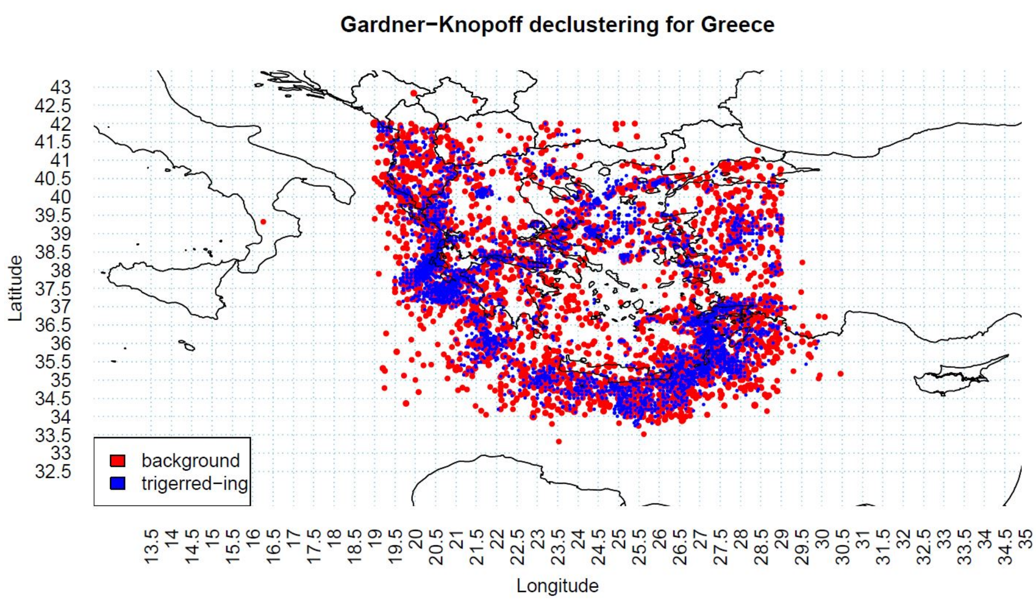

Figure 2.1: Declustering of the earthquakes in Greece

43]. The initial historical catalogue contains 7051 events with magnitude greater than 4 M_w since 1968. The number of background (after declustering) events is 2981 (37 events are outside of the specified spatial window), and 58% of the events are identified as triggering or triggered events and deleted. We obtained a similar proportion of aftershocks as in Gardner and Knopoff (1974), where they arrived at 66%. For reasons of robustness, we also checked the success of the declustering for the highly seismogenic region of Kefallinia. Applying a Box–Pierce test (1970) to the inter-event times of the declustered catalogue, the null hypothesis of independence was not rejected with a high p-value of 0.9146. The area is partitioned to different polygons, with each one governed by a unique b-value. We introduced the notion of Voronoi polygons to perform this spatial separation. First, we defined central points having a distance of 0.5 degrees ($\sim$ 50km), resulting in 360 different points covering the whole examined area. The Dirichlet tile associated with a particular data point x_i (in this application, any of the above centres) was defined as the region of space that is closer to x_i than to any other point in the pattern $\underline{x}$ (the remaining centres). Thus, the mathematical definition is as follows:

$$D(x_i|\underline{x}) = \{v \in \Re^2 : ||v - x_i||_2 = \min_j ||v - x_j||_2\}$$

These tiles form the Dirichlet tessellation or Voronoi diagram. The procedure for estimating the Voronoi diagram that best fits our data (i.e., which centres to use) follows below. The whole investigation area is partitioned into k polygons (number of centers=k). In each polygon, r, with the number of data points, N_r, the maximum likelihood estimator (MLE) of the b-value, b_r, is estimated using the following equation:

$$\hat{b_r} = \frac{N_r}{\sum_{i=1}^{N_r} m_i - N_r m_{min} \ln(10)},$$

because the Gutenberg–Richter distribution is:

$$f(m) = b \ln(10) 10^{-b(m - m_{min})}, \qquad m > m_{min}$$

and the total log-likelihood is given by

$$l(\underline{b}) = \sum_{r=1}^{k} \ln L_r(b_r|\underline{m_r}),$$

where L_r is the likelihood of the magnitude in the $r-th$ polygon and $\underline{b} = (b_1, ..., b_r, ...)$.

The ideal Dirichlet tessellation corresponds to the lowest Bayesian information criterion (BIC):

$$\text{BIC} = -2l(\underline{b}) + kln\ (n),$$

where k is the number of different centres and n is the total number of events.

As the number of different tessellations that are produced with these 360 centres is huge, the procedure for the Voronoi analysis was undertaken as outlined below. Starting from the assumption that there are only two different tiles, 20000 different random tessellations were tested. We repeated the procedure with the assumption that there are three different tiles, continuing until 20000 different random tessellations had been tested given that there were 20 different tiles.

Thus, the resulting Voronoi diagram better describes the seismicity of Greece based on the Bayesian Information Criterion (BIC). The tessellation with the lowest BIC value was chosen and is presented together with the b-values of its tiles in Figure 2.2.

The best tessellation produced a BIC equal to 219.65, while the BIC under no separation (null model) was 261.85. We assigned to each cell, j, of the grid its rate, λ_j (number of events observed over the time horizon), and its b-value (the same b-value as the polygon to which the cell belongs). This 20x18 grid confirmed that our partition preserved the average seismicity behaviour. The average b-value is 1.21 very close to 1.26, which was the estimated b-value obtained using the ZMAP software (Wiemer, 2001) for the same data.

The number of events inside each polygon were as follows: Polygon 1: 601

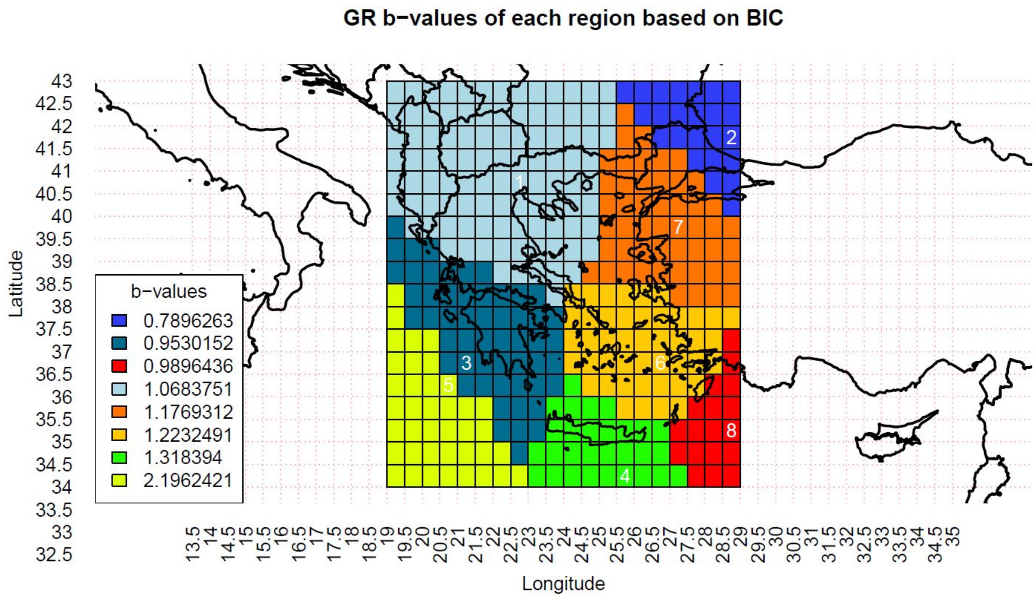

Figure 2.2: G–R b-values based on the BIC Voronoi algorithm

events; Polygon 2: 20 events; Polygon 3: 664 events; Polygon 4: 511 events; Polygon 5: 134 events; Polygon 6: 448 events; Polygon 7: 342 events; and Polygon 8: 224 events.

The annual activation rate of each cell of the (20x18) grid was estimated and is presented in Figure 2.3.

To get a smoother version of the rates, the Frankel kernel was applied. The Frankel kernel is obtained with respect to the correlation distance, c, between each area. The correlation distance is estimated according to the methodology of Valentini et al., (2017) equal to $c = 30$ km.

Having estimated c, the Frankel kernel intensity function maps the smoothed annual rate, $\overline{\lambda_j}$, of earthquakes in each cell, j, of the grid, and it is defined as follows:

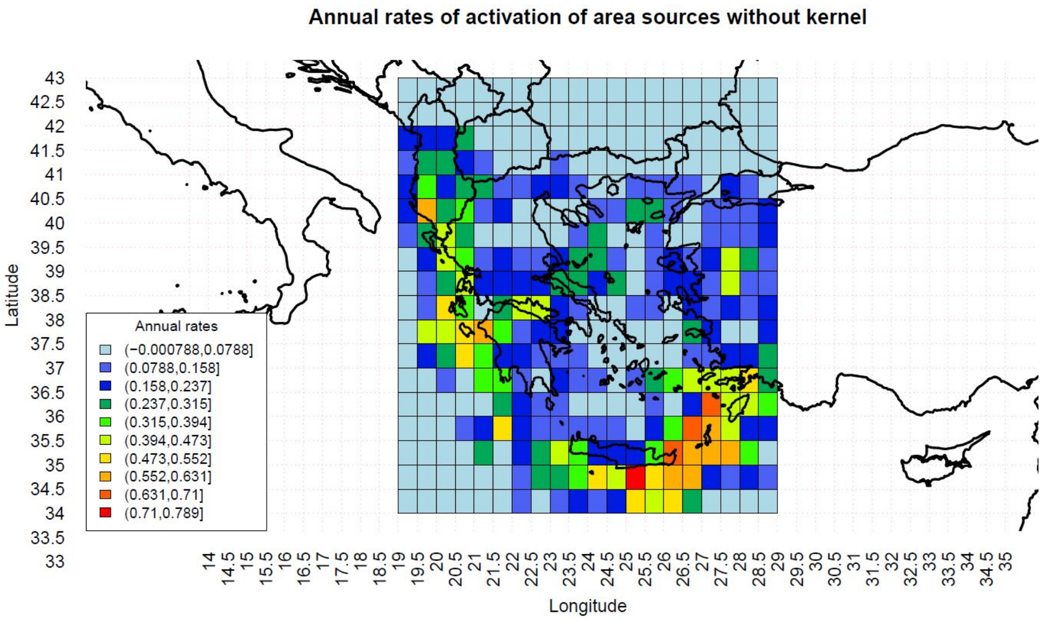

Figure 2.3: Annual rates of earthquakes without kernel

$$\overline{\lambda_j} = \frac{\sum_i \lambda_i \exp\left(\frac{-\Delta_{ji}^2}{c^2}\right)}{\sum_i \exp\left(\frac{-\Delta_{ji}^2}{c^2}\right)},$$

where Δ_{ji} is the distance between the centroids of cells j and i. The result is provided in Figure 2.4.

The smoothed rates obtained by Frankel's kernel are affected by a weighting function that depends on the distances from the center of each cell to the nearest point of the projected surface of all faults (Valentini et al., 2017). The distance from a point of interest to the nearest point of the projected surface is called the Joyner–Boore (JB) distance. This function must be applied to the area rates based on the assumption that the closer an earthquake occurs to a fault on the same range with the potential magnitudes of the fault, the more probable it is that it has occurred due to that fault. Thus, the rates obtained by the Frankel kernel are trans-

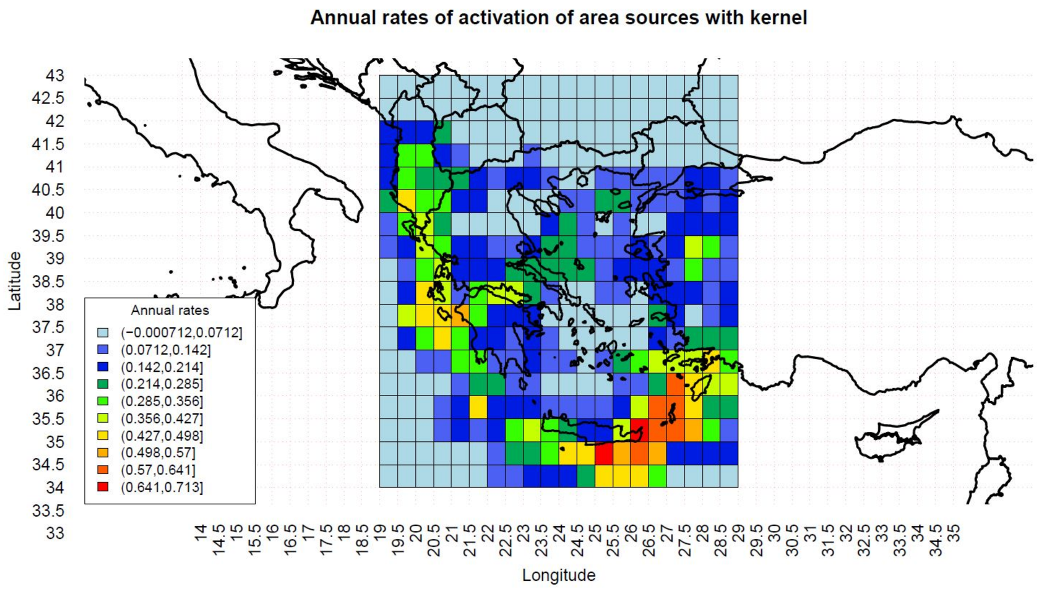

Figure 2.4: Annual rates of earthquakes with kernel

formed as outlined below.

Denoting the weighting function by h, the total rate of earthquakes (denoted by v_j) arriving in an area source j where a higher than 6 M_w earthquake has been observed in the past can be done as follows:

$$v_j(m > 4) = \overline{\lambda_j}(4 < m \leq 6) + \overline{\lambda_j}(6 < m \leq m_{up})\times h$$

$$= \overline{\lambda_j}(m > 4)P(4 < m \leq 6) + \overline{\lambda_j}(m > 4)P(m_{up} \geq m > 6)h =: k_1 + k_2,$$

where m follows the GR distribution, the lowest potential moment magnitude produced by all faults is approximately 6 M_w, m_{up} is the maximum observed magnitude in each area source, and $\overline{\lambda_j}$ is estimated by the Frankel kernel. The total rate of each area source is the summation of the rate of the weaker earthquakes, k_1, and the strongest but more improbable, k_2. The weighting function ensures that there is no double counting of events due to the simultaneous use of area and fault sources. The next section analyses the properties of fault sources.

2.2 Statistical analysis of fault sources

Historical catalogues typically cover several decades or even a few hundred years; thus, they reveal only a small amount of information compared to the thousands of years that make up the typical period of occurrence for earthquakes in a specific geographic area. The current research approach progresses by adding fault sources, which is the recent trend of geophysical research in order to address the missing information. Geological and geomorphological studies of faults may provide information for up to 15 thousand years in the past (Deligiannakis et al. 2018). On the other hand, the information contained in all available historical catalogues is very poor, with the oldest recorded events in Greece—which has one of the longest historical catalogues worldwide—reported as occurring in 1500 AD, corresponding to events between 7.3 and 8 M_w (Papazachos et al. 2000).

The fault database of GreDaSS (Pavlides et al. 2010) was used for the analysis of faults in Greece (Figure 2.5).

For the properties of faults, the main analysis is in line with Pace et al. (2016), assuming the default geological values of shear modulus $\mu = 3 \times 10^{10}$ Pa (see also Youngs and Coppersmith, 1985 for more details) and strain drop equal to $k = 3 \times 10^{-5}$.

Elementary, though not very precise, estimates of the maximum magnitude that each fault produces along with its mean reactivation time are obtained as follows. The average displacement (in meters) over all the area of each fault is equal to

$$D = kL_{str},$$

where L_{str} is the along-strike length of the fault in meters. The maximum scalar seismic moment is computed by Aki (1966) from the formula

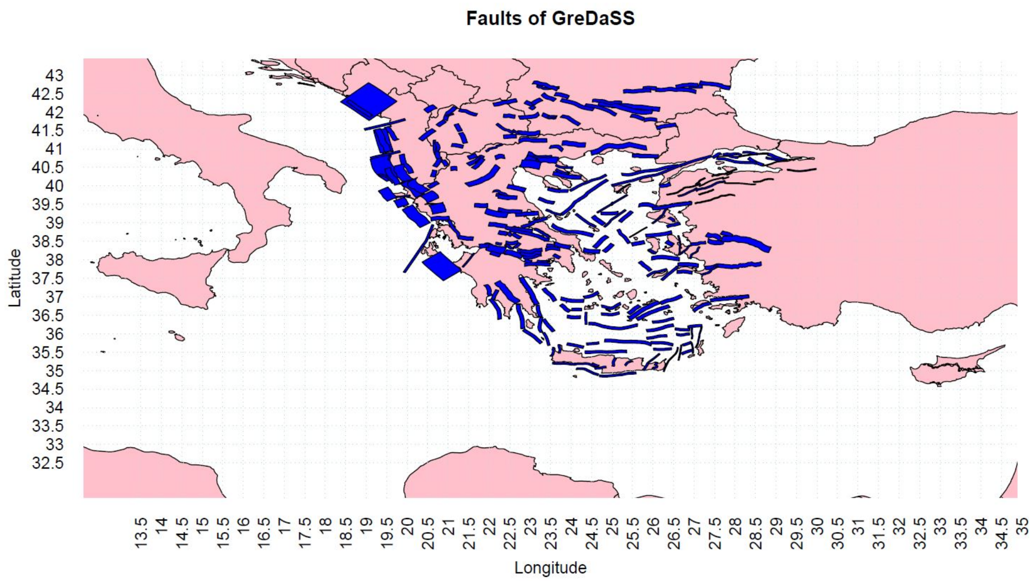

Figure 2.5: Faults of GreDaSS plotted in R

$$M_0 = \mu L_{str} DW,$$

where W is the down-dip width estimated by

$$W = \frac{\text{seismogenic thickness}}{\text{dip}}$$

.

Then, by Hanks and Kanamori (1979) the maximum expected magnitude

$$E(M_f) = \frac{2}{3} ln(M_0 - 9.1)$$

.

The average reactivation time is equal to:

$$E(T_f) = \frac{D}{\text{SR}},$$

where SR is the slip rate of the fault in meters/year. However, this elementary method was not adopted in the present thesis and the following more advanced procedure was applied.

As beforementioned, the magnitude of a fault has a mean value denoted by $E(M_f)$ and a standard deviation of $\sigma(M_f)$; it is reactivated at an average time of $E(T_f)$ with a standard deviation of $\sigma(T_f)$. These four parameters are estimated via the MB tool of the FiSH (version 1.02) Package in MATLAB (Pace et al. 2016), which uses many empirical equations relating $E(M_f)$ to certain properties of the faults, as for example the equation of Hanks and Kanamori (1979) or Wells and Coppersmith (1994) and returns overall values of $E(M_f)$, $\sigma(M_f)$, $E(T_f)$ and $\sigma(T_f)$.The mean reactivation time $E(T_f)$ is estimated by the following ratio:

$$E(T_f) = \frac{10^{9.1+1.5E(M_f)}}{\mu \mathrm{SR} L_{str} W}$$

where SR is the slip rate of the fault, L_{str} is the is the along-strike length of the fault, and W is the down-dip width of the fault. This is the ratio between the seismic moment released from a maximum event and the seismic moment rate defined by slip rate (García-Mayordomo et al. 2017). Its standard error referred to as $\sigma(T_f)$ is obtained via the law of error propagation (Peruzza et al. 2010).

The first two moments of T_f are estimated, but still there is no knowledge for the true probability distribution of the inter-event times. Thus, we introduce a time-dependent model. We assume that the inter-event time of each fault, t, that is related to a previous earthquake follows a left-truncated at the latest elapsed time lognormal distribution (Papanikolaou et al. 2013) with parameters:

$$\mu_t = -0.5(ln(\sigma(T_f)^2 + E(T_f)^2) - 4ln(E(T_f)))$$

and

$$s_t = \sqrt{2(ln(E(T_f)) - \mu_t)},$$

as these faults have been recently decompressed.

The remaining faults are handled with exponential inter-event times (Poisson process). The magnitude of all faults is supposed to follow a truncated below at $m_{low} = E(M_f) - \sigma(M_f)$ and above at $m_{high} = E(M_f) + \sigma(M_f)$ Normal distribution with parameters $E(M_f)$ and $\sigma(M_f)$. The magnitudes of the faults are respectively simulated under a truncated normal distribution using R statistical software and the inverse simulation method. Lastly, the energy released is very intense at the generating point of an occurring earthquake, attenuating as the distance from the epicenter increases. This phenomenon is modelled by ground motion prediction equations (GMPEs), as described further in section 2.3.

2.3 Ground motion prediction equations (GMPEs)

When an earthquake occurs, the produced seismic wave attenuates as the distance from the epicenter to a cite increases. The ground motion caused by that wave propagation is measured by some intensity measures described by ground motion prediction equations (GMPEs). There are many different GMPEs (or attenuation equations) describing this phenomenon. The online resource created by Douglas (2014) summarizes most of the empirical GMPEs and their errors. In the present work, we use the GMPE by Rinaldis et al. (1998), as shown below in equation (2.7). Yucemen (2005) presented research regarding annual insurance pricing for different seismic zones in Turkey using damage probability matrices and based on the expected annual damage ratio caused by the Mercalli intensity scale (MMI). The MMI employs personal reports and observations to measure earthquake intensity, but peak ground acceleration (PGA) is measured by instruments such as accelerographs. PGA can be correlated to macro seismic intensities on the MMI, but these correlations are associated with large uncertainty (Cua et al. 2010). Therefore, PGA is directly used in this work, as in Asprone et al. (2013) and Lin (2018). The major quantity of interest is the peak ground acceleration (PGA), which is assumed to be a random variable as modelled by equation (2.8). The GMPE predicts the mean of the natural logarithm of the PGA accepted by a point on the surface of the earth R kilometres away from the epicentre of an earthquake of magnitude M:

$$E[ln\ (PGA)] = 0.82M - 1.59ln(R + 15) + 5.25, \tag{2.7}$$

where PGA is in units of $\mathrm{cm/sec}^2$. The magnitude of the earthquake that will be produced by a seismic source remains unknown. The natural logarithm of PGA given M,R is normally distributed as in equation (6):

$$ln(PGA \,\big|\, M, R) \sim Normal(E[ln\ (PGA)], \sigma(ln(PGA))), \tag{2.8}$$

where $\sigma = \sigma(ln(PGA)) = 0.68$ for this GMPE constant for all magnitudes and distances. The ground motion accepted by a building due to a strong earthquake then results to economic losses as presented in chapter 3. The vulnerability of each building determines its potential damage state with respect to the size of the ground acceleration. This is modelled via fragility curves. The next section 2.4 presents an epidemic-type model for earthquake recurrence along with the distribution of the associated produced magnitudes.

2.4 ETAS model for historical catalogues

While most actuarial considerations, including section 2.1 of the present thesis, concentrate on models with simpler dynamics, such as the homogeneous Poisson model, this section examines modeling the historical catalog as a branching process (by the epidemic-type aftershock sequence or ETAS model). Based on the stochastic declustering of the events, the maximum magnitude of the background events (modeled as a homogeneous Poisson process) and their descendants is used as the input for the premium rating process in chapter 3.

The Gardner and Knopoff (1972 and 1974) declustering of seismic past events offers the advantage of being a quite simple technique for implementation. Earthquakes are sorted in descending order with respect to their magnitude and the rest events around the larger magnitude earthquakes are discarded in order to result in an independent Poisson catalogue. Thus, the stronger and more catastrophic events are considered as the main events. This declustering constitutes a good approximation but in practice the seismic phenomenon is better described and modeled by a branching process. The epidemic-type aftershock sequence (ETAS) model is the most widely used statistical model to describe earthquake catalogs (Jalilian, 2019; Ogata, 1988; Ogata, 1998 and many others). The problem that arises when estimating actuarial values using the ETAS model is that the actual losses are underestimated with respect to the background rates of the events (Kohrangi et al. 2021). The need of a Poisson distribution for long-term prediction forces the risk analyst to discard aftershocks that include strong events (potentially stronger than the background event that triggers them). In this work, we deal with this problem using the probability distribution of the largest magnitude of the whole cloud of events triggered by a background event including itself (Zhuang et al., 2006, Zhuang, 2012). Zhuang et al. (2006) provided a formula for the distribution function of this largest magnitude considering events larger than a lower bound. We provide the analytical formula considering there is also an upper magnitude limit. This computation is quite useful as events larger than 6 M_w are suggested to be imputed to fault sources instead of area sources. As the upper limit reaches infinity, the provided formula coincides with the formula of Zhuang et al. (2006). Therefore, this analysis is also considering fault sources to deal with the incompleteness of the historical catalogues.

2.4.1 Description of the spatiotemporal ETAS model for historical catalogues

The conditional intensity function of the spatiotemporal ETAS model is described by:

$$\lambda_{B,\theta}(t,x,y,m \,\big|\, H_t) = f_B(m)\lambda_\theta(t,x,y \,\big|\, H_t),$$

where $f_B(m) = Bexp\left(-B(m-m_0)\right), B > 0$ is the probability distribution function (PDF) of the magnitude of an event and

$$\lambda_\theta(t,x,y|H_t) = \tilde{u}(x,y) + \sum_{i:t_i<t} k_{A,a}(m_i)g_{c,p}(t-t_i)f_{D,g,q}(x-x_i,y-y_i;m_i)$$

under the following definitions:

- $\tilde{u}(x,y)$ is the background seismicity rate
- $k_{A,a}(m_i)$ is the expected number of triggered events generated from an event of magnitude m_i equal to:

$$k_{A,a}(m_i) = A\,\text{exp}(a(m-m_0)),$$

- $g_{c,p}(t-t_i)$ is the PDF of the occurrence time of a triggered event by an event of magnitude m_i occurring at time t_i. Based on the modified Omori's law:

$$g_{c,p}(t-t_i) = \begin{cases} \frac{p-1}{c}\left(1+\frac{t-t_i}{c}\right)^{-p} & ,t-t_i > 0 \\ 0 & ,t-t_i \leq 0 \end{cases}$$

- $f_{D,w,q}(x-x_i,y-y_i;m_i)$ is defined as:

$$f_{D,w,q}(x-x_i,y-y_i;m_i) = \frac{q-1}{\pi D\,\text{exp}(w(m_i-m_o))}\left(1+\frac{(x-x_i)^2+(y-y_i)^2}{D\,\text{exp}(w(m_i-m_0))}\right)^{-q}$$

The background events are generated by a stationary in time Poisson process with intensity $\tilde{u}(x, y)$. Previous events, background or triggered, generate further events according to a non-stationary Poisson process with intensity function:

$$\sum_{i:t_i<t} k_{A,a}(m_i) g_{c,p}(t - t_i) f_{D,w,q}(x - x_i, y - y_i; m_i).$$

The unknown parameters c, p, A, a, d, w, q are estimated by the Davidon-Fletcher-Powell algorithmic method (Ogata, 1988) as described in Jalilian (2019) by the ETAS package in R. The simpler case where the conditional intensity function is only dependent on time can be found on Harte (2010).

The total spatial intensity function in an interval of length T is approximated by:

$$\Lambda(x, y) = \tilde{u}(x, y) + \frac{1}{T} \sum_{i:t_i<t} k_{A,a}(m_i) f_{D,w,q}(x - x_i, y - y_i; m_i),$$

used to obtain the clustering coefficient

$$\omega(x, y) = 1 - \frac{\tilde{u}(x, y)}{\Lambda(x, y)}$$

at the point (x, y).

The historical catalogue is presented in the figure 2.6.

Following the declustering, each event is categorized as either background or triggered based on its respective probability, as shown in the Figure 2.7 below.

The probabilities associated with these events are estimated as follows.

Let p_{ij} the probability that event j is triggered by i. Zhuang et al. (2002) proposed that

$$p_{ij} = \begin{cases} \frac{k_{A,a}(m_i) g_{c,p}(t_j - t_i) f_{D,w,q}(x_j - x_i, y_j - y_i; m_i)}{\lambda_\theta(t_j, x_j, y_j | H_t)} & , t_j - t_i > 0 \\ 0 & , \text{otherwise} \end{cases}$$

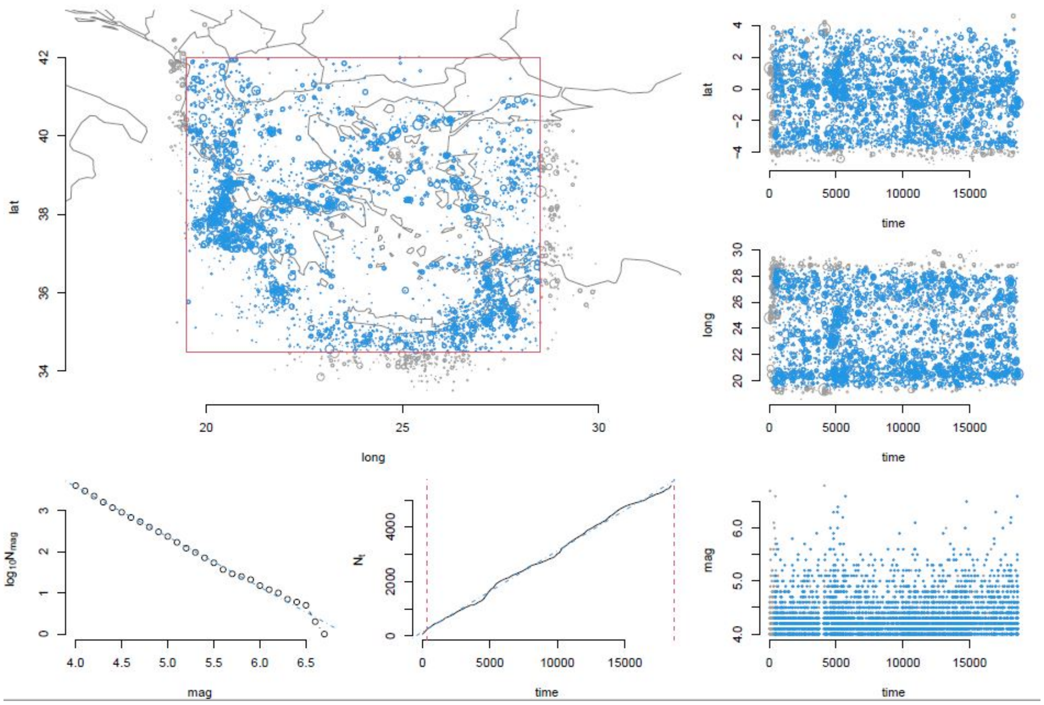

Figure 2.6: Description of the historical catalogue where the ETAS model is fitted

Therefore, $p_j = \sum_{i:t_i<t_j} p_{ij}, j = 1, \ldots N$ is the probability that j is a triggered event and

$$1 - p_j = \frac{\tilde{u}(x_j, y_j)}{\lambda_\theta(t_j, x_j, y_j|H_t)}$$

is the probability that j is a background event.

For computational reasons, it is also assumed that $\tilde{u}(x, y) = \mu u(x, y)$.

The unknown parameters $B, \theta = (\mu, A, a, c, p, D, w, q), u(x, y)$ are estimated via the Davidon-Fletcher-Powell algorithm (Ogata, 1998; Jalilian, 2019). Once these quantities are estimated,

$$\hat{u}(x, y) = \frac{1}{T}\sum_{j-1}^{N}(1 - \hat{p}_j)\phi_k(x - x_j, y - y_j; h_j),$$

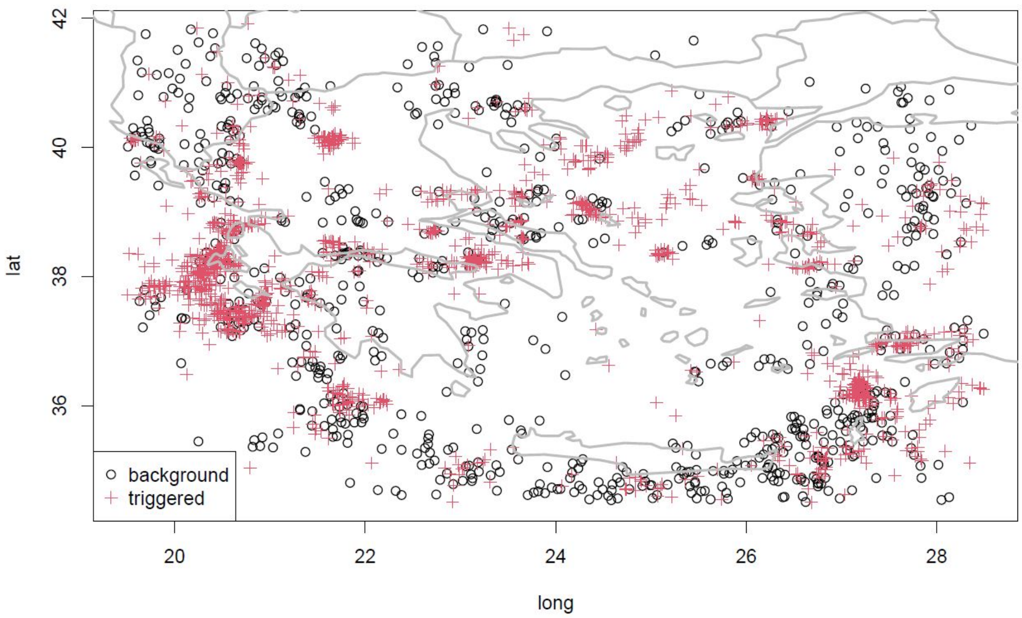

Figure 2.7: Background vs triggering events after stochastic declustering

where

$$\phi_k(x, y; h) = \frac{1}{2\pi h^2} exp(-\frac{x^2 + y^2}{2h^2})$$

is the isotropic Gaussian kernel and h is described in Jalilian (2019).
Then,

$$\hat{\Lambda}(x, y) = \frac{1}{T} \sum_{j=1}^{N} \phi_k(x - x_j, y - y_j; h_j)$$

and

$$\hat{\omega}(x, y) = 1 - \frac{\hat{u}(x, y)}{\hat{\Lambda}(x, y)}.$$

The bandwidth h_j can be found in Jalilian (2019). The estimates provided by the package ETAS are: B=3.61, equivalently the total b-value= 1.56, $a = 1.58$,

$\mu = 0.98$, $A = 0.28, c = 0.03, p = 1.15, D = 0.002, q = 1.64, w = 0.39$. Below, there is a graphic representation of all rates in Figure 2.8.

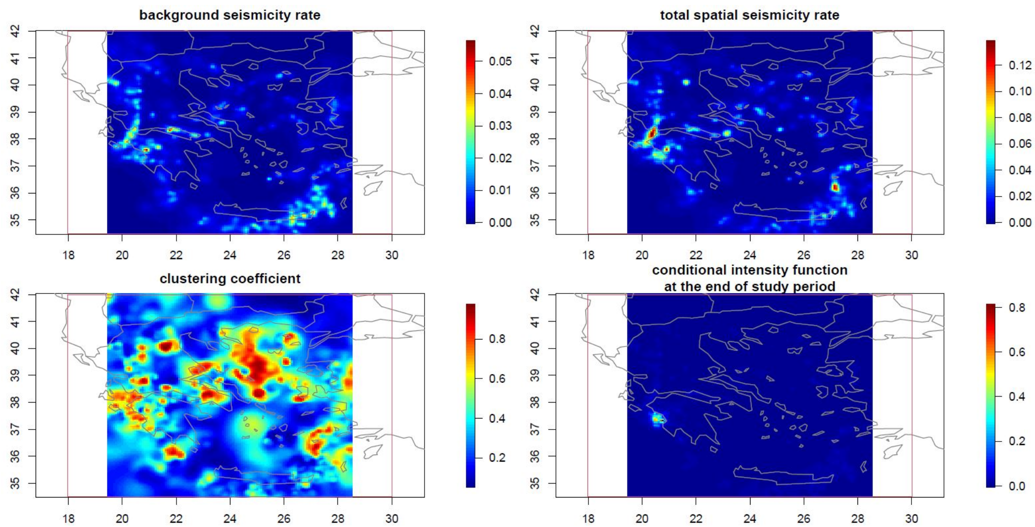

Figure 2.8: Estimated rates by the ETAS model

If we only used background events and their generated magnitudes, there would be a substantial underestimation of the actual risk. Therefore, it is necessary to also obtain the PDF of the maximum magnitude of the cloud of events that a background event carries. This is analytically estimated in section 2.4.2.

2.4.2 PDF of the maximum magnitude of a cloud of events generated by a background event including itself.

The purpose of this section is the analytical estimation of the PDF ϕ of the maximum magnitude over a whole cluster of events generated by a background event. The cumulative distribution function (CDF) according to the Gutenberg-Richter law for magnitudes truncated from both sides is:

$$F_M(m) = \frac{1 - \exp(-B(m - m_{min}))}{1 - \exp(-B(m_{max} - m_{min}))}.$$

The respective PDF, namely s(m), is obtained by differentiating F with respect to m yielding:

$$s(m) = \frac{dF_M(m)}{dm} = \frac{B\,\text{exp}(-B(m - m_{min}))}{1 - \text{exp}(-B(m_{max} - m_{min}))}.$$

Following Zhuang & Ogata (2006), the probability for the largest with respect to magnitude earthquake in an arbitrary cluster including the initial event and all its descendants to be greater than m is computed by:

$$F_c(m) = 1 - \int_{m_{min}}^{m} s(v)\,\text{exp}\left(-k(v)F_C(m)\right) dv.$$

Equivalently,

$$F_C(m) = 1 - \frac{1}{1 - \text{exp}(-B(m_{max} - m_{min}))}[[\Gamma_{-B/a}(AF_C(m))$$

$$-\Gamma_{-B/a}(AF_C(m)\,\text{exp}(a(m - m_{min})))]\frac{B}{a}(AF_c(m))^{B/a}],$$

where $\Gamma_a(x) = \int_x^\infty exp(-u)u^a du$ is the Gamma function.

Considering the fact that:

$$\Gamma_\phi(x_1) - \Gamma_\phi(x_2) = \sum_{n=0}^{\infty} \frac{(-1)^n(x_2^{n+\phi} - x_1^{n+\phi})}{n!(n+\phi)},$$

we get:

$$F_C(m) = 1 - K\left[\frac{B}{a}(AF_C(m))^{B/a}\sum_{n=0}^{\infty}(-1)^n X_n(m)\right],$$

where

$$K = \frac{1}{1 - \text{exp}(-B(m_{max} - m_{min}))}$$

and

$$X_n(m) = \begin{cases} \frac{[AF_C(m)\exp(a(m-m_{min}))]^{n-B/a} - [AF_C(m)]^{n-B/a}}{n!(n-B/a)} & , n \neq B/a \\ \frac{a(m-m_{min})}{n!} & \text{, otherwise} \end{cases}$$

Thus, under Taylor approximation keeping the first two terms:

$$F_C(m) = 1 - K\left[\frac{B}{a}(AF_C(m))^{B/a}(X_0(m) - X_1(m))\right].$$

After some algebra, we get:

$$F_C(m) = 1 - K[-exp - B(m - m_{min})) + 1$$

$$-\frac{B}{a-B}AF_C(m)(-1 + \exp((a-B)(m - m_{min})))].$$

For the special case of K=1 i.e.: $m_{max} \to \infty$, the equation (24) of Zhuang & Ogata (2006) is obtained.

Solving for $F_C(m)$:

$$F_C(m) = \frac{r(m)}{h(m)},$$

where

$$r(m) = 1 + K[\exp(-B(m - m_{min})) - 1]$$

and

$$h(m) = 1 + K\frac{AB}{a-B}[1 - \exp((a-B)(m - m_{min}))].$$

For the special case of K=1 i.e.: $m_max \to \infty$, the equation (25) of Zhuang & Ogata (2006) is obtained.

The PDF of interest is the derivative of $1 - F_C(m)$ with respect to m, as presented below:

$$\phi(m) = -\frac{dF_C(m)}{dm},$$

where

$$\frac{dF_C(m)}{dm} = \frac{r'(m)h(m) - r(m)h'(m)}{h^2(m)}.$$

The shape of ϕ along with generated values from its density using the Rejection Algorithm are presented in Figure 2.9 below. These values are useful for the premium rating and the SCR evaluation in the algorithms of chapter 3 further.

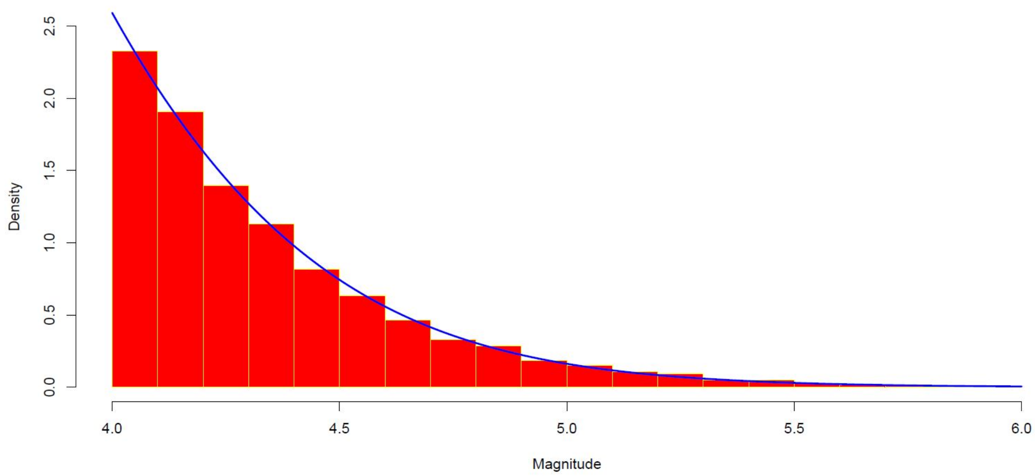

Figure 2.9: PDF of ϕ with a histogram of generated values from this density using the Rejection Algorithm

Continuing in chapter 3, the estimation of losses arising from seismic hazard and their quantification in actuarial terms is discussed.

Chapter 3

Insurance pricing & capital requirements

There are similar works regarding premium rating for seismic risk including Asprone et al. (2013), Lin (2018), Tao et al. (2010) and Yucemen (2005) as well as a published work for SCR calculation by Deligiannakis et al. (2021). None of the models apart from Deligiannakis et al. (2021), consider fault sources. Moreover, their models produce results in terms of zone pricing (or at the municipality and city levels), in opposition to our model producing coordinate precision pricing for the whole country of Greece. Furthermore, their estimates depend on expected loss, which is indeed useful for insurance pricing but always under the certainty equivalence principle of premium calculation. However, in a more reliable risk analysis, the main interest should be concentrated on the total probability measure of the damage ratio, including the extremal events. This was achieved in the present risk assessment using simulations, resulting in a reliable SCR estimate where convex risk measures were also applied, apart from Value at Risk (VaR).

3.1 Fragility curves

There are two main types of empirical methods for conducting the seismic vulnerability assessment of buildings based on the damage observed after earthquakes (Calvi et al., 2006). The first type involves the damage probability matrices (DPM), which express in a discrete form the conditional probability of obtaining a damage level j, due to a ground motion of discrete intensity i according to Whitman et al. (1973). An example of the structure of damage probability matrices is presented in Table 3.1 provided by Meslem & Lang (2017).

The second type involves fragility curves, which are continuous functions expressing the probability of exceeding a given damage state, considering a continuous earthquake intensity measure. Numerous different forms of fragility curves are widely used for vulnerability assessment (Meslem & Lang, 2017). These curves correspond to cumulative distribution functions in their majority, but this is not strict. Examples of fragility curves are presented in Table 3.2 according to a literature review processed by Meslem & Lang (2017). The variable D_S denotes the damage state, while the variable IM denotes the continuous intensity measure.

Table 3.1: An example of DPM

Damage state	Structural Damage	Non- structural Damage	Damage Ratio (%)	Seismic Intensity				
				V	VI	VII	VIII	IX
0	None	None	0-0.05	92	79	33	6	0
1	None	Minor	0.05-0.3	8	18	34	17	6
2	None	Localized	0.3-1.25	0	3	20	39	18
3	Not noticeable	Widespread	1.25-3.5	0	0	10	11	22
4	Minor	Substantial	3.5-7.5	0	0	3	5	27
5	Substantial	Extensive	7.5-20	0	0	0	11	12
6	Major	Nearly total	20-65	0	0	0	6	10
7	Building condemned		100	0	0	0	5	4
8	Collapsed		100	0	0	0	0	1

Table 3.2: Examples of fragility curve functions

Type	$P(D_S \geq d_s(i) \mid IM = x)$	Parameters
Log-Normal CDF	$\Phi\left(\frac{ln(x)-\lambda}{\zeta}\right)$	λ, ζ
Normal CDF	$\Phi\left(\frac{x-\mu}{\sigma}\right)$	μ, σ
Logistic PDF	$\frac{1}{1+exp(-[\theta_0+\theta_1 x])}$	θ_0, θ_1
Exponential Functcion	$1 - exp(-\theta_0 x^{\theta_1})$	$\theta_0, \theta 1$

The levels of damage states (1-9) characterizing fragility curves refer to different damage percentages according to different proposed taxonomies (Pitilakis et al. 2014) or damage scales . Different proposed taxonomies are summarized in Table 3.3.

In this work, the fragility curves outlined by Kappos et al. (2006) and Kappos (2013) for buildings of type RC3.1LL (Reinforced Concrete with regularly infilled frames, Low Rise, Low Code seismic design level) and RC3.1LM (Reinforced Concrete with regularly infilled frames, Low Rise, Moderate Code seismic design level) are used, as they refer to the region of Greece, account for many damage states, and are defined in terms of PGA. The fragility curves by Kappos et al. have the following form:

$$P(D_S \geq d_s(i) \mid \text{PGA}) = \Phi\left[\frac{1}{\beta_{d_s(i)}} \ln\left(\frac{\text{PGA}}{\overline{\text{PGA}}[d_s(i)]}\right)\right],$$

where: $\overline{PGA}[d_s(i)]$ is the median value of peak ground acceleration where the building reaches the threshold of the damage state, $d_s(i)$; $\beta_{d_s(i)}$ is the standard deviation of the natural logarithm of peak ground acceleration for damage state, $d_s(i)$; Φ is the standard normal cumulative distribution function; and D_S is the discrete random variable denoting the damage state.

Table 3.3: Different proposed taxonomies

ATC-13			HAZUS99		
Damage States	Characterization	Damage Ratio	Damage States	Characterization	Damage Ratio
1	None	0	1	None	0-2
2	Slight	0-1	2	Slight	2-10
3	Light	1-10	3	Moderate	10-50
4	Moderate	10-30	4	Extensive	50-100
5	Heavy	30-60	5	Complete	100
6	Major	60-100			
7	Destroyed	100			
FEMA273			VISION2000		
Damage States	Characterization	Damage Ratio	Damage States	Characterization	Damage Ratio
1	Very Light	0-1	9,10	Negligible	0-2
2	Light	1-10	7,8	Light	2-10
3	Moderate	10-30	5,6	Moderate	10-50
4	Severe	30-100	3,4	Severe	50-100
			1,2	Complete	100

The damage states are described in Table 3.4 below. We have defined a loss index uniformly distributed inside each interval. If one wants to be more risk averse, a beta distribution with negative skewness is the ideal distribution for the loss index inside each interval.

Table 3.4: Damage states of Kappos et al.'s fragility curves

Damage State	Range of Loss Index (%)
ds_0	0
ds_1	(0–1]
ds_2	(1–10]
ds_3	(10–30]
ds_4	(30–60]
ds_5	(60–100]

The parameters for the fragility curves are introduced in Table 3.5 and their shapes in Figure 3.1. The probability of occurrence of a certain damage state ds_j is described by:

$$P(DS = d_s(j)|\text{PGA}) = P(DS \geq d_s(j)|\text{PGA}) - P(DS \geq d_s(j+1)|\text{PGA}).$$

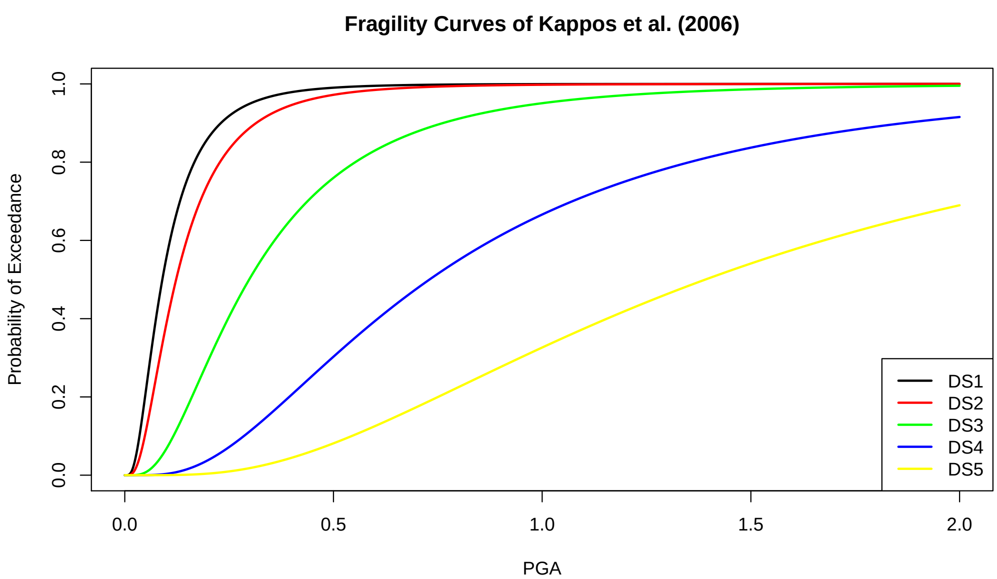

Figure 3.1: Fragility curves of Kappos et al. (2006)

Table 3.5: Parameters of fragility curves of Kappos (2013)

RC3.1 Low Rise Low Code									
DS1		DS2		DS3		DS4		DS5	
Median	Beta	Median	Beta	Median	Beta	Median	Beta	Median	Beta
0.021	0.733	0.101	0.733	0.201	0.733	0.257	0.733	0.343	0.733
RC3.1 Low Rise Moderate Code									
DS1		DS2		DS3		DS4		DS5	
Median	Beta	Median	Beta	Median	Beta	Median	Beta	Median	Beta
0.090	0.733	0.123	0.733	0.298	0.733	0.730	0.733	1.391	0.733

3.2 Loss random variable & risk measures

The pecuniary value of the damage incurred to a building i, denoted by L_i, is defined as the random variable (3.1), where w_i is the random variable corresponding to the maximum loss percentage of the total insured value of the building, TIV_i, over the year:

$$L_i = w_i \times TIV_i \tag{3.1}$$

When the loss due to the N_b insured buildings of an insurance company portfolio (needed for the SCR) is investigated, the loss random variable is as follows:

$$L_T = \sum_{i=1}^{N_b} w_i TIV_i. \tag{3.2}$$

In this work, the case study considers buildings of the same type, thus assuming they have the same $TIV_i = TIV$, so:

$$L_T = N_b TIV \overline{w}$$

and

$$\overline{w} = \frac{\sum_{i=1}^{N_b} w_i}{N_b}.$$

It is obvious from the equations that both loss random variables are completely defined by the random loss indices w_i and $\overline{w}$. These losses are highly affected by large uncertainties because of the insufficient historical data of seismic events, as they are infrequent (Dong et al. 1996; Kunreuther, 1996). The proper risk management of risk L_T ensures the solvency of the insurance company. Therefore, convex risk measures are applied to $\overline{w}$ to deal with these uncertainties. An advantage of convex risk measures is that they can admit a representation of a whole set of probability measures dealing with uncertainty. Therefore, we can avoid the unpleasant position of choosing a unique probability measure, especially as we are not sure if it is sufficiently representative, as is the case with the use of VaR.

Risk measures are mappings from the space of bounded random variables to the real numbers representing the amount of capital that the holder of risk X should add to its position and safely invest to satisfy a regulator (Mc Neil et al., 2015). A commonly used measure of risk (though not a risk measure in the above sense in general) is VaR. According to Mc Neil et al., VaR has certain difficulties; the most important being that it does not always coincide with one of the basic principles in finance—that risk is reduced by diversification between the portfolio assets. Therefore, we define the larger class of convex risk measures (Föllmer & Schied, 2011). Convex risk measures account for uncertainty due to their robust representation (Follmer, 2011). If $\mathcal{M}$ is a set of probability measures on $(\Omega, \mathcal{F})$ and $\alpha : \mathcal{M} \rightarrow \Re \cup +\{\infty\}$, then:

$$\rho(L) = \sup_{\mathcal{Q} \in \mathcal{M}} (E^Q(L) - \alpha(Q))$$

A convex risk measure can be thought as the worst-case scenario which is adjusted by a penalty function showing our belief in each probability measure. A risk measure ρ is said to be convex if it is a risk measure according to Artzner et al. (1999) (properties 1–3 below) and also satisfies the fourth property of convexity. Let L_1, L_2 represent random variables denoting losses.

Properties of convex risk measures:

1. Monotonicity: $\rho(L_1) \leq \rho(L_2)$ for $L_1 \leq\ L_2$.

2. Translation invariance: $\rho(L + c) = \rho(L) + c$

3. Normalization: $\rho(0) = 0$

4. Convexity: $\rho(\lambda L_1 + (1 - \lambda)L_2) \leq \lambda\rho(L_1) + (1 - \lambda)\rho(L_2), \ \ \lambda \in [0, 1]$

A commonly used way of quantifying risk is Value at Risk (VaR), defined as:

$$VaR_a(L) = \text{inf } \{l \in \Re : F_L(l) \geq a\}.$$

A related quantity, which is a coherent risk measure, is Expected Shortfall, defined as:

$$ES_a(L) = E(L|L \geq VaR_a(L)).$$

Positive Homogeneity that coherent risk measures satisfy has been doubted in general by many authors, for example Mc Neil et al., (2015). The reason is that due to positive homogeneity a coherent risk measure does not adapt to liquidity problems that may occur. Therefore, the above problems lead us to define the larger set of convex risk measures. The most common convex risk measure is the entropic risk measure (Föllmer & Schied, 2011). It is defined as follows: Let the exponential utility function:

$$u(x) = 1 - exp(-\theta x)$$

for some $\theta > 0$. The entropic measure of risk induced by this utility function is given by:

$$\rho_\theta(L) = \frac{1}{\theta} lnE(exp(\theta L)).$$

The implementation of a larger θ leads to higher risk aversion. In this thesis, (convex) risk measures are used only with the intention of reserving against losses, while pricing is processed under the certainty equivalence principle (risk premium). A premium calculation principle is defined as a function $\pi : X \rightarrow \Re$ representing the price (premium) that an insurer would charge for insuring risk $X \in \mathcal{X}$ (Tsanakas & Desli, 2005). There is also a variety of methods where risk measures can also be applied for insurance pricing under uncertainty (Peng, 2011 and Escobar & Pflug, 2018) by means of a distortion function. The proposed algorithms for the processed insurance pricing follow in section 3.3.

3.3 The algorithm

1. Generate $N_0 = 100$ uniformly spatially distributed points inside each area source (grid cell).

2. Compute the distances from each insured building of the portfolio to each point of every area source and the centroid of each fault.

3. Generate the number of earthquakes for each area and fault based on their annual activation rates (chapter 2).

4. Generate as many magnitudes as depicted in step 3 based on the b-value for areas and the geometrical properties for faults. For area sources, distances are also generated by resampling from the matrices of step 2.

5. Attenuate each produced magnitude to all buildings using equations (2.7) and (2.8).

6. Transform all peak ground accelerations to damage indices using fragility curves (equation (7)) with respect to the properties of the buildings.

7. Keep the maximum produced annual damage of each building from all sources (random variable L_i) and the summation of these losses for the whole portfolio (random variable L_T).

8. Go to step 3 and repeat until a large number of simulations has been performed.

9. The expected value of L_i is the risk premium, while risk measures are applied to L_T for the SCR evaluation.

The respective algorithm considering the ETAS model for historical catalogues is provided in the Appendix. These algorithms were applied to 100 buildings with specific coordinates in Greece. The portfolio was constructed based on the building census that was carried out by the Hellenic Statistical Authority (ELSTAT) in 2011. It contains 53 buildings in Attica (ATT), 21 buildings in Central Macedonia (CMC), 5 in Thessaly (THE), 4 in Western Greece (WGR), 3 in Easter Macedonia and Thrace (EMT), 3 in Crete (CRE), 3 in Peloponnese (PEL), 2 in Western Macedonia (WMC), 2 in Epirus (EPI), 2 in Central Greece (CGR), 1 in the Ionian islands (ION), and 1 in the South Aegean (SAE).

The distance considered from an area source to a building is a random variable because there is no prior knowledge where on the cell surface the epicenter will be. In the first case, the sources of the source-to-site distance are considered as uniformly distributed in each grid square, while the sources for the ETAS case are chosen as the centroids of each grid.

As a large number of simulations is needed for the actuarial estimations due to the low frequency of seismic events, one may also use the Rcpp package in R for faster computations by cointegrating R and C++ languages (Eddelbuettel, D. & François, R., 2011 and Eddelbuettel, D., 2013).

3.4 Insurance pricing

Risk premium is estimated for low-rise buildings built under low code (up to 1985), moderate code (1986–1995), and high code (similar to the latest seismic Eurocode 8). The risk premium is estimated for each building of an insured value of 200,000 euros based on different models. On average, the moderate-code buildings accept 1/4 of the damage of the low-code buildings. The moderate-code buildings have a strength derived by interpolation between the values estimated for the L and H cases, assuming the distance from the low-code value is twice that of the high-code value (Kappos et al., 2006). Therefore, multiplying the damage of the moderate-code buildings by 2/4, we get an approximation for the high-code buildings.

The expected annual loss (risk premium) and the median annual loss in euros are presented in Table 3.6 and Table 3.7 for Greece and the region of Attica, respectively. We denote A0 as the model with areas without kernel and A1 as the model with areas and kernel. AF00 denotes the model using faults without the weighting function and areas without kernel. In this model, areas are activated up to 6 M_w, and their contributions to faults are cancelled after this threshold. The next model is denoted as AF01, describing the co-integration of areas and faults without kernel but with the weighting function. The model AF10 uses areas with kernel and faults without the weighting function, while the full model AF11 is used to co-integrate areas with the kernel smoother and the faults with the weighting function.

If an insurance company decides to avoid high-risk buildings (adverse selection), then the mean risk premium drops even further. Our algorithm produces risk premiums with coordinate precision. Thus, two buildings have a different premium even if they are too close. This is crucial for identifying buildings at a short distance from active faults or seismically active areas to avoid insuring them. We observed that the weighting function offers an increase in the results because of the activation of stronger events of areas far from faults. Moreover, the kernel smoother lowers the occurrence rate of high-rate tiles while simultaneously slightly increasing the rate of low-rate tiles. A sample of the estimated premium rates for the whole portfolio under the AF10 model are also provided in Table A1 of the appendix. Table A2 includes the respective rates when the ETAS model is applied (AFe model).

In Tables A1 and A2 of the appendix, the first column refers to the longitude degrees, the second to the latitude degrees, the third to the region, and the fourth to the postal code. Furthermore, L denotes the premium ratings of the low-code buildings, M the moderate-code buildings, and H the high-code buildings.

It is worth mentioning that the insurance pricing process must be carried including fault sources. Inspecting Table 3.6 and the simpler models A0 and AF00, the addition of faults for the seismic hazard increases the risk premium by an amount of 10%. The increment may even reach 18% (AF11 vs. A1). Moreover,

Table 3.6: Overall risk premium estimation for Greece

Risk Premium	Low Code		Mode Code		High Code	
Model	Mean	Median	Mean	Median	Mean	Median
A0	797	730	202	169	101	84
A1	800	720	207	178	103	89
AF00	846	768	222	193	111	96
AF01	891	817	238	207	119	103
AF10	887	780	235	211	117	105
AF11	930	871	246	224	123	112

Table 3.7: Overall risk premium estimation for the region of Attica

Risk Premium	Low Code		Mode Code		High Code	
Model	Mean	Median	Mean	Median	Mean	Median
A0	630	638	150	147	75	73
A1	648	635	155	165	77	82
AF00	691	702	169	172	84	86
AF01	737	757	186	188	93	94
AF10	717	737	184	172	92	86
AF11	835	819	211	213	105	106

in the region of Attica (Table 3.7), model AF11 adds 36% to the risk premium estimated by model A1 for the moderate code.

Therefore, insurance companies should use area and fault co-integration models to estimate the premium rating. This is highlighted even further when using other premium principles rather than the certainty equivalence.

The addition of faults is important for the results, as seen in Figure A1 and Figure A2 of the Appendix comparing the model A1 considering area sources and AF10 adding faults. Thanks to their coordinate precision, these algorithms are useful for an insurance company in order to identify and avoid or handle adverse selection but simultaneously remain competitive. Our estimates are in accordance with the premiums the insurance companies are charged in Greece by the largest reinsurance companies, which are roughly 100 euros annually (for 200000 euros of insured value at most, using historical catalogues).

The model proposed offers a significant advantage for insurance pricing. Premium rates are no longer constant over large areas, but their size is determined with respect to their exact coordinates. Therefore, buildings close to seismically active areas or faults that have not been recently decompressed have higher premium rates. For the use of the ETAS model, the respective premium rates are presented in Table 3.8 and Table 3.9 for Greece and the region of Attica respectively.

Table 3.8: Overall risk premium estimation for Greece using the ETAS model

Risk Premium	Low Code		Mode Code		High Code	
Model	Mean	Median	Mean	Median	Mean	Median
Ae	1035	862	282	214	141	107
AFe	1118	964	307	240	153	120

Table 3.9: Overall risk premium estimation for Attica using the ETAS model

Risk Premium	Low Code		Mode Code		High Code	
Model	Mean	Median	Mean	Median	Mean	Median
Ae	836	789	225	208	112	104
AFe	931	902	253	239	126	119

Comparing Table 3.6 with Table 3.8 and Table 3.7 with Table 3.9, it is obvious that the ETAS model yields higher premium estimates than the model based on GK declustering. These increases must be considered by the respective companies as the earthquake process is better described by an epidemic behaviour. It is also worth mentioning that the actual rates of the reinsurance market are in line with the ETAS model in median terms opposed to the GK declustering model where they are in line in average terms. Inspecting Table 3.8 and the models Ae and AFe, the addition of faults for the seismic hazard increases the risk premium by an amount of 9%. The increment may even reach 12.5% in the region of Attica (Table 3.9).

3.5 Capital requirements

In terms of the related insurance, companies must conform to the requirements of Solvency II and hold sufficient reserves. The European Directive Solvency II, which came into effect in 2016, sets the quantitative requirements for the measurement of risks. It also allows insurance companies to use internal models completely or partially (along with or substituting the SF) to calculate the solvency capital requirement (SCR) induced by a certain type of risk. Each risk must be separately computed so that the total SCR is an overall value depending on these risks and the dependence among them. The standards for capital requirements are no longer related to the number of premiums, as with Solvency I, but rather to the magnitude of the risks. Thus, the existence and use of a seismic risk model is crucial for the quantification of this risk component. A robust model is also essential for any given society, as it guarantees that insurance companies are capable of compensating their customers. Consequently, the balance of the economy is preserved by considering the potential extreme losses caused by earthquakes.

The SCR for this portfolio was calculated equal to 405000 euros by the SF of Solvency II. The SCR is defined as the unexpected loss. Thus,

$$\text{SCR(Loss)} = \text{Risk Measure(Loss)-Expected(Loss)}.$$

Our algorithm is applied to evaluate the SCR for different models (A0 to AF11) and Ae, AFe, building codes, and risk measures. The entropic risk measure is a function of the risk aversion parameter $p(\theta)$. We estimate the mean value of this function in an interval $[r_1, r_2]$, where r_1 is the risk aversion parameter assigning the entropic risk measure a value equal to VaR (99.5%), while r_2 is the parameter assigning the entropic risk measure a value equal to ES(99.5%). The results are presented in the tables below.

It is worth mentioning that the SF for the estimation of the SCR does not consider the construction year of each building leading to overestimation or underestimation. The proposed model deals with that weakness. Therefore, the results presented in Table 3.10 prove that the SCR for moderate-code buildings using VaR and model AF10 is 67% of the SCR induced by the SF. This weakness of the standard formula forces an insurance company that in general undertakes modern buildings to hold reserves without substantial reason, while at the same time underestimates the essential reserves for an insurance company undertaking a risky portfolio.

The first consequence has a negative impact on the investment of these capitals, but the second could even make it insolvent towards its customers. For example, the SCR estimated using high-code buildings, model AF10, and VaR is only 33% of the SCR proposed by the SF. The same estimates are processed using the cointegration of the ETAS model with fault sources. For the use of the ETAS model, the respective SCR is presented in Table 3.11 below.

Comparing Tables 3.10 and 3.11, the ETAS model yields higher estimates of the SCR than the Gardner-Knopoff declustering combined with Voronoi polygons. However, these estimates are always lower than the estimate of the SF for the low-code and moderate-code buildings when the VaR is used as a risk measure. The SCR induced by the ETAS model is closer to the SCR induced by the SF than the SCR induced by the first A0 to AF11 models.

Table 3.10: SCR evaluation in thousand euros

Low Code	VaR	ES	Entropic
A0	538	775	669
A1	612	892	759
AF00	646	1168	924
AF01	751	1277	1027
AF10	699	1137	932
AF11	824	1413	1135
Mode Code	VaR	ES	Entropic
A0	196	289	246
A1	201	336	274
AF00	241	402	325
AF01	252	496	381
AF10	271	425	354
AF11	259	475	371
High Code	VaR	ES	Entropic
A0	98	144	123
A1	100	168	137
AF00	120	201	162
AF01	126	248	190
AF10	135	212	177
AF11	129	237	185

For the use of the Voronoi model, the SCR compared to the standard formula is presented in Table 3.12. For the use of the ETAS model respectively, the SCR compared to the SF is presented in Table 3.13 below. While the difference of the induced SCR for the A0 to AF11 models compared to the standard formula's SCR ranges from -52% to +22%, the respective range for the Ae and AFe models is from -29% to +22%.

Table 3.11: SCR evaluation in thousand euros using the ETAS model

Low Code	VaR	ES	Entropic
Ae	897	1244	1084
AFe	988	1503	1279
Mode Code	VaR	ES	Entropic
Ae	288	420	360
AFe	338	496	425
High Code	VaR	ES	Entropic
Ae	144	210	180
AFe	169	248	212

Convex risk measures used in this work extend the most common measure of VaR and propose a larger SCR. Insurance companies should consider those higher values as convex risk measures cover the large uncertainty of losses due to natural catastrophes.

Table 3.12: SCR compared to the standard formula using Voronoi polygons and faults

Low Code	VaR	ES	Entropic
A0	+33%	+91%	+65%
A1	+51%	+120%	+87%
AF00	+60%	+188%	+128%
AF01	+85%	+215%	+154%
AF10	+73%	+181%	+130%
AF11	+103%	+249%	+180%
Mode Code	VaR	ES	Entropic
A0	-52%	-29%	-39%
A1	-50%	-17%	-32%
AF00	-40%	-1%	-20%
AF01	-38%	+22%	-6%
AF10	-33%	+5%	-13%
AF11	-36%	+17%	-8%
High Code	VaR	ES	Entropic
A0	-76%	-64%	-70%
A1	-75%	-59%	-66%
AF00	-70%	-50%	-60%
AF01	-69%	-39%	-53%
AF10	-67%	-48%	-56%
AF11	-68%	-41%	-54%

Table 3.13: SCR compared to the standard formula using the ETAS model and faults

Low Code	VaR	ES	Entropic
Ae	+121%	+207%	+168%
AFe	+144%	+271%	+216%
Mode Code	VaR	ES	Entropic
Ae	-29%	+4%	-11%
AFe	-17%	+22%	+5%
High Code	VaR	ES	Entropic
Ae	-64%	-48%	-56%
AFe	-58%	-39%	-48%

Chapter 4

CAT bond pricing

Since 2004, private property consists primarily of expensive and valuable buildings constructed according to the latest seismic code, Eurocode 8. Every citizen and owner of a building located within a developed European country would like to ensure that his investment is secured and protected against a possible large loss. As a result, they purchase an insurance policy for natural disasters. In most cases, homeowner insurance policies do not cover earthquake damage, with the exception of Turkey, where earthquake insurance is mandatory. The low-frequency, high-severity nature of this catastrophe risk encourages insurance companies to engage in reinsurance contracts. However, the balance sheets and creditworthiness of reinsurance companies are highly sensitive to catastrophic events (European Central Bank, 2005)[1]. Although the available capital of (re)insurance companies is very large, it is still insufficient to deal with extreme catastrophic losses. Thus, the insurance industry developed a new financial instrument called a catastrophe bond or CAT bond (Polacek 2018). Thus, the insurance risk is transferred to the capital market (Stupfler et al. 2018).

[1] Financial Stability Review (2005), European Central Bank
https://www.ecb.europa.eu/pub/financial-stability/fsr/focus/2005/pdf/ecbf̃8e9aaa84d.fsrbox200505_15.pdf?acc257e30d1d876a16086c7a30a11c51

A CAT bond is a special kind of defaultable bond (Duffie, 1999). Its default risk is primarily associated with the natural hazard, but it may also include credit risk in special cases. Therefore, induced payments for the product are uncertain, and the likelihood of their occurrence is subject to a parameter of the earthquake. This could be a natural earthquake parameter as the local magnitude or depth as in Zimbidis et al., 2007 and Shao et al., 2015, or a damage index as in Loubergé et al. (1999). It should be noted that in the latter work, the default due to natural hazard only affects the sequence of coupon payments (while the face value is assured), whereas in the former works, including the present thesis, the coupon rates and the par value are both under risk. These bonds are inherently risky, rated BB in general, (see for example the Greek Government bonds rated by Fitch[2]). This is also justified from the fact that the regulation (EU) 2016/1799 of the Official Journal of the European Union defines the mid-default rate in a 3-time horizon to be equal to 7.5%[3] (varying from 2.40% to 11%), while the default rate in the 3-time horizon of the proposed CAT bond is 6% as seen further in this work.

In the present empirical analysis, bond is priced from the capital market perspective. With regard to the bond pricing, there are several methods that have been employed and most of them are based on Monte Carlo simulation as in Vaugirard (2003) and Romaniuk (2003). Arbitrage-free pricing is employed in several works as in Zimbidis et al. (2007), Shao et al. (2015) and Jarrow (2010) while in other works risk neutral pricing is employed (see Ma et al., 2013 and Nowak et al., 2013). In this work, we additionally employ CAT bond pricing based upon credit risk, in the simplified scenario where cat bonds are issued by the insurers themselves, as described in Lee & Yu (2002).

CAT bonds require a third-party mediator known as a Special Purpose Vehicle (SPV). The analytical transaction between the parties involved in a CAT bond,

[2]http://www.worldgovernmentbonds.com/country/greece/

[3]https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R1799&from=en# d1e32-12-1

namely the sponsor, the SPV, and the investor, has been detailed in several works, such as Galeotti et al. (2013) and Stupfler et al. (2018). Härdle et al. (2010) examine the calibration of a CAT bond issued by CAT-MEX Ltd in 2006 and sponsored by the Mexican government. In practice, the insurer (e.g., a government) enters into a reinsurance agreement with a reinsurance company. The reinsurance company reaches an agreement with the SPV for the larger potential losses while remaining capable of paying the lower magnitude losses alone. In return for the capital received from the investors, the SPV can return interest rates and the face value if the bond is not triggered. If the bond is triggered, the SPV will pay to the reinsurer the invested capital used to protect the insureds against earthquake losses (see Figure 4.1). Härdle et al (2010) highlighted the importance of a modelled loss trigger mechanism as well as the missing information of losses. In order to overcome these problems, we use the recurrent activation times of seismic faults, i.e., the source of large earthquakes that can trigger the CAT bond. Simulations are used to model losses based on fragility curves.

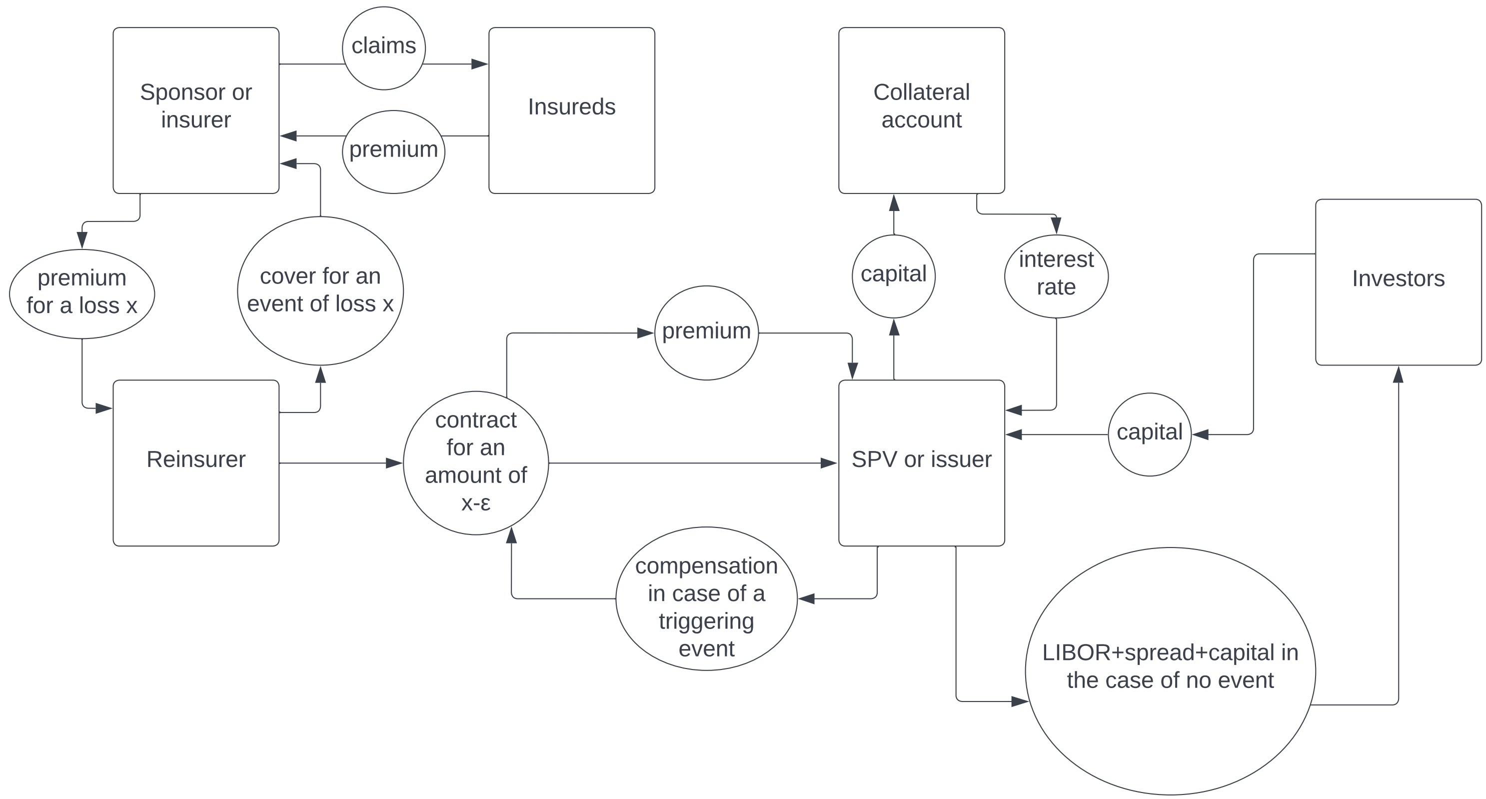

Figure 4.1: CAT bond description

4.1 Description of the CAT bond

The payment scheme of the CAT bond by the capital market aspect is presented in Figure 4.2 below. The coupon payment in the case of no triggering event consists of the coupon rate C_t which is determined by the 12-month Libor rate r_t increased by a spread c_t and multiplied by the face value FV (assumed to be 100 euros) i.e.,

$$C_t = (r_t + c_t)FV \ , t = 1, \ldots 10 \text{ years}$$

.

There are two special cases to consider. The first is the case where the investors are also the insureds and the second is where the insurer is also the issuer. The 12-month Libor Rate is described by the Vasicek's interest short rate model. The data are obtained by the Federal Reserve Bank of St. Louis (Figure 4.3) containing the daily announced 12-month LIBOR rates from 2000 up to 2015, a long time-period including the financial crisis.

Because of the break in the financial series due to the financial crisis, two different time-periods are considered for the estimation. The first time-period is from 2000 up to and including 2009, while the second is from 2010 up to 2015.

The SDE of the Vasicek interest rate model is:

$$dr_{1t} = a_v(b_v - r_{1t})dt + \sigma_v dW_{1t}$$

with the Euler discretization (Iacus, 2009):

$$r_{1,k+1} - r_{1,k} = a_v(b_v - r_{1,k})\Delta k + \sigma_v[W_1(k+1) - W_1(k)],$$

where $W_1(t)$ is a standard Brownian motion and r_k denotes the value of the 12-month LIBOR on day k. This SDE has the solution:

$$r_{1t} = r_{10} \exp(-\alpha_v t) + b_v(1 - \exp(-a_v t)) + \sigma_v \exp(-a_v t) \int_0^t \exp(-a_v s) dW_{1s}.$$

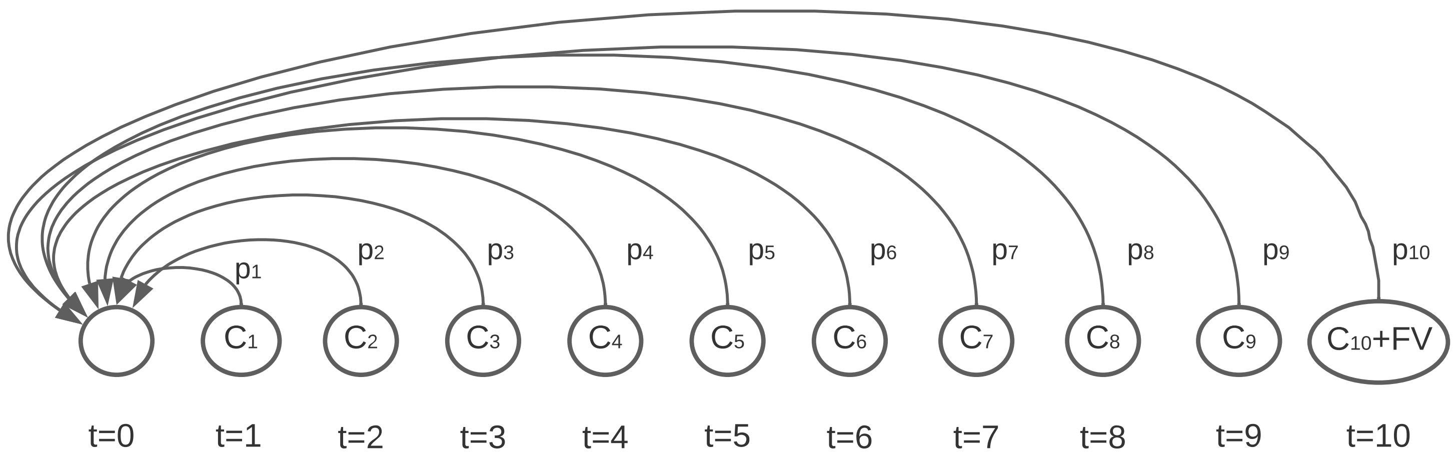

Figure 4.2: Payment Scheme of the CAT bond, C_t: Coupon Rate at each time t, FV: Capital and p_t the probability of occurrence of the payment at time t

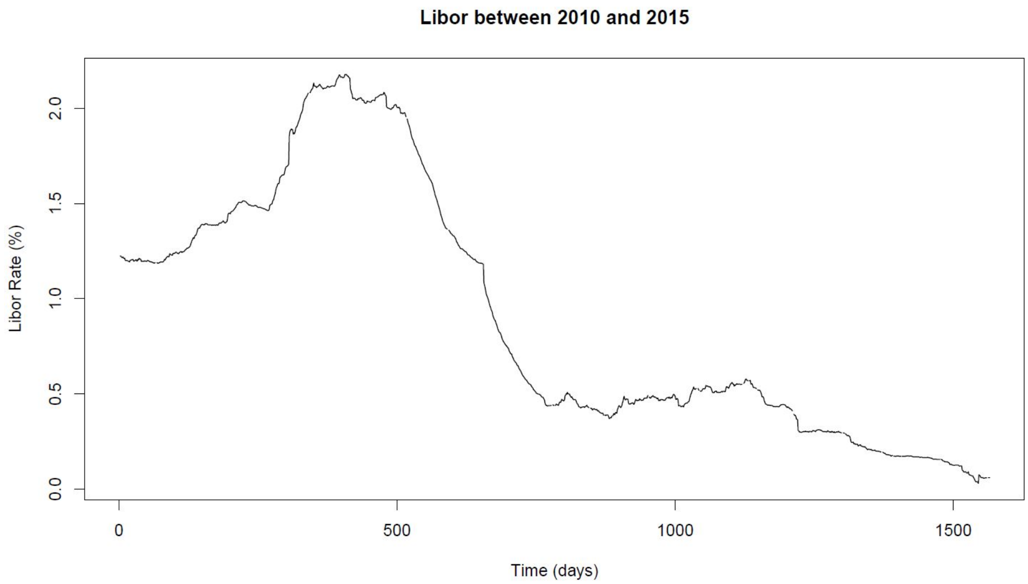

Figure 4.3: 12-month LIBOR announced on a daily basis from 2000 to 2015.

Therefore,

$$r_{1t} \sim N\left(r_{10}\exp(-\alpha_v t) + b_v(1 - \exp(-a_v t)), \frac{\sigma_v^2}{2a}(1 - \exp(-2a_v t))\right).$$

The parameters of the SDE are estimated under the Bayesian methodology using HMC by the package “rstan” (see the respective code in the Appendix) or by MLE. Bayesian Monte Carlo algorithms are also used for the parameter estimation of SDEs in other studies as in Golightly et al. (2010). The average of each parameter’s posterior sample is considered as the estimate. The posterior mean is the Bayes estimator under the squared error loss function. The same exact results are also obtained using maximum likelihood estimation (MLE). The transition density is normal and the parameters can be obtained by differentiating the log-likelihood function with respect to each of the three parameters on the observed rates. The equations for obtaining the parameters are provided in Fergusson, K. & Platen, E. (2015).

The main parameter of interest is b_v which is the value of the 12-month Libor Rate in the long run. Therefore, the expected value of each of the t coupons (without any other risk considered) would be equal to $b_v + c_t$. However, a CAT bond is a defaultable bond. The expected payments are affected by the probabilities of "default" due to seismic events. Thus, these probabilities are multiplied by b_v and the last element of the probability vector also by the nominal value (100 euros) to get the expected payments at times $t = k,\ k = 1, \ldots, 10$. From 2000 up to 2009, the parameters are estimated as:

$$a_v = 0.008, b_v = 3.23\%, \sigma_v = 0.15\%,$$

while the respective parameters from 2010 to 2015 are:

$$a_v = 0.01, b_v = 0.85\%, \sigma_v = 0.11\%.$$

If a triggering event occurs, then the remaining coupons and the face value are not paid to the bondholder (investor). Due to the risk involved in such an investment, CAT bonds offer greater yields than traditional bonds. The trigger parameter in this study is based upon the total material loss of the most destructive earthquake that occurred in Greece. It was the earthquake of Athens in 1999, resulting in approximately 3-4 billion dollars in damages (approximately 3 billion euros).

4.2 Discounting process

The price of any bond strongly depends on the discounting process. The most common methods for bond pricing consist of using short rate processes or forward rate processes. There are two categories for the forward rate processes (De La Grandville, O., 2003), the empirical forward rate curves (e.g., Nelson-Siegel, Svenson or splines) and the general Heath-Jarrow-Morton (Heath et al. 1992) model of forward rate curves (HJM) of which the Vasicek model is a special case.

Let $r_{2,t}$ the short rate used for discounting. In this work, this rate is the 3-month Interbank rate of Germany. The stochastic differential equation of Vasicek is:

$$dr_{2,t} = a_d(b_d - r_{2,t})dt + \sigma_d dW_{2t}.$$

The respective pricing kernel or discount factor is:

$$B(t) = E^Q[exp(-\int_0^t r_2(u)du))].$$

A closed form solution of the Vasicek interest rate model is presented in Nowak (2013). More specifically, the discounting factor at t=0 of a payment on time T denoted by B(T) is equal to:

$$B(T) = exp(-TR(T, r_0)),$$

where:

$$R(x,y) = R_\infty - \frac{1}{a_d x}\left[(R_\infty - y)(1 - exp(-a_d x)) - \frac{\sigma_d^2}{4a^2}(1 - exp(-a_d x))^2\right]$$

and

$$R_\infty = b_d - \frac{\lambda \sigma_d}{a_d} - \frac{\sigma_d^2}{2a_d^2}.$$

For comparison purposes, we proceed in an arbitrage-free context, assuming that the market price of interest rate risk $\lambda = 0$. Short-rate discounting models are not well suited for long-maturity bonds, as demonstrated below. Simulation-based pricing and the relation of the HJM with the Vasicek model are also detailed in the Appendix. Comparisons between the results using various discounting methods follow in the next sections. CAT bond pricing estimation with respect to the overnight Libor is also investigated. However, it does not yield sufficient estimates as described in the Appendix.

4.2.1 Short-rate Vasicek model for the 3-month interbank rate of Germany

A much more rational pricing is provided considering the 3-month interbank rate of Germany as the short rate under consideration. Specifically, the estimated parameters of the Vasicek model using "rstan" are:

$$a_d = 0.03, b_d = 0.88\%, \sigma_d = 0.5\%,$$

$B(4) = B(1\text{year}) = 0.97$ euros at time T=0 for a face value of 1 euro at time T=1 year.

The respective parameters for the time period 2010-2015 are:

$$a_d = 0.03, b_d = -0.71\%, \sigma_d = 0.17\%.$$

The prices of zero-coupon bonds with a face value of 1 euro for maturities 1 to 10 with respect to the two different historical periods are presented in Table 4.1.

Table 4.1: Prices of zero-coupon bonds using the 3-month IBR of Germany

Maturity	2000-2009	2010-2015
1	0.97	1.00
2	0.94	1.01
3	0.92	1.02
4	0.90	1.03
5	0.88	1.05
6	0.86	1.07
7	0.84	1.09
8	0.83	1.11
9	0.82	1.13
10	0.81	1.16

4.2.2 Yield curves

Short rate models as the one used by Vasicek do not also provide forecasts, but the pricing process assumes that short rates will follow the same behaviour as historically. Therefore, the pricing of the CAT bond in this work is mainly based on the time-dependent yield curves of the European Central Bank (ECB), which reflect a much more rational bond pricing able to adapt to the negative interest periods as well. Moreover, these curves offer different yields based on the maturity time. Therefore, their induced discount factors may be larger than 1 for the first one or two years if the economic conditions are not favorable and then become lower than 1. This does not happen when discounting with short-rate models. The data of the ECB euro-area yield curves are presented in the graph below (Figure 4.4). The models explaining the yield curves are summarized below for both the time-independent and time-dependent cases.

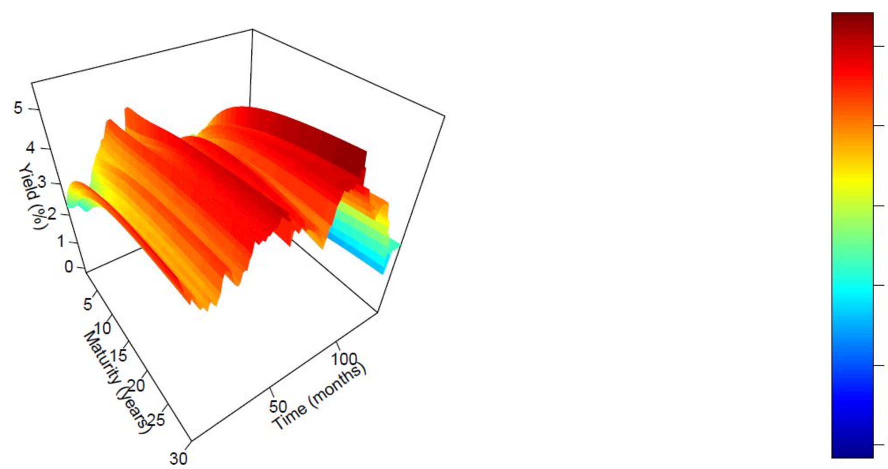

Figure 4.4: ECB yield curves. In this graph, x-axis shows the time the yield curves announced with 1 denoting September 2004 up to 136 denoting December 2015. The y-axis depicts the maturity in years while the z-axis refers to the yields (%).

The forward rate curves of Nelson-Siegel are analyzed below. The forward rate curve is described as:

$$f(T) = b_0 + b_1 \exp(-\frac{T}{\lambda}) + \frac{b_2}{\lambda} T \exp(-\frac{T}{\lambda}),$$

where b_0, b_1, b_2, λ are parameters to be estimated and T is the time to maturity. The future spot rate curve or yield curve is obtained through averaging the forward rate curve and equal to:

$$R(T) = b_0 + (b_1 + b_2) \frac{1 - \exp(-\frac{T}{\lambda})}{\frac{T}{\lambda}} - b_2 \exp(-\frac{T}{\lambda}).$$

The pricing kernel using the yield curves is

$$B(T) = exp\,(-TR(T)).$$

Following the reasoning of Diebold and Li (2006), the lambda shape parameter is pre-determined to result in order to achieve reduction to a linear problem. This parameter determines the maturity at which the loading on the medium-term, or curvature, factor achieves it maximum. Diebold and Li (2006) state that two or three-year maturities are commonly used in that regard and thus they consider the average. In order to provide a more objective interpretation, the lambda parameter is estimated by the package rstan, and it is graphically illustrated in our work to facilitate a better adaptation to our data set. It appears that lambda is equal to 10. The graphs in Figures 4.5 and 4.6 below illustrate that the fit is of high quality.

We assume there is variability between R(T) and our data described by a normal distribution as in Das (2019). More specifically,

$$R_i(\tau_j) = \mu_i(\tau_j) + e_i j,\ e_{ij} \sim N(0, \sigma_3^2).$$

The total likelihood is:

$$L(\text{data}|\theta) = \prod_{i=1}^{n} \prod_{j=1}^{m} f(\mu_i(\tau_j), \sigma_3^2).$$

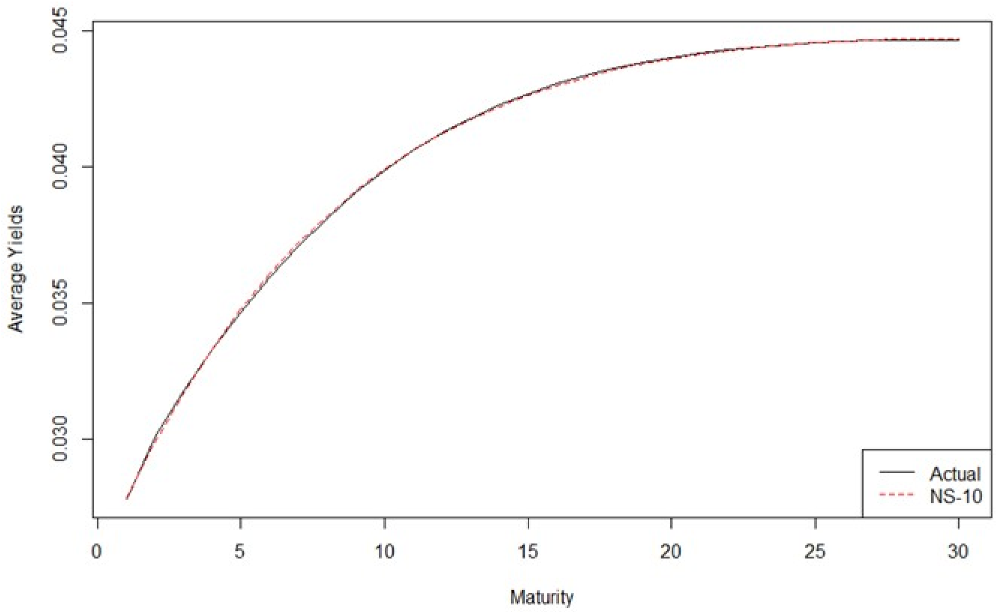

Figure 4.5: Average Actual Yield Curves vs Yield Curve fitted with the NS-model (shape parameter=10) for the time-period 2004-2009.

After defining the likelihood and by using flat priors, the average of the posterior samples obtained from "rstan" yields us the estimates of all parameters of the defined model. The same exact estimates are obtained by MLE. In order to check the goodness of fit, we compare the one step ahead cross validation out of sample MSE of the yield curves across the different used models later in the thesis.

For the sample period between 2004 and 2009, the estimates of the model are:

$$b_0 = 0.04, b_1 = -0.015, b_2 = 0.033, \sigma_3 = 0.005, \lambda = 10.$$

For the period between 2010 and 2015, the estimates of the model are:

$$b_0 = 0.03, b_1 = -0.03, b_2 = 0.06, \sigma_3 = 0.01, \lambda = 10.$$

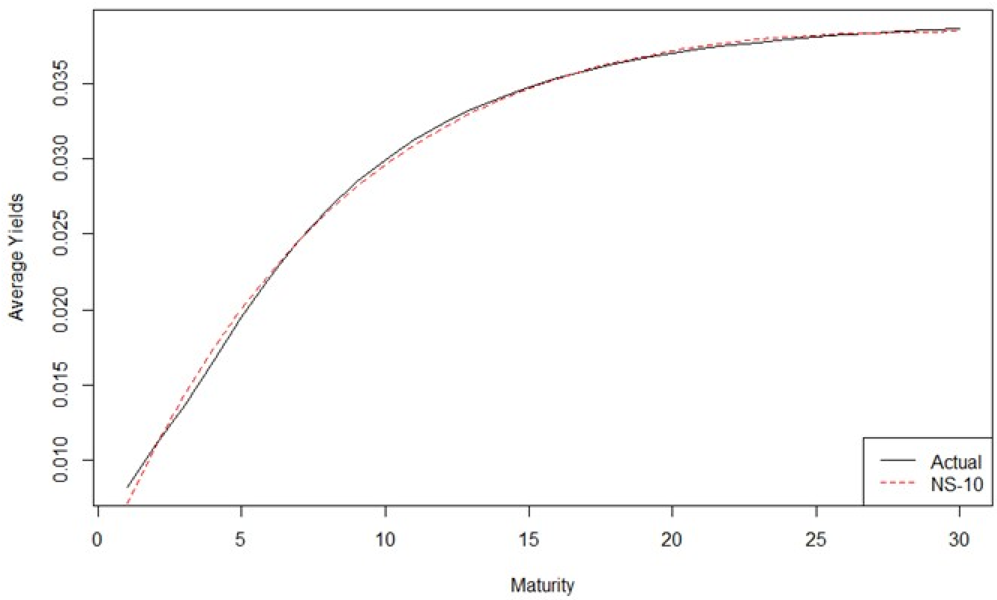

Figure 4.6: Average Actual Yield Curves vs Yield Curve fitted with the NS-model (shape parameter=10) for the time-period 2010-2015.

The respective zero-coupon prices for maturities from 1 to 10 years ahead are summarized in Table 4.2.

Inspecting Table 4.1, we observe that there is a constantly increasing spread between the different time-span discount factors being functions of the 3-month IBR of Germany and the Vasicek model. On the other hand, the respective spread made by the different time-span ECB yield curves (Table 4.2) is not so volatile and pessimistic for the 2nd time period.

The Svensson (1995) model adds flexibility to the Nelson-Siegel formulation by adding a potential extra hump in the forward curve. Its equation is:

$$f(T) = b_0 + b_1 \exp(-\frac{T}{\lambda_1}) + b_2 \frac{T}{\lambda_1} \exp(-\frac{T}{\lambda_1}) + b_3 \frac{T}{\lambda_2} \exp(-\frac{T}{\lambda_2}).$$

Table 4.2: Prices of zero-coupon bonds using the yield curves of the ECB and the Nelson-Siegel model

Maturity	2004-2009	2010-2015
1	0.97	0.99
2	0.94	0.98
3	0.91	0.96
4	0.88	0.93
5	0.84	0.90
6	0.81	0.87
7	0.77	0.84
8	0.74	0.81
9	0.70	0.78
10	0.67	0.74

This model is widely used in practice: It was used by most central banks, including the Bank of England, between 1995 and the end of the 20th century. Averaging the forward rate curve, the spot rate curve is equal to:

$$R(T) = b_0 + b_1 \frac{1 - \text{exp}(-\frac{T}{\lambda_1})}{T\lambda_1} + b_2 \left(\frac{1 - exp(-\frac{T}{\lambda_1})}{T\lambda_1} - \text{exp}(-\frac{T}{\lambda_1}) \right)$$

$$+ b_3 \left(\frac{1 - exp(-\frac{T}{\lambda_2})}{T\lambda_2} - \text{exp}(-\frac{T}{\lambda_2}) \right).$$

The extra shape hump appears to be statistically insignificant for the analyzed data for all positive λ_2. Specifically, the lowest p-value for the analogous t-test is equal to 0.17 for the respective term corresponding to λ_2 being equal to 1.27 for the time period 2010-2015 and the lowest p-value for the analogous t-test is equal to 0.55 for the respective term corresponding to λ_2 being equal to 0.8 for the time period 2004-2009. Therefore, only a single hump is needed in order to model the shape of the specific yield curves. However, we still need to model the upward or downward movement of the curves for each period. This is processed by time-dependent algorithms as described in the following section.

4.2.3 Time dependent (dynamic) cases

In this point, we also consider the announced month t of the yield curve. By using time dependent models, we may predict the yield curve for each time period (month) in order to have a forecast for the future price of the CAT bond.

Following the model of Nelson and Siegel (Das, 2019), the unobserved latent factors b_t are assumed to follow the system equation:

$$\underline{b}_t = \underline{\theta}_0 + Z\underline{b}_{t-1} + \underline{h}_t,$$

where

$$\underline{b}_t = (b_{1t}, b_{2t}, b_{3t})^{,} \ \ \underline{\theta}_0 = (\theta_{01}, \theta_{02}, \theta_{03})^T,$$

$$Z = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{pmatrix}$$

and $h_t \sim N_3(0, S)$ i.e., a VAR(1) model. The general multivariate ARMAX model is discussed in Shumway & Stoffer (2010) and the VAR implementation in R by Cowpertwait & Metcalfe (2009). The yield curves at time t obey the observation equation:

$$\underline{y}_t = \Phi\underline{b}_t + \underline{\epsilon}_t,$$

where

$$\Phi = \begin{pmatrix} 1 & f_1(1) & f_2(1) \\ \vdots & \vdots & \vdots \\ 1 & f_1(30) & f_2(30) \end{pmatrix}.$$

The functions f_1 and f_2 are defined as:

$$f_1(u) = \frac{1 - \exp(-\frac{u}{\lambda})}{\frac{u}{\lambda}},$$

$$f_2(u) = f_1(u) - \exp(-\frac{u}{\lambda})$$

and

$$\underline{\epsilon}_t \sim N_{30}(0, \sigma_\epsilon^2 I_{30}).$$

The unobservable beta factors in the Nelson-Siegel formulation determine the dynamics of the yield curve over time. This means that if one can forecast these factors, one is directly able to forecast the yield curve and the price of the CAT bond as well. Diebold and Li (2006) recognize that the factors in the Nelson-Siegel equation are strongly dependent over time, which suggests that they are forecastable.

Applying ordinary least squares to the yield data for each month gives us a time series of estimates of b_{1t}, b_{2t}, b_{3t} and a corresponding panel of residuals, or pricing errors (Diebold and Li, 2006). The estimation of betas allows us to model them by any arbitrary model (statistical or machine learning) and then use the observation equation for the forecast of the yield curve. In this manner, we also use neural networks for the latent factors and compare the results of the different models by out-of-sample MSE applied on the yields.

The VAR model is based on the assumption that the series are co-stationary implying that each of the time series is separately stationary. However, each of the considered time series $\{b_{1t}\}_t, \{b_{2t}\}_t, \{b_{3t}\}_t$ has a stochastic trend based on the augmented Dickey-Fuller (2009) and Philipps-Perron (1988) tests. Therefore, immediate regression between the series is not advisable and the respective first differences should be taken first as in Chalamandaris & Tsekrekos (2011). This is also justified by the fact that the MSE of the integrated model is lower than the MSE of the typical dynamic VAR(1) model of the literature. The non-dynamic case corresponds to a high MSE of 0.37 for 2000-2009 initial training period and 3.74 for the 2010-2015 initial training period.

4.2.4 Neural network auto-regression (NNAR) for betas

A neural network is briefly described in Figure 4.7. There are a number of interconnected nodes that make up a neural network, known as neurons. The input layer nodes correspond to the number of features you wish to feed into the neural network, while the number of output nodes correspond to the number of items you wish to predict or classify. Each input is an explanatory variable of the output. There is at least one extra introduced layer in neural networks called the hidden layer except from the input and output layer. The number of nodes (or neurons) in the hidden layer is arbitrarily chosen by the analyst usually by trial and error looking for a better fit on their data. The outputs of the nodes in one layer are inputs to the next layer. The inputs to each node are combined using a weighted linear function. The result is then modified by an activation function (usually nonlinear) before producing the output. For our data, experimentation suggests that the linear activation function performs better than the sigmoid. The NNAR is described in the work of Hyndman & Khandakar (2008). The NNAR is a feed-forward neural network where the input layer consists of the lagged values of the time series itself.

The weights used are initially random and modified at each step based on a learning function. The starting randomness of the weights is corrected by repeating the algorithm thousands of times and getting an overall forecast value for reasons of robustness. The learning or error function in this work is the sum of square errors (sse). Neural networks are as black boxes non-interpretable, but they are

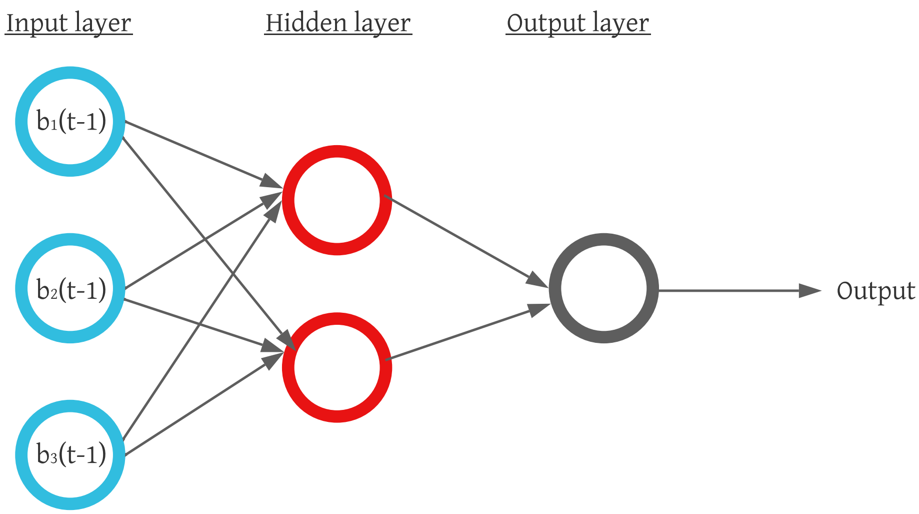

Figure 4.7: The scheme of a typical neural network with three inputs, one hidden layer with two nodes and an output. In our problem, this figure presents the typical NNARX(1) with two nodes.

very commonly used for forecast reasons usually outperforming other methods. The other crucial benefit of neural networks is that they are not based on statistical assumptions. The inputs of the 3 NNARs used here are $\{b_{1t}\}_t, \{b_{2t}\}_t, \{b_{3t}\}_t$ or their lagged values. An inclusion of the lagged values of the other series (NNAR with external regressors) as inputs is also tested as an analogue of the statistical full VAR(1) model. The same exact models are also tested for the increments even though stationarity is not necessary for applying neural networks.

4.2.5 Recurrent neural networks (RNN) for betas

Recurrent neural networks are composed of neurons that connect back to other neurons; information flow is multi-directional, so that neurons' activation can cycle in a loop. This type of neural network has a sense of time and memory of earlier networks states which enables it to learn sequences which vary over time (Lewis 2015). A brief description of those neural networks is provided in Figure 4.8 (Lewis 2015):

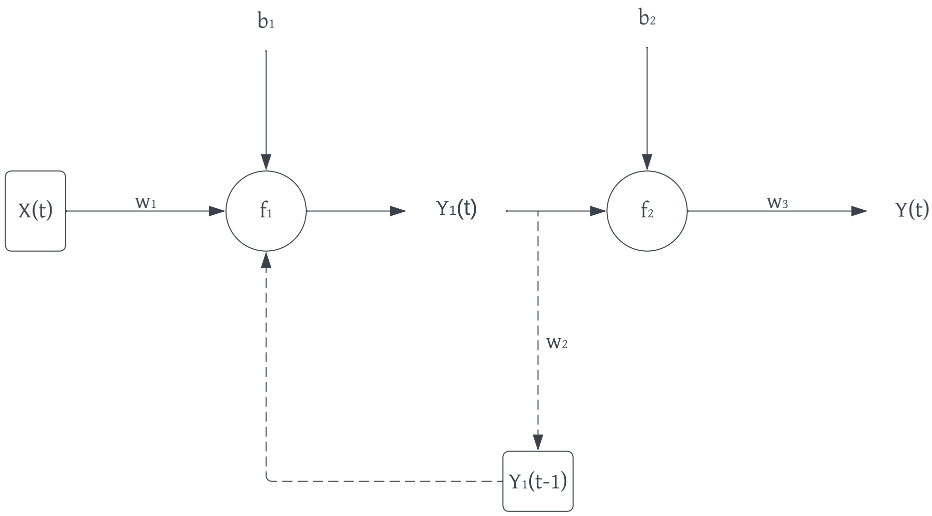

Figure 4.8: The scheme of a typical recurrent neural network as described by Lewis (2015).

In the Elman case, we have the following equations based on Figure 4.8:

$$Y_1(t) = f_1(w_1 X(t) + b_1) \text{ and}$$

$$Y(t) = f_2(w_3 f_1(w_1 X(t) + w_2 Y_1(t-1) + b_1) + b_2).$$

Two simple cases of RNNs include the Elman and Jordan networks. Unlike Jordan networks, Elman networks are configured such that context units receive input not from output units, but from hidden units. Furthermore, there is no direct feedback in the context units. In an Elman net, the number of context units and hidden units needs to be the same. Compared to Jordan nets, Elman nets provide a greater degree of flexibility as the number of context units is not directly determined by the output size (as in Jordan nets). It is determined by the number of hidden units, which is more flexible, as it is easy to add/remove hidden units, but not output units.

Based on one-step ahead cross-validation (RMSE) in Table 4.3, we conclude that the best predictive model is the VAR(1) used on the increments (first differences) considering the first sample period, while the best predictive model is the integrated NNARX(1) considering the second financial period including economic crisis.

Table 4.3: Summary of out of sample MSEs for all used models

Train time period	2004-2009		2010-2015	
Model Differences	0	1	0	1
VAR(1)	0.030	**0.020**	0.027	0.022
NNAR(1)	0.028	2.32	0.044	0.022
NNARX(1)	0.37	0.10	0.029	**0.016**
Elman	0.035	0.024	0.06	0.027
Jordan	0.23	0.025	0.055	0.029

The sufficient fit of the models is justified in Figure 4.9 and Figure 4.10 below, presenting the close distance between the predicted and the actual yield curve for December 2010 and December 2016 respectively.

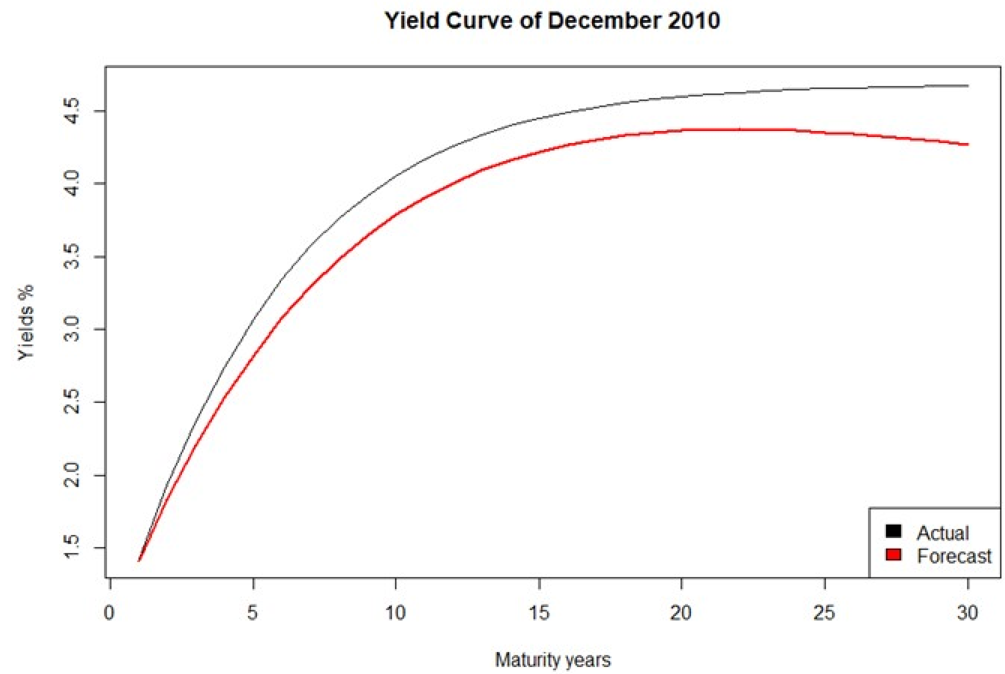

Figure 4.9: Predicted using integrated VAR(1) and actual yield curve of December 2010.

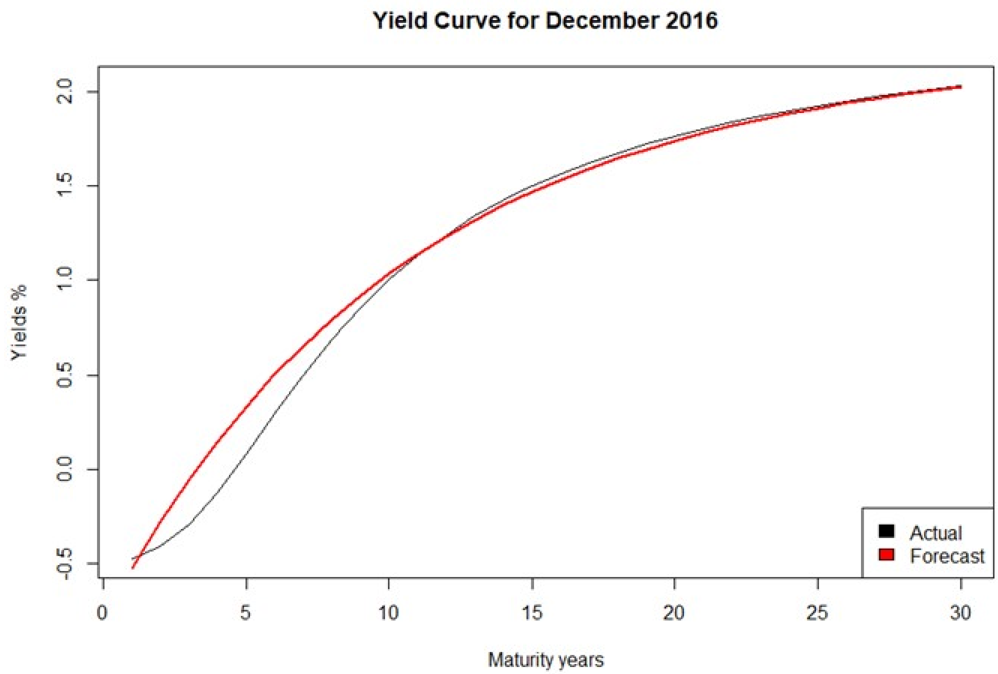

Figure 4.10: Predicted using integrated NNARX(1) and actual yield curve of December 2016.

4.3 Estimating the seismic “default” probabilities

In this section, the statistical method used for the 10-year simulation of the maximum loss by a single event in Greece is analyzed. Only catastrophic events with moment magnitude larger than 6 M_w can cause extreme losses and trigger the CAT bond. Therefore, unlike other studies that use historical catalogues, fault sources are used in this work. The insufficiency of the historical catalogues is eliminated with the use of the geometrical properties of faults in both the reactivation times and the construction losses.

The Simulation Algorithm for the annual damage follows.

1. Fill a (1,000,000x267) matrix **D** with each column representing 1,000,000 different realizations of the number of earthquakes the specific fault produced according to the preceding theory.

2. Fill a (1,000,000x267) matrix **F** with each column representing 1,000,000 different realizations of the magnitude of earthquakes the specific fault is able to produce according to the preceding theory.

3. Create an empty zero (1,000,000x267) matrix $\mathbf{G_1}$ with each column containing the total (over all buildings) losses from each fault and an empty zero (1,000,000x267) matrix $\mathbf{G_2}$ containing the maximum magnitude from each fault.

4. For each fault source, simulate events with replacement from **D** and **F** and keep the maximum annual magnitude for each simulation in $\mathbf{G_2}$.

5. Attenuate each of these magnitudes according to the attenuation equation of to all buildings according to the attenuation equation of Rinaldis et al. (1998), transform the seismic intensity to damage percentage using the fragility curves and store the total losses due to each fault source to $\mathbf{G_1}$.

6. Estimate the proportion of simulations where the total damage exceeds the defined threshold of 3 billion euros due to each fault from $\mathbf{G_1}$.

The same algorithm is repeated for a 2-year time interval, a 3-year time interval and so on until 10 years are reached. Based on the fragility curves of a typical Low-Rise Moderate Code building (construction years 1986-1995) and assuming that each building corresponds to an average cost of 150 thousand euros, the annual loss per building will be 68 euros. In accordance with ELSTAT 2011, there are approximately 4 million residential buildings in Greece resulting in an estimated annual loss of 272 million euros due to all faults combined. In the largest event of Athens 1999, an amount of C = 3 billion euros was estimated to occur. Therefore, C = 3 billion euros will be our threshold for default. If losses are over C, then the remaining coupons and the face value of the bond are not paid to bondholders. Thus, the payments of each year are multiplied with the summation of the 267 probabilities Pr[the fault causes a damage larger than C] in the time interval k,k=1,…,10. These probabilities for a year are depicted in Figure 4.11. According to our seismic model there is a 2% chance the bond defaults due to the potential activation of all faults. 126 euros and 4% are the expected damage by building in two years and the respective probability of exceedance of the threshold C.

It is likely that the payments will occur based on the following probabilities in Table 4.4. The probabilities in this table increase in a linear manner over time. Considering one million simulations, the probability of default for a year is 0.0196 and 0.1946 for a decade. Clearly, this result is attributed to the fact that 1 to 10 years is a relatively short interval in comparison to the reactivation time of the recently activated faults. Consequently, these specific faults are unlikely to be activated and the time-dependent model serves as a barrier to their recurrence. The probabilities are theoretically increasing over time. However, rounding to a precision of two decimal digits determines the above conclusions yielding a linear approximation over time.

There is a need for an equivalent parametric triggered bond so as to be more ro-

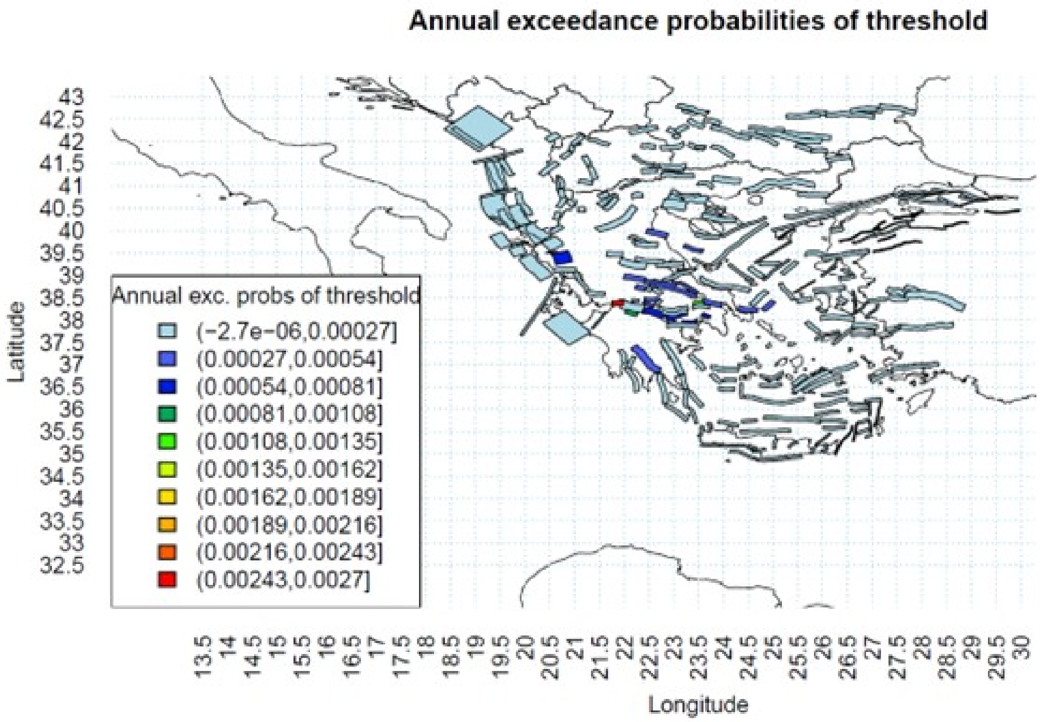

Figure 4.11: Probabilities of exceedance of C = 3 billion euros for each fault

bust against the investors and eliminate moral hazard risk (Cummins et al. 2004). Thus, the modeled-loss index bond can be replaced by an equivalent parametric-triggered bond. The main reason for this change is that the magnitude of an earthquake is announced by an independent organization, usually a geophysical institute, which has no financial interest in announcing a stronger or weaker earthquake. It is also important to note that the model used to estimate losses is a closed model, unknown to the investors while the magnitude threshold offers transparency to all parties involved. Thus, we split the region in two zones with different thresholds as presented in Figure 4.12.

In this figure, it is worth observing that in the region of Attica and north Peloponnese, there are even weaker earthquakes that may exceed the threshold of 3 billion euros as was in the case of Athens, 1999. By comparing the two seis-

Table 4.4: Probabilities of occurrence of payments of the CAT bond

Years	Probabilities
1st year	0.98
2nd year	0.96
3rd year	0.94
4th year	0.92
5th year	0.90
6th year	0.88
7th year	0.86
8th year	0.84
9th year	0.82
10th year	0.80

mic risk maps with the probability of loss threshold exceedance and the minimum produced magnitude on an annual basis, we can determine that the bond should be triggered in the manner outlined below. All remaining payments to investors should be canceled if an earthquake of at least 6.4 M_w happens located in Attica or north Peloponnese or an earthquake of at least 6.7 M_w happens located in Macedonia or south Peloponnese or Ionian islands or Crete. Therefore, on an annual basis, P(at least one earthquake in zone1 > 6.4) + P(at least one earthquake in zone 2 > 6.7) = 2%, as estimated by the matrix G_2 from the proposed algorithm.

The CAT bond pricing based on these probabilities is presented in the following tables considering a variety of discounting methods starting from the time independent models.

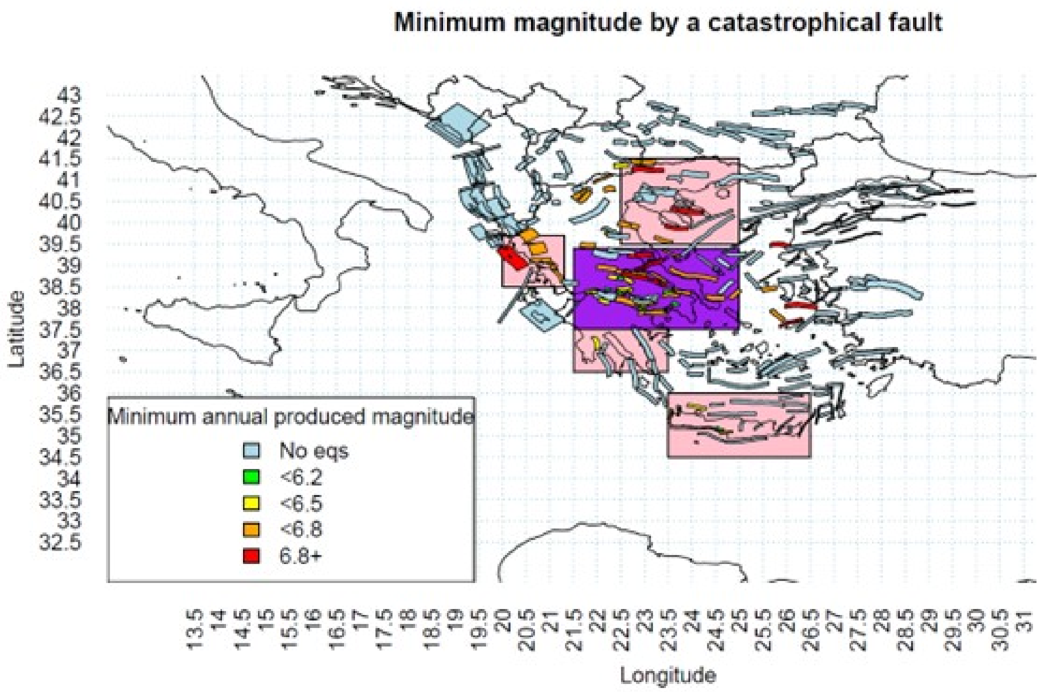

Figure 4.12: Zone 1 (purple region with threshold 6.4 M_w) and zone 2 (pink regions with threshold 6.7 M_w).

4.3.1 Time independent models

1. 3-month IBR of Germany

   (a) 2000-2009 ($\frac{\text{Expected Coupon}}{\text{Face Value}} = 3.23\% + spread$)

Table 4.5: CAT bond pricing using the 3-month IBR of Germany for the time period 2000-2009

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	97.89	97.45	97.19	97.11	97.20	97.43	97.80	98.28
2%	100.61	100.99	101.52	102.20	103.01	103.95	104.99	106.12
3%	103.33	104.54	105.86	107.29	108.83	110.46	112.18	113.96
4%	106.05	108.09	110.19	112.38	114.65	116.98	119.37	121.80
5%	108.77	111.63	114.53	117.47	120.46	123.49	126.55	129.64

   (b) 2010-2015 ($\frac{\text{Expected Coupon}}{\text{Face Value}} = 0.85\% + spread$)

Table 4.6: CAT bond pricing using the 3-month IBR of Germany for the time period 2010-2015

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	101.48	102.34	103.34	104.48	105.74	107.10	108.55	110.08
2%	104.40	106.21	108.15	110.23	112.42	114.71	117.09	119.55
3%	107.32	110.08	112.97	115.98	119.11	122.33	125.64	129.02
4%	110.24	113.94	117.78	121.73	125.79	129.94	134.18	138.49
5%	113.15	117.81	122.59	127.49	132.48	137.56	142.72	147.96

We observe that prices are lower during the more promising financial period 2000-2009 corresponding to Table 4.5 than during 2010-2015 corresponding to Table 4.6 despite the fact that the first bond also provides greater payments as based on a higher total coupon rate. This observation does not hold for the case of the ECB yield curves as presented below.

2. ECB Yield Curves

   (a) 2004-2009 ($\frac{\text{Expected Coupon}}{\text{Face Value}} = 3.23\% + spread$)

Table 4.7: CAT bond pricing using the ECB yield curves for the time period 2004-2009

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	96.94	95.39	93.71	91.95	90.16	88.37	86.60	84.88
2%	99.66	98.91	97.99	96.94	95.81	94.63	93.45	92.26
3%	102.37	102.43	102.26	101.92	101.45	100.90	100.29	99.64
4%	105.08	105.95	106.53	106.90	107.10	107.16	107.13	107.02
5%	107.79	109.46	110.81	111.88	112.74	113.43	113.97	114.39

   (b) 2010-2015 ($\frac{\text{Expected Coupon}}{\text{Face Value}} = 0.85\% + spread$)

Table 4.8: CAT bond pricing using the ECB yield curves for the time period 2010-2015

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	95.23	92.60	89.70	86.62	83.46	80.28	77.15	74.10
2%	98.05	96.27	94.18	91.87	89.43	86.94	84.44	81.99
3%	100.86	99.95	98.66	97.12	95.41	93.60	91.74	89.87
4%	103.67	103.62	103.15	102.38	101.39	100.25	99.03	97.76
5%	106.48	107.29	107.63	107.63	107.37	106.91	106.32	105.65

In general, bonds derived by the ECB yield curves offer lower prices as the maturity time increases. Moreover, prices are lower during the financial period that includes the economic crisis in this case. The forecasts derived by the best performing time-dependent models are summarized in the tables below.

4.3.2 Time dependent models

1. using the forecasted “December 2010” yields by the integrated VAR(1) model ($\frac{\text{Expected Coupon}}{\text{Face Value}} = 3.23\% + spread$)

Table 4.9: CAT bond pricing using the yields for December 2010 as forecasted by the integrated VAR(1) model for betas

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	99.71	98.38	96.71	94.83	92.84	90.81	88.80	86.84
2%	102.48	101.98	101.09	99.95	98.64	97.25	95.82	94.41
3%	105.25	105.58	105.48	105.06	104.44	103.68	102.85	101.99
4%	108.02	109.19	109.86	110.18	110.24	110.12	109.88	109.56
5%	110.79	112.79	114.25	115.30	116.04	116.55	116.90	117.13

2. using the forecasted “December 2016” yields by the integrated NNARX(1) model ($\frac{\text{Expected Coupon}}{\text{Face Value}} = 0.85\% + spread$)

Table 4.10: CAT bond pricing using the yields for December 2016 as forecasted by the integrated NNARX(1) model for betas

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	99.50	98.49	97.19	95.65	93.93	92.08	90.13	88.12
2%	102.39	102.30	101.88	101.19	100.30	99.23	98.04	96.75
3%	105.29	106.11	106.57	106.74	106.66	106.39	105.95	105.38
4%	108.18	109.91	111.26	112.29	113.03	113.54	113.86	114.01
5%	111.07	113.72	115.95	117.83	119.40	120.70	121.77	122.65

Comparing Tables 4.9 and 4.10, it is obvious that prices are fairly close for the two different time-period structured CAT bonds. However, the first one is more profitable due to the increased total coupon rate. Furthermore, it is worth mentioning that the fixed spread of 4% makes the CAT bond price roughly constant in the first case and equal to 110 euros, while the fixed spread of 3% makes the CAT bond price roughly equal to 105 euros in the second case. The time-dependent models suggest higher prices than the respective time-independent models, because they adapt to the current state of the economy each time.

4.4 Credit risk

In the present section, the case where the insurers are the bond issuers themselves is analysed. Then, SPV does not exist, and credit risk is introduced. Suppose there are three bond issuers (states of an issuer) concerned as “Good”, ”Bad” and “Bankrupt” respectively. The price of each issuer’s CAT bond varies based on the credibility power of the issuer. The dynamic behaviour of the different states is modeled by a Markov Chain (see Privault, 2013 for a gentle introduction to Markov Chains). According to the below transition probability matrix, the likelihood of a transition from each state to another state next year independently of the seismic events is:

$$P = \begin{pmatrix} 0.95 & 0.045 & 0.005 \\ 0.25 & 0.50 & 0.25 \\ 0 & 0 & 1 \end{pmatrix}$$

According to P, each good issuer remains to the same state in a year with probability 95%, while moves to the state of a bad issuer with probability 4.5% and bankrupts with probability 0.5%. Since the CAT bond is assumed to have a 3 to 10-year maturity, it is necessary to estimate all state probabilities for the next 10 years. The matrices are arranged in the following order:

$$P^2 = \begin{pmatrix} 0.914 & 0.065 & 0.021 \\ 0.363 & 0.261 & 0.376 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\vdots$$

$$P^{10} = \begin{pmatrix} 0.73 & 0.07 & 0.2 \\ 0.385 & 0.037 & 0.578 \\ 0 & 0 & 1 \end{pmatrix}$$

For a “good” bond issuer (state 1) every k year’s payment is multiplied with $p_{11}^k + p_{12}^k = 1 - p_{13}^k$, i.e., the probability that this issuer remains solvent (remains to the same state 1 or moves to state 2). As a result of the credit risk of the issuers, the bond’s price is reduced even further, increasing its attractiveness and offering a higher yield. The pricing formula is provided in the equation below:

$$PV(t,T,j) = \sum_{h=1}^{T} e^{-hy_t(h)}(1-p_h)P_{j,h}(b_v + c)FV + FVe^{-Ty_t(T)}(1-p_T)P_{j,T},$$

where FV: Face Value= 100 euros and $P_{j,h}$ the probability of issuer j not bankrupting after h years based on matrix P. Specifically,

$$P_{j,h} = 1 - P^h(j,3).$$

The probability vector that the 1st issuer will be able to fulfil its obligations each of the ten years of the bond’s life is presented in Table 4.11, while the probability vector that the 2nd issuer will be able to fulfil its obligations each of the ten years of the bond’s life is presented in Table 4.12.

Table 4.11: Probabilities of occurrence of bond payments with respect to the credit risk of issuer 1

Years	Probabilities
1st year	0.995
2nd year	0.979
3rd year	0.958
4th year	0.935
5th year	0.912
6th year	0.888
7th year	0.865
8th year	0.843
9th year	0.821
10th year	0.799

Table 4.12: Probabilities of occurrence of bond payments with respect to the credit risk of issuer 2

Years	Probabilities
1st year	0.750
2nd year	0.624
3rd year	0.557
4th year	0.518
5th year	0.493
6th year	0.474
7th year	0.459
8th year	0.446
9th year	0.434
10th year	0.422

The prices of the CAT bonds issued by each issuer of different credibility are summarized below incorporating their creditworthiness.

4.4.1 Time independent cases

1. For the 3-month IBR of Germany

   (a) 2000-2009

      i. Issuer 1

Table 4.13: CAT bond pricing using the 3-month IBR of Germany for issuer 1 incorporating credit risk for the time-period 2000-2009

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	94.02	91.63	89.46	87.54	85.87	84.43	83.19	82.12
2%	96.68	95.06	93.61	92.37	91.33	90.47	89.78	89.24
3%	99.34	98.50	97.77	97.19	96.78	96.51	96.38	96.35
4%	102.00	101.93	101.92	102.01	102.23	102.55	102.97	103.47
5%	104.66	105.36	106.07	106.84	107.68	108.59	109.57	110.58

      ii. Issuer 2

Table 4.14: CAT bond pricing using the 3-month IBR of Germany for issuer 2 incorporating credit risk for the time-period 2000-2009

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	55.52	51.95	49.75	48.26	47.16	46.30	45.60	45.01
2%	57.28	54.13	52.32	51.19	50.43	49.88	49.47	49.16
3%	59.04	56.32	54.90	54.13	53.69	53.46	53.34	53.31
4%	60.80	58.51	57.47	57.06	56.96	57.04	57.21	57.45
5%	62.56	60.69	60.05	59.99	60.23	60.61	61.08	61.60

CAT bond prices by both issuers are decreasing by maturity time when spreads are low (1%-3%). The above does not hold when spreads are larger (4% and 5%). Prices also increase by the increase of the spread. This increase is almost the same for both issuers. For example, there is an increase of 35% when we get from spread=1% to spread=5% for T=10 considering the 1st issuer, while the same increase is 37% for the 2nd issuer.

(b) 2010-2015

i. Issuer 1

Table 4.15: CAT bond pricing using the 3-month IBR of Germany for issuer 1 incorporating credit risk for the time-period 2010-2015

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	96.58	95.66	94.35	93.48	92.68	91.84	91.70	91.17
2%	99.17	98.86	98.77	98.40	98.37	98.32	98.93	99.26
3%	101.59	101.88	102.57	103.37	104.16	105.04	106.00	107.11
4%	104.93	106.64	108.02	109.28	111.15	113.19	115.04	116.93
5%	107.35	109.56	111.96	114.74	117.48	119.78	122.10	124.99

ii. Issuer 2

Table 4.16: CAT bond pricing using the 3-month IBR of Germany for issuer 2 incorporating credit risk for the time-period 2010-2015

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	57.07	54.15	52.23	51.18	50.31	50.04	49.54	49.32
2%	58.74	55.95	54.49	53.73	53.33	53.38	53.72	53.70
3%	60.62	58.23	57.35	57.11	57.30	57.64	58.11	58.70
4%	62.23	60.50	60.07	60.32	60.84	61.62	62.43	63.25
5%	64.96	63.65	63.69	64.54	65.63	66.80	68.02	69.21

2. ECB yield curves modeled with NSTI

   (a) 2004-2009

      i. Issuer 1

Table 4.17: CAT bond pricing using the ECB yield curves for issuer 1 incorporating credit risk for the time-period 2004-2009

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	93.11	89.71	86.29	82.95	79.77	76.76	73.95	71.33
2%	95.76	93.12	90.38	87.68	85.07	82.58	80.24	78.05
3%	98.42	96.52	94.48	92.40	90.36	88.40	86.53	84.77
4%	101.07	99.93	98.57	97.13	95.66	94.22	92.82	91.50
5%	103.72	103.33	102.67	101.85	100.96	100.04	99.12	98.22

      ii. Issuer 2

Table 4.18: CAT bond pricing using the ECB yield curves for issuer 2 incorporating credit risk for the time-period 2004-2009

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	55.00	50.88	48.03	45.80	43.91	42.23	40.71	39.31
2%	56.75	53.05	50.57	48.68	47.10	45.69	44.42	43.24
3%	58.51	55.22	53.12	51.56	50.28	49.15	48.13	47.18
4%	60.26	57.40	55.66	54.44	53.47	52.61	51.84	51.12
5%	62.02	59.57	58.21	57.32	56.65	56.08	55.55	55.06

CAT bond prices by both issuers are decreasing by maturity time whatever the spreads are. Prices also increase by the increase of the spread. This increase is almost the same for both issuers. For example, there is an increase of 38% when we get from spread=1% to spread=5% for T=10 (Table 4.17) considering the 1st issuer, while the same increase for the 2nd issuer is 37.5% (Table 4.18).

(b) 2010-2015

i. Issuer 1

Table 4.19: CAT bond pricing using the ECB yield curves for issuer 1 incorporating credit risk for the time-period 2010-2015

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	90.27	85.95	81.50	77.24	72.96	68.53	64.54	60.89
2%	93.35	89.66	86.15	82.35	78.24	75.02	71.38	68.33
3%	96.12	93.44	90.19	87.03	84.00	81.25	78.21	75.64
4%	98.85	96.40	93.84	91.42	88.99	87.03	84.83	82.69
5%	101.45	99.97	98.15	96.43	94.33	92.37	90.53	89.07

ii. Issuer 2

Table 4.20: CAT bond pricing using the ECB yield curves for issuer 2 incorporating credit risk for the time-period 2010-2015

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	53.64	48.85	45.31	42.54	40.17	37.84	35.75	33.76
2%	55.54	51.25	48.24	45.52	43.10	41.24	39.37	37.63
3%	58.29	54.28	51.28	49.20	47.24	45.73	43.94	42.62
4%	59.55	56.05	53.67	51.83	50.31	49.01	47.64	46.62
5%	60.58	57.81	56.25	54.97	53.70	52.56	51.53	50.64

It is worth mentioning that ECB yield curves suggest lower prices than the respective IBR of Germany. The CAT bond priced by the 2nd issuer results to lower prices, because the 1st one is more credible. Moreover, both issuers provide a lower CAT bond price than the SPV as a result of the introduction of credit risk. We continue by presenting the results by the time dependent models.

4.4.2 Time dependent cases

In this section, the results of the CAT bond pricing incorporating credit risk considering the time dependent forecasts of the discounting yield curves are presented:

1. using the forecasted "December 2010" yields by the integrated VAR(1) model

   (a) Issuer 1

Table 4.21: CAT bond pricing for issuer 1 using the yields for December, 2010 as forecasted by the integrated VAR(1) model for betas incorporating credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	95.76	92.51	89.04	85.54	82.13	78.87	75.82	72.98
2%	98.48	96.00	93.24	90.39	87.57	84.85	82.28	79.88
3%	101.19	99.48	97.44	95.24	93.01	90.83	88.74	86.78
4%	103.90	102.97	101.64	100.09	98.45	96.81	95.21	93.68
5%	106.61	106.46	105.84	104.94	103.90	102.79	101.67	100.58

   (b) Issuer 2

Table 4.22: CAT bond pricing for issuer 2 using the yields for December, 2010 as forecasted by the integrated VAR(1) model for betas incorporating credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	56.55	52.45	49.54	47.21	45.20	43.38	41.73	40.21
2%	58.35	54.67	52.15	50.17	48.46	46.93	45.53	44.25
3%	60.14	56.90	54.76	53.12	51.73	50.49	49.34	48.29
4%	61.93	59.12	57.36	56.08	55.00	54.04	53.15	52.33
5%	63.72	61.34	59.97	59.03	58.27	57.59	56.96	56.37

2. using the forecasted “December 2016” yields by the integrated NNARX(1) model

   (a) Issuer 1

Table 4.23: CAT bond pricing for issuer 1 using the yields for December, 2016 as forecasted by the integrated NNARX(1) model for betas incorporating credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	94.51	91.43	88.43	85.13	81.95	78.61	75.59	72.28
2%	97.31	95.30	92.99	90.43	87.84	85.36	83.06	80.46
3%	99.98	98.72	97.35	95.75	93.61	92.06	90.36	88.75
4%	103.74	103.09	102.36	101.31	100.14	99.34	98.24	96.85
5%	106.29	107.01	107.06	106.90	106.48	106.04	105.29	104.46

   (b) Issuer 2

Table 4.24: CAT bond pricing for issuer 2 using the yields for December, 2016 as forecasted by the integrated NNARX(1) model for betas incorporating credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	56.15	52.01	49.06	46.91	45.02	43.07	41.37	39.82
2%	57.45	53.76	51.41	49.70	48.10	46.51	45.04	43.76
3%	60.09	56.77	54.62	53.30	52.09	51.19	50.19	49.27
4%	62.54	59.90	58.27	57.24	56.54	55.82	55.01	54.10
5%	63.51	61.54	60.48	59.69	59.36	59.04	58.82	58.30

Time dependent prices are increased compared to their average time independent counterparts, as the state of the economy rapidly worsens since 2010. Moreover, we infer the following. The 3-month Interbank Rate of Germany results to a more pessimistic pricing by the view of the investor considering the worse financial period. Specifically, the prices are increased compared to the respective prices of the more promising financial period despite the fact that the total coupon rate is much lower in the 2nd time-period.

4.5 Dynamic credit risk with respect to events

We also introduce an interdependence between the credit risk of the issuer and the seismic sequence. A negative shock is experienced in the solvency capacity of the issuer in case of a significant event since the obligation to pay claims comes into conflict with the obligation to return the payments to the investors. The transition probability matrix is assumed to change as follows given that a strong and larger than $6M_w$ but weaker than 6.4 M_w in zone 1 and 6.7 M_w in zone 2 occurs during the lifetime of the bond in the insured zones:

$$P = \begin{pmatrix} 0.95-\epsilon & 0.045+2\frac{\epsilon}{3} & 0.005+\frac{\epsilon}{3} \\ 0.25-2\frac{\epsilon}{3} & 0.50-\frac{\epsilon}{3} & 0.25+\epsilon \\ 0 & 0 & 1 \end{pmatrix}$$

For $\epsilon = 9\%$ and after many lines of coding and simulations, the CAT bond price under a Monte-Carlo approach for each different discounting model is presented in the tables below. The less complicated time independent models are provided in the Appendix for this case (Tables A3-A10), while we present the time dependent models in this section.

4.5.1 Time dependent cases

In this section, the results of the CAT bond pricing incorporating dynamic credit risk considering the time dependent forecasts of the discounting yield curves are presented:

1. using the forecasted “December 2010” yields by the integrated VAR(1) model

   (a) Issuer 1

Table 4.25: CAT bond pricing for issuer 1 using the yields for December, 2010 as forecasted by the integrated VAR(1) model for betas incorporating dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	94.07	90.80	87.22	83.50	79.88	76.17	72.93	70.05
2%	96.85	94.11	90.71	87.77	84.62	81.84	79.26	76.49
3%	99.56	97.23	94.91	92.62	90.11	87.44	84.82	82.67
4%	101.58	100.53	98.75	96.57	94.65	93.10	91.20	89.21
5%	104.52	104.05	102.96	101.59	100.23	98.77	97.25	95.77

   (b) Issuer 2

Table 4.26: CAT bond pricing for issuer 2 using the yields for December, 2010 as forecasted by the integrated VAR(1) model for betas incorporating dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	56.24	52.05	48.94	46.31	44.04	42.13	40.29	38.69
2%	56.73	53.30	50.83	48.83	46.97	45.23	43.83	42.45
3%	59.73	56.14	53.80	51.97	50.26	48.81	47.47	46.15
4%	61.93	58.39	56.36	54.93	53.62	52.60	51.41	50.51
5%	62.97	60.62	59.00	57.96	56.95	56.20	55.35	54.68

2. using the forecasted “December 2016” yields by the integrated NNARX(1) model

   (a) Issuer 1

Table 4.27: CAT bond pricing for issuer 1 using the yields for December, 2016 as forecasted by the integrated NNARX(1) model for betas incorporating dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	93.60	90.33	86.64	82.97	79.09	75.53	71.82	68.33
2%	96.43	93.67	91.10	87.59	84.61	81.67	78.58	75.34
3%	99.25	97.06	94.69	92.61	90.27	87.74	85.19	82.94
4%	102.43	101.49	99.95	98.44	96.43	94.43	92.94	91.15
5%	105.49	104.80	103.81	102.99	102.20	100.73	99.38	98.04

   (b) Issuer 2

Table 4.28: CAT bond pricing for issuer 2 using the yields for December, 2016 as forecasted by the integrated NNARX(1) model for betas incorporating dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	55.35	50.92	47.91	45.47	42.97	40.81	38.78	37.19
2%	56.43	52.45	49.86	47.61	45.69	43.96	42.22	40.58
3%	59.64	56.33	54.08	52.30	50.57	48.80	47.33	46.05
4%	60.07	57.02	55.02	53.50	52.36	51.44	50.53	49.42
5%	63.77	60.92	59.81	58.76	57.76	56.79	56.04	55.36

As expected, the dynamic with respect to credit risk CAT bond pricing yields even lower prices due to the shape of the transition probability matrix.

4.6 CAT bond design

In this section, we describe the process of determining the ideal spread over the coupon of a conventional bond. Data containing the prices of the standard government bond and its coupons as quoted by the Bank of Greece are used as a benchmark for our design and calculations. For example, on November 17, 2022, the price announced was 80.41 euros with a coupon of 4.29%. The effective interest rate of this bond is estimated to be 7% as seen in Figure 4.13. Incorporating the seismic hazard default risk, the proposed spread is the one that yields the same price with the government bond. In this case, we assume that only the coupons are at risk, while the capital is assured to be paid to the investor. After discounting the CAT bond by the same discount rate and using the 2% annual probability of default due to seismic risk (according to Table 4.4), the proposed spread is found to be equal to 1.99% as presented in Figure 4.14. It is thus concluded that a total coupon of 4.29%+1.99%=6.28% is the ideal for the CAT bond. This same spread for January 2010 is estimated as 2.41%. In the following chapter, the results of the thesis are summarized, as well as recommendations for future research in this topic.

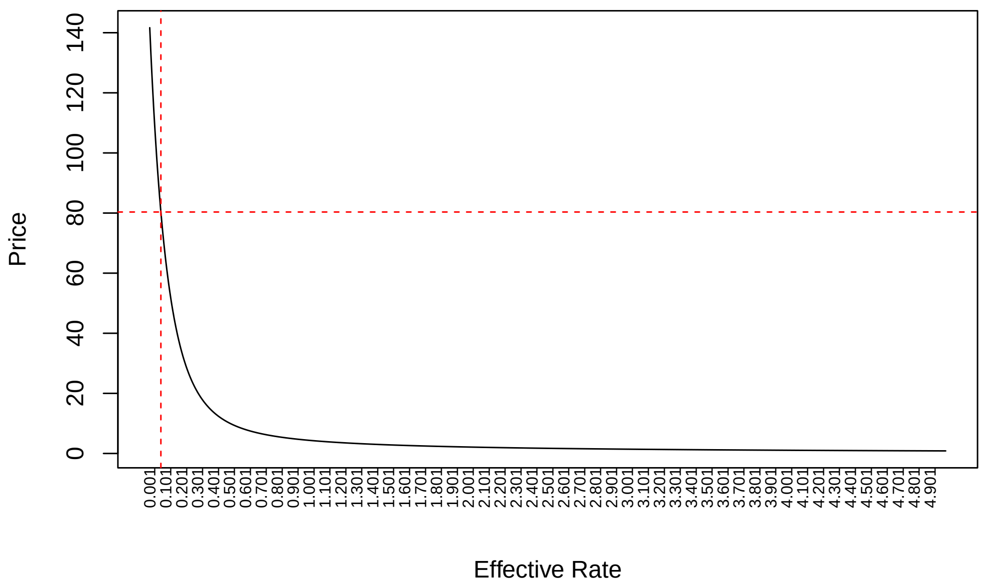

Figure 4.13: Effective interest rate estimation of the government bond

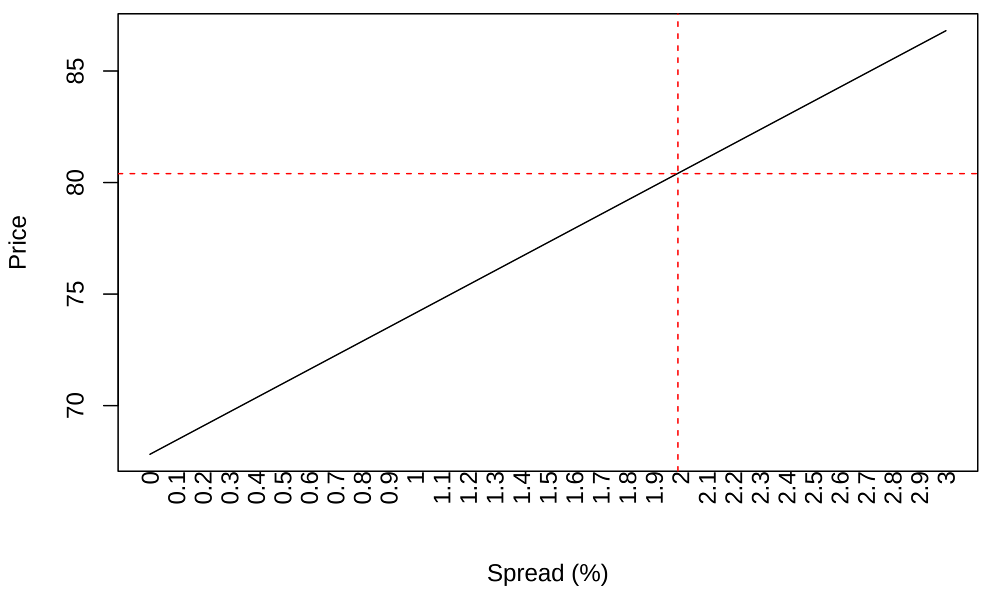

Figure 4.14: Spread determination for the CAT bond issuance

Chapter 5

Conclusion and further research

The main objective of this thesis is to present an earthquake model that can be used not only for insurance pricing and reserving, but also for SCR calculations in the context of Solvency II and the design and pricing of a CAT bond. The algorithms used in the present study may produce premium rates with coordinate precision in contrast to other studies that refer to regional pricing. After comparing different models, encoded as A0 to AF11, and Ae or AFe, we present the results that deserve special attention. The model that uses the weighting function, AF01, offers a slight increase of about 6% in the premium rate over the simpler model AF00 as presented in Table 3.6. The kernel smoother models lower the earthquake occurrence rate of the high-intensity tiles, but increase the earthquake occurrence rate of the low-intensity tiles resulting to the increase of the premium rates and SCR (AF10 vs AF00 or AF11 vs AF01). This high-resolution pricing could be beneficial to insurance companies in terms of adverse selection.

The incorporation of faults contributes to premium rating by a significant percentage in the region of Greece, equal to roughly 10% when the typical models are used (A0 vs AF00), but also when the ETAS model is used (Ae vs AFe). This contribution becomes slightly greater for Attica as it is a region containing plenty of faults as presented in Figure 2.5. More specifically, it offers an addition of

12% to 13% when the typical Voronoi model or the ETAS model are used for the historical catalogues. These contributions are significant, as faults complete the data gaps that exist in the historical catalogues. The ETAS model should also be considered by insurance institutions and analysts since earthquakes exhibit an epidemic behaviour in practice. The ETAS model yields higher premium estimates than the simpler model, which is based on GK declustering, by an amount of 39% when the expected value is used. The respective contribution when the median value is used is estimated as 27%. The actual rates of the reinsurance market are about 100 euros annually per 200,000 euros of insured value. These actual market rates coincide with those of the ETAS model in median terms as the median rate of the Ae model is found to be 107 euros, opposed to the GK declustering model, where they coincide in average terms as the average rate of the A0 model is found to be 101 euros.

Contrary to this work, the standard formula of Solvency II for the estimation of the SCR does not consider the construction year of each building leading insurance companies to over-reserving. The results of the proposed models prove that the SCR based on VaR should be lower than the one calculated using the standard formula considering moderate and high code buildings which is the vast majority of the insured cases. For example, the proposed SCR by the AF10 model is only 67% of the standard formula's considering moderate code buildings (Table 3.10). If we do not consider faults (A0 model), then the proposed SCR is only 48% of the standard formula's (Table 3.10), which also considers historical catalogues. The SCR induced by the ETAS model ranges from -29% up to 22% compared to the standard formula's SCR depending on the construction year (Table 3.12) versus the A0 to AF11 models where it ranges from -52% up to 22% of the standard formula's SCR (Table 3.13). Convex risk measures propose larger SCR values, but they should be strongly considered by insurance institutions as they deal with model uncertainty thanks to their robust representations. For example, the SCR induced by the AF00 model and the Expected Shortfall coincides with the one by the SF (Table 3.12).

As the magnitude of equity capital in insurance and reinsurance companies is potentially small compared to the large-scale claims caused by extreme catastrophic events making them unable to compensate them, the industry employs catastrophe (CAT) bonds to transfer this risk to investors in capital markets. In this thesis, CAT bond pricing is processed to deal with events such that of 1999 in Athens, which resulted to $c = 3$ billion euros of total material loss[1] by the activation of the Parnitha fault. Therefore, the bond is used as a reinsurance alternative. The proposed fault-specific model estimates the probability of overcoming this threshold c equal to 2% on an annual basis based on the geometrical properties of faults as described by the GreDaSS database (Table 4.4). A CAT bond design that could be potentially issued by the Greek government is provided. The proposed spread over a typical coupon of the Greek government bond is estimated as 2% to undertake the risk of the natural catastrophe of earthquake.

Various discounting methods are considered for the estimation of the present value of the bond. The 3-month IBR of Germany (modelled by the model of Vasicek) yields more pessimistic estimates for the investor than the ECB yield curves and with a wider price range between the two different investigated time-periods. For example, the price of the 10-year zero-coupon bond discounted with ECB yield curves (Face Value=1) ranges from 0.67 to 0.74 for the two different time periods (Table 4.2). Respectively, the price of the 10-year zero-coupon bond (Face Value=1) discounted with 3-month IBR of Germany ranges from 0.81 to 1.16 for the different time periods (Table 4.1). The Svensson model appears to be statistically insignificant proving that the simpler Nelson-Siegel model is sufficient for modelling the ECB yield curves. In order to deal with the non-stationarity of the beta factors of the Nelson-Siegel model, the differenced time-series is used. Comparing the different models based on cross-validation, the best predictive model for the first (more promising) financial period is the VAR(1) used on the incre-

[1] https://en.wikipedia.org/wiki/1999_Athens_earthquake

ments of betas with an MSE of 0.020. The best predictive model for the second (including the financial crisis) financial period respectively is the NNARX(1) used on the increments of betas with an MSE of 0.016 (Table 4.3).

CAT bond pricing by the Vasicek model on the 3-month IBR of Germany yields higher CAT bond prices during the second financial period compared to the first despite the lower coupon payments. On the other hand, the exactly opposite happens with ECB yield curves fitted by the Nelson-Siegel model. We also observed that there can be found different spread fixed values that keep the CAT bond price roughly constant across different maturities. Moreover, the prices with respect to time-dependent methods offer the benefit of providing forecasts and they are increased compared to their time independent counterparts as the financial crisis appeared since 2010.

In the absence of the mediator SPV, credit risk is introduced as the insurer is also the issuer. An insurance company must be able to simultaneously receive premiums for its policies and pay claims, receive capital and pay returns for the issuance of the CAT bond. Dynamic and non-dynamic credit risk is investigated as two issuers are introduced where the first appears to be more solvent than the second based on a transition probability matrix. As expected, the CAT bond pricing involving credit risk yields lower prices than the one issued by an SPV, and the bond issued by the first issuer appears to be more expensive than the second. Finally, the prices are more reduced in the dynamic case where there is a dependence between the seismic sequence and the creditworthiness of the issuers.

The present research could be extended in many aspects. Further research could include more inter-event time distributions and more GMPEs. Different premium principles may also be used to account for uncertainty. The claims in this work are also supposed to be equal to the estimated losses. This could be evolved by considering deductions and inflation costs to premium rating. A CAT bond could be also issued for different hazard phenomena considering the proposed discounting processes (wildfire, hurricane, flood etc.). However, all proposed methods are again applicable after such changes.

References

1. Aki, K. (1966). Generation and propagation of g waves from the niigata earthquake of june 16, 1964.: Part 2. estimation of earthquake moment, released energy, and stress-strain drop from the g wave spectrum. Bulletin of the Earthquake Research Institute, University of Tokyo, 44(1):73–88.

2. Artzner P, Delbaen F, Eber J-M, Heath D (1999) Coherent measures of risk. Mathematical Finance 9(3):203–228

3. Asprone D, Jalayer F, Simonelli S, Acconcia A, Prota A, Manfredi G (2013) Seismic insurance model for the Italian residential building stock. Structural Safety 44:70–79

4. Box GE, Pierce DA (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association 65(332):1509–1526

5. Calvi GM, Pinho R, Magenes G, Bommer JJ, Restrepo-Vélez LF, Crowley H (2006) Development of seismic vulnerability assessment methodologies over the past 30 years. ISET Journal of Earthquake Technology 43(3):75–104

6. Chalamandaris, G., & Tsekrekos, A. E. (2011). How important is the term structure in implied volatility surface modeling? Evidence from foreign ex-

change options. Journal of International Money and Finance, 30(4), 623-640.

7. Cifuentes, A., & Charlin, V. (2016). Operational risk and the Solvency II capital aggregation formula: implications of the hidden correlation assumptions. Journal of Operational Risk, Forthcoming.

8. Console R, Murru M, Falcone G (2017) Earthquake occurrence: short- and long-term models and their validation. John Wiley & Sons

9. Cowpertwait, P. S., & Metcalfe, A. V. (2009). Introductory time series with R. Springer Science & Business Media.

10. Cua G, Wald D, Allen T, Garcia D, Worden C, Gerstenberger M, Lin K, Marano K (2010) Best practices for using macroseismic intensity and ground motion-intensity conversion equations for hazard and loss models in gem1. Technical report, GEM Technical Report 2010–4, GEM Foundation, Pavia, Italy

11. Cummins, J. D., Lalonde, D., & Phillips, R. D. (2004). The basis risk of catastrophic-loss index securities. Journal of Financial Economics, 71(1), 77-111.

12. Das, S. (2019). Modeling Nelson–Siegel Yield Curve Using Bayesian Approach. New Perspectives and Challenges in Econophysics and Sociophysics, 169.

13. De La Grandville, O. (2003). Bond pricing and portfolio analysis: Protecting investors in the long run. MIT Press.

14. Deligiannakis G, Papanikolaou I, Roberts G (2018) Fault specific gis based seismic hazard maps for the Attica region, Greece. Geomorphology 306:264–282

15. Deligiannakis G, Zimbidis A, Papanikolaou I (2021) Earthquake loss and solvency capital requirement calculation using a fault-specific catastrophe model. The Geneva Papers on Risk and Insurance-Issues and Practice 1–26

16. Dong W, Shah H, Wong F (1996) A rational approach to pricing of catastrophe insurance. Journal of Risk and Uncertainty 12(2):201–218

17. Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. Journal of econometrics, 130(2), 337-364.

18. Douglas J (2014) Ground motion prediction equations 1964–2014. PEER Report 2011:102

19. Duffie, D., & Singleton, K. J. (1999). Modeling term structures of defaultable bonds. The review of financial studies, 12(4), 687-720.

20. Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. Journal of statistical software, 40, 1-18.

21. Eddelbuettel, D. (2013). Seamless R and C++ integration with Rcpp. New York: Springer.

22. Elnashai, A. S., & Di Sarno, L. (2008). Fundamentals of earthquake engineering (p. 347). New York: Wiley.

23. Episcopos, A. (2000). Further evidence on alternative continuous time models of the short-term interest rate. Journal of International Financial Markets, Institutions and Money, 10(2), 199-212.

24. Escobar DD, Pflug GC (2018) The distortion principle for insurance pricing: properties, identification and robustness. Annals of Operations Research 1–24

25. Fergusson, K., & Platen, E. (2015). Application of maximum likelihood estimation to stochastic short rate models. Annals of Financial Economics, 10(02), 1550009.

26. Föllmer H, Schied A (2011) Stochastic finance: an introduction in discrete time. Walter de Gruyter

27. Fuller, W. A. (2009). Introduction to statistical time series (Vol. 428). John Wiley & Sons.

28. Galeotti, M., Gürtler, M., & Winkelvos, C. (2013). Accuracy of premium calculation models for CAT bonds—an empirical analysis. Journal of Risk and Insurance, 80(2), 401-421.

29. García-Mayordomo J, Martín-Banda R, Insua-Arévalo JM, Álvarez-Gómez JA, Martínez-Díaz JJ, Cabral J (2017) Active fault databases: building a bridge between earthquake geologists and seismic hazard practitioners, the case of the qafi v. 3 database. Natural Hazards and Earth System Sciences 17(8):1447–1459

30. Gardner J, Knopoff L (1974) Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? Bulletin of the Seismological Society of America 64(5):1363–1367

31. Girsanov, I. V. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. Theory of Probability & Its Applications, 5(3), 285-301.

32. Golightly, A., & Wilkinson, D. J. (2010). Markov chain Monte Carlo algorithms for SDE parameter estimation. Learning and Inference for Computational Systems Biology, 253-276.

33. Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. Bulletin of the Seismological Society of America 34(4):185–188

34. Hanks TC, Kanamori H (1979) A moment magnitude scale. Journal of Geophysical Research: Solid Earth 84(B5):2348–2350

35. Härdle, W. K., & Cabrera, B. L. (2010). Calibrating CAT bonds for Mexican earthquakes. Journal of Risk and Insurance, 77(3), 625-650.

36. Harte, D. (2010). PtProcess: An R package for modelling marked point processes indexed by time. Journal of Statistical Software, 35, 1-32.

37. Heath, D., Jarrow, R., Morton, A., & Spindel, M. (1992). Easier done than said. Risk, 5(9), 77-80.

38. Heep-Altiner, M., Mullins, M., & Rohlfs, T. (2018). Solvency II in the Insurance Industry. Springer International Publishing.

39. Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. Journal of statistical software, 27, 1-22.

40. Iacus, S. M. (2009). Simulation and inference for stochastic differential equations: with R examples. Springer Science & Business Media.

41. Jalilian A (2019) Etas: an r package for fitting the spacetime etas model to earthquake data. Journal of Statistical Software 88(1):1–39

42. Kagan YY (2013) Earthquakes: models, statistics, testable forecasts. John Wiley & Sons

43. Kappos AJ (2013) Seismic vulnerability and loss assessment for buildings in Greece. In Seismic Vulnerability of Structures, P. Gueguen (Ed.). https://doi.org/10.1002/9781118603925.ch3

44. Kappos AJ, Panagopoulos G, Panagiotopoulos C, Penelis G (2006) A hybrid method for the vulnerability assessment of r/c and urm buildings. Bulletin of Earthquake Engineering 4(4):391–413

45. Knopoff L, Gardner J (1972) Higher seismic activity during local night on the raw worldwide earthquake catalogue. Geophysical Journal International 28(3):311–313

46. Kohrangi, M., Papadopoulos, A. N., Kotha, S. R., Vamvatsikos, D., & Bazzurro, P. (2021). Earthquake Catastrophe Risk Modeling, Application to the Insurance Industry: Unknowns and Possible Sources of Bias in Pricing. In Advances in Assessment and Modeling of Earthquake Loss (pp. 239-274). Springer, Cham.

47. Kumar, K. (2008). Basic Geotechnical Earthquake Engineering, New Age International Publishers, 2008: Basic Geotechnical Earthquake Engineering (Vol. 1). Bukupedia.

48. Kunreuther H (1996) Mitigating disaster losses through insurance. Journal of Risk and Uncertainty 12(2):171–187

49. Lee, J. P., & Yu, M. T. (2002). Pricing default-risky CAT bonds with moral hazard and basis risk. Journal of Risk and Insurance, 25-44.

50. Lewis, N. D. (2015). Build Your Own Neural Network Today. With Step by Step Instructions Showing You How to Build Them Faster than You Imagined Possible Using R. Create Space Independent Publishing Platform.

51. Lin JH (2018) Earthquake insurance pricing: a risk-based approach. Disasters 42(2):392–404

52. Loguinova, K. (2019). A Critical Legal Study of the Ideology Behind Solvency II. Springer.

53. Loubergé, H., Kellezi, E., & Gilli, M. (1999). Using catastrophe-linked securities to diversify insurance risk: A financial analysis of CAT bonds. Journal of Insurance Issues, 125-146.

54. Louloudis, E., Zimbidis, A. & Yannacopoulos, A. Stochastic assessment of seismic risk using faults to address the incomplete information in historical catalogues. Eur. Actuar. J. (2022). https://doi.org/10.1007/s13385-022-00324-2

55. Ma, Z. G., & Ma, C. Q. (2013). Pricing catastrophe risk bonds: A mixed approximation method. Insurance: Mathematics and Economics, 52(2), 243-254.

56. Marshak, S. (1997). Earth structure: An introduction to structural geology and tectonics. WCB/McGraw-Hill.

57. McGuire, R. K. (2004). Seismic hazard and risk analysis. Earthquake Engineering Research Institute.

58. McNeil AJ, Frey R, Embrechts P (2015) Quantitative risk management: concepts, techniques and tools-revised edition. Princeton University Press

59. Meslem A, Lang DH (2017) Physical vulnerability in earthquake risk assessment. In: Oxford Research Encyclopedia of Natural Hazard Science. DOI: 10.1093/acrefore/9780199389407.013.71

60. Mikosch, T. (1998). Elementary stochastic calculus with finance in view. World scientific.

61. Nowak, P., & Romaniuk, M. (2013). Pricing and simulations of catastrophe bonds. Insurance: Mathematics and Economics, 52(1), 18-28.

62. Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical association, 83(401), 9-27.

63. Ogata Y (1998) Space-time point-process models for earthquake occurrences. Annals of the Institute of Statistical Mathematics 50(2):379–402

64. Pace B, Visini F, Peruzza L (2016) Fish: Matlab tools to turn fault data into seismic-hazard models. Seismological Research Letters 87(2A):374–386

65. Papanikolaou ID, Roberts GP, Deligiannakis G, Sakellariou A, Vassilakis E (2013) The sparta fault, southern Greece: from segmentation and tectonic geomorphology to seismic hazard mapping and time dependent probabilities. Tectonophysics 597:85–105

66. Papazachos B, Comninakis P, Karakaisis G, Karakostas B, Papaioannou CA, Papazachos C, Scordilis E (2000) A catalogue of earthquakes in Greece and surrounding area for the period 550bc–1999. Publ Geoph Lab, Univ of Thessaloniki 1:333

67. Pavlides S, Caputo R, Sboras S, Chatzipetros A, Papathanasiou G, Valkaniotis S (2010) The greek catalogue of active faults and database of seismogenic sources. Bulletin of the Geological Society of Greece 43(1):486–494

68. Peng Y (2011, May) Insurance pricing under uncertainty by means of distortion function. In: 2011 International Conference on E-Business and E-Government (ICEE), IEEE, pp 1–4

69. Peruzza L, Pace B, Cavallini F (2010) Error propagation in time-dependent probability of occurrence for characteristic earthquakes in Italy. Journal of Seismology 14(1):119–141

70. Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. Biometrika, 75(2), 335-346.

71. Pitilakis, K., Franchin, P., Khazai, B., & Wenzel, H. (Eds.). (2014). SYNER-G: systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities: methodology and applications (Vol. 31). Springer.

72. Polacek,A, (2018). Catastrophe Bonds: A Primer and Retrospective. Chicago Fed Letter, No. 405, 2018.

73. Privault, N. (2012). An elementary introduction to stochastic interest rate modeling. World Scientific.

74. Rinaldis D, Berardi R, Theodulidis N, Margaris B (1998) Empirical predictive models based on a joint Italian & Greek strong-motion database. In: Proceedings of Eleventh European Conference on Earthquake Engineering

75. Shao, J., Pantelous, A., & Papaioannou, A. D. (2015). Catastrophe risk bonds with applications to earthquakes. European Actuarial Journal, 5(1), 113-138.

76. Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). Time series analysis and its applications (Vol. 3). New York: springer.

77. Starita, M., & Malafronte, I. (2014). Capital requirements, disclosure, and supervision in the European insurance industry: new challenges towards Solvency II. Springer.

78. Stupfler, G., & Yang, F. (2018). Analyzing and predicting cat bond premiums: A financial loss premium principle and extreme value modeling. ASTIN Bulletin: The Journal of the IAA, 48(1), 375-411.

79. Svensson, L. E. (1995). Estimating forward interest rates with the extended Nelson & Siegel method. Sveriges Riksbank Quarterly Review, 3(1), 13-26.

80. Talbi A, Nanjo K, Satake K, Zhuang J, Hamdache M (2013) Comparison of seismicity declustering methods using a probabilistic measure of clustering. Journal of Seismology 17(3):1041–1061

81. Tao Z, Wu DD, Zheng Z, Tao X (2010) Earthquake insurance and earthquake risk management. Human and Ecological Risk Assessment 16(3):524–535

82. Towhata, I. (2008). Geotechnical Earthquake Engineering. Springer.

83. Tsanakas, A., & Desli, E. (2005). Measurement and pricing of risk in insurance markets. Risk Analysis: An International Journal, 25(6), 1653-1668.

84. Valentini A, Visini F, Pace B (2017) Integrating faults and past earthquakes into a probabilistic seismic hazard model for peninsular Italy. Natural Hazards Earth System Sciences.

85. Varotsos, P., Alexopoulos, K., Nomicos, K. (1981a). Seismic electric currents. Proceedings of the Academy of Athens 56: 277–286

86. Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of financial economics, 5(2), 177-188.

87. Wells DL, Coppersmith KJ (1994) 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

88. Whitman, R. V., Reed, J. W., & Hong, S. T. (1973, June). Earthquake damage probability matrices. In Proceedings of the Fifth World conference on earthquake engineering (Vol. 2, pp. 2531-2540). Rome, Italy: Palazzo DEI CONGRESSI (EUR).

89. Wiemer S (2001) A software package to analyze seismicity: ZMAP. Seismological Research Letters 72(3):373–382

90. Yeats, R. (2012). Active faults of the world. Cambridge University Press.

91. Youngs RR, Coppersmith KJ (1985) Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. Bulletin of the Seismological society of America 75(4):939–964

92. Yucemen M (2005) Probabilistic assessment of earthquake insurance rates for Turkey. Natural Hazards 35(2):291–313

93. Zhuang, J. (2012). Long-term earthquake forecasts based on the epidemic-type aftershock sequence (ETAS) model for short-term clustering. Research in Geophysics, 2(1), e8-e8.

94. Zhuang, J., & Ogata, Y. (2006). Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks. Physical Review E, 73(4), 046134.

95. Zimbidis, A. A., Frangos, N. E., & Pantelous, A. A. (2007). Modeling earthquake risk via extreme value theory and pricing the respective catastrophe bonds. ASTIN Bulletin: The Journal of the IAA, 37(1), 163-183.

Appendix

Table A1: Premium rating under AF10 model

LON	LAT	REG	POS	L	M	H
23.70203	37.95552	ATT	17672	570	139	69
23.70381	38.08099	ATT	13341	1087	250	125
23.81193	38.07281	ATT	14561	746	274	137
23.75344	37.86282	ATT	16674	478	110	55
23.69955	37.93208	ATT	17561	541	147	73
23.75725	37.93191	ATT	16342	588	154	77
23.74914	37.91107	ATT	16452	601	121	60
23.75792	38.13064	ATT	13676	1151	279	140
23.76154	37.97615	ATT	15771	521	177	88
23.96096	38.15526	ATT	19007	822	194	97
23.88193	38.00473	ATT	15351	725	174	87
23.85231	37.95401	ATT	19002	560	171	86
24.05712	37.71496	ATT	19500	294	57	29
23.54212	38.04323	ATT	19200	1072	269	135
23.49645	37.96421	ATT	18900	815	230	115
23.8075	38.05457	ATT	15124	907	261	130
23.8019	37.81836	ATT	16672	418	70	35
23.64693	37.94292	ATT	18535	618	162	81

23.83581	38.04966	ATT	15127	832	262	131
23.85963	38.14042	ATT	14565	803	250	125
23.71506	37.91368	ATT	17455	499	149	74
23.73473	37.94994	ATT	17235	737	157	78
23.79551	38.00501	ATT	15561	764	179	90
23.75581	38.01821	ATT	11147	780	223	112
23.75353	37.95507	ATT	16232	578	159	80
23.77612	38.06198	ATT	14122	858	240	120
23.65941	37.99038	ATT	12351	771	173	86
23.72977	38.0507	ATT	13562	937	281	141
23.67848	37.99223	ATT	12241	719	215	108
23.68491	38.04189	ATT	13231	990	247	123
23.66446	38.00506	ATT	12461	778	195	97
23.75214	38.04316	ATT	14231	841	275	138
23.712	37.93188	ATT	17123	591	148	74
23.72896	37.98618	ATT	10432	586	160	80
23.75852	38.06114	ATT	14452	860	234	117
23.71031	38.01227	ATT	12133	694	213	107
23.87291	37.90004	ATT	19400	472	94	47
23.94367	37.8881	ATT	19003	402	93	47
24.01216	38.01726	ATT	19009	541	128	64
23.49539	38.07622	ATT	19600	1196	306	153
23.72409	37.92426	ATT	17342	518	149	74
23.70673	38.0348	ATT	13122	1024	222	111
23.77491	38.00356	ATT	15451	603	192	96
23.73209	37.99412	ATT	10434	799	161	80
23.82762	38.03645	ATT	15235	751	216	108
23.76568	37.84049	ATT	16673	462	125	62

23.65314	37.94207	ATT	18534	785	162	81
23.81007	38.07659	ATT	14561	1071	230	115
23.76461	37.88315	ATT	16561	522	124	62
23.71178	37.97872	ATT	11854	741	161	80
23.68544	37.9435	ATT	17674	662	123	61
23.76534	37.99299	ATT	11523	605	154	77
23.66657	37.97476	ATT	18233	752	163	82
23.70523	40.66364	CMC	57014	698	168	84
22.94226	40.63101	CMC	54623	1341	334	167
23.00504	40.72466	CMC	57200	1159	395	197
22.24465	40.49624	CMC	59100	743	229	115
22.8745	40.99389	CMC	61100	1050	323	161
23.82878	41.03015	CMC	62042	382	92	46
23.28228	41.18158	CMC	62400	859	229	115
22.93031	40.67187	CMC	56431	1126	316	158
23.86562	40.17014	CMC	63078	636	144	72
22.95452	40.58541	CMC	55133	1102	314	157
23.0323	40.58168	CMC	55236	1006	308	154
22.9226	40.6438	CMC	54627	1175	362	181
22.90862	40.66949	CMC	56224	1154	307	154
22.94631	40.66158	CMC	56533	1160	340	170
23.54281	41.08831	CMC	62125	633	154	77
22.50866	40.27148	CMC	60132	437	59	30
23.10803	40.29715	CMC	63080	505	114	57
23.44028	40.3778	CMC	63100	518	123	61
22.05069	40.8024	CMC	58200	1060	214	107
22.94309	40.63478	CMC	54624	1061	302	151
23.8633	40.89849	CMC	62041	349	119	59

22.41598	39.63878	THE	41222	550	102	51
21.76825	39.55562	THE	42132	698	164	82
21.92381	39.36755	THE	43131	958	289	145
22.94659	39.3615	THE	38221	945	286	143
22.74508	39.38092	THE	37500	781	181	91
21.118	39.04306	WGR	30500	1446	380	190
22.0811	38.25058	WGR	25100	2355	675	337
21.73507	38.24621	WGR	26221	2423	664	332
21.44249	37.67254	WGR	27131	2655	768	384
24.14679	41.15005	EMT	66133	264	45	22
24.40948	40.9365	EMT	65302	164	34	17
24.88367	41.1356	EMT	67133	241	34	17
25.13475	35.33122	CRE	71305	1107	299	149
24.01881	35.51624	CRE	73132	910	210	105
24.47945	35.36557	CRE	74131	1050	305	153
22.80866	37.56818	PEL	21100	779	211	105
22.93203	37.93858	PEL	20131	1345	345	173
22.10912	37.02737	PEL	24131	1435	493	247
21.78596	40.30103	WMC	50131	1087	264	132
21.42584	40.08355	WMC	51100	1276	378	189
20.84021	39.67453	EPI	45333	1732	501	251
20.98617	39.15907	EPI	47132	1538	386	193
22.43368	38.8999	CGR	35132	1616	412	206
23.3229	38.31986	CGR	32200	1958	529	264
20.43696	38.20129	ION	28200	3079	803	401
25.16147	37.53807	SAE	84200	162	18	9

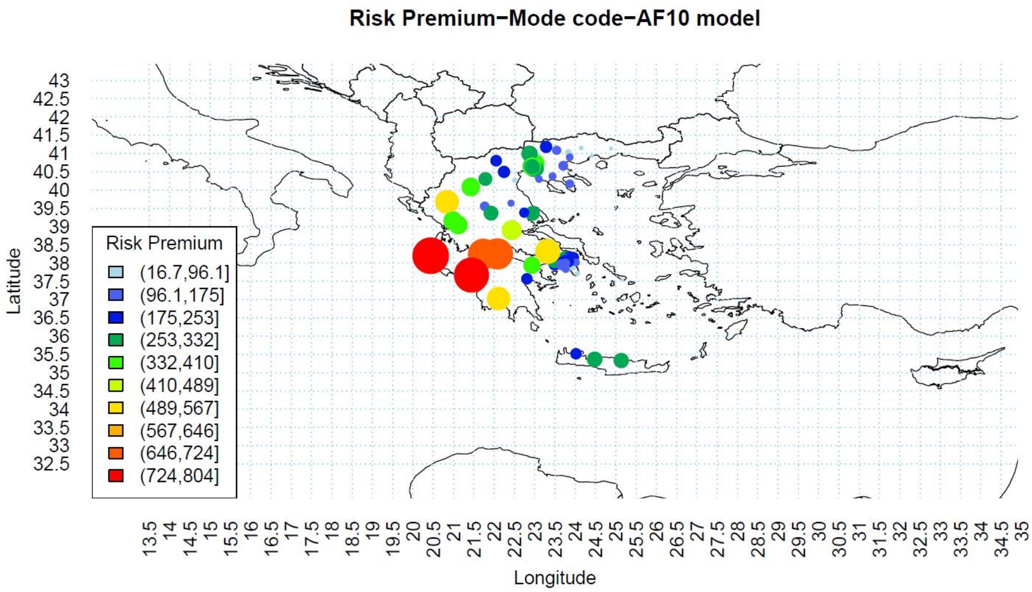

Figure A1: Premium mode code rates under model AF10.

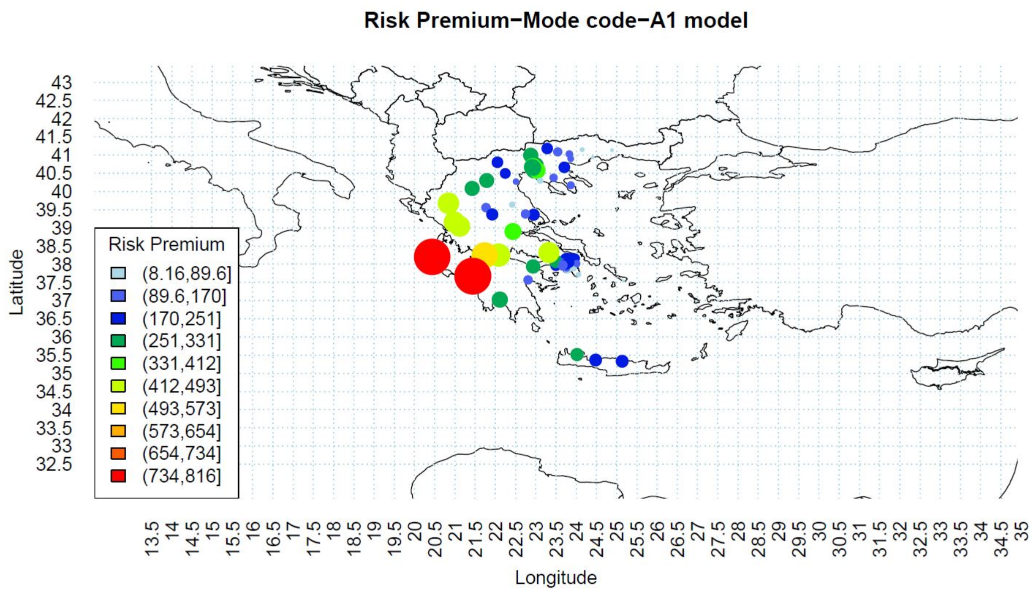

Figure A2: Premium mode code rates under model A1.

Table A2: Premium mode code rating under AFe model

LON	LAT	REG	POS	L	M	H
23.70203	37.95552	ATT	17672	878	195	97
23.70381	38.08099	ATT	13341	1116	285	142
23.81193	38.07281	ATT	14561	1103	261	131
23.75344	37.86282	ATT	16674	468	112	56
23.69955	37.93208	ATT	17561	679	188	94
23.75725	37.93191	ATT	16342	902	200	100
23.74914	37.91107	ATT	16452	752	174	87
23.75792	38.13064	ATT	13676	916	240	120
23.76154	37.97615	ATT	15771	1029	340	170
23.96096	38.15526	ATT	19007	511	119	59
23.88193	38.00473	ATT	15351	661	241	121
23.85231	37.95401	ATT	19002	629	162	81
24.05712	37.71496	ATT	19500	263	57	29
23.54212	38.04323	ATT	19200	890	221	111
23.49645	37.96421	ATT	18900	957	232	116
23.8075	38.05457	ATT	15124	1221	343	171
23.8019	37.81836	ATT	16672	409	87	43
23.64693	37.94292	ATT	18535	774	174	87
23.83581	38.04966	ATT	15127	1162	233	117
23.85963	38.14042	ATT	14565	725	170	85
23.71506	37.91368	ATT	17455	779	183	92
23.73473	37.94994	ATT	17235	889	253	126
23.79551	38.00501	ATT	15561	1271	377	189
23.75581	38.01821	ATT	11147	1519	547	273
23.75353	37.95507	ATT	16232	1043	203	101
23.77612	38.06198	ATT	14122	1438	516	258

23.65941	37.99038	ATT	12351	858	236	118
23.72977	38.0507	ATT	13562	1366	410	205
23.67848	37.99223	ATT	12241	1022	314	157
23.68491	38.04189	ATT	13231	1170	327	163
23.66446	38.00506	ATT	12461	1051	276	138
23.75214	38.04316	ATT	14231	1786	518	259
23.712	37.93188	ATT	17123	664	210	105
23.72896	37.98618	ATT	10432	1219	324	162
23.75852	38.06114	ATT	14452	1325	405	203
23.71031	38.01227	ATT	12133	1323	391	196
23.87291	37.90004	ATT	19400	565	116	58
23.94367	37.8881	ATT	19003	395	69	35
24.01216	38.01726	ATT	19009	410	107	54
23.49539	38.07622	ATT	19600	998	297	149
23.72409	37.92426	ATT	17342	655	169	85
23.70673	38.0348	ATT	13122	1245	340	170
23.77491	38.00356	ATT	15451	1442	397	198
23.73209	37.99412	ATT	10434	1343	334	167
23.82762	38.03645	ATT	15235	1041	269	134
23.76568	37.84049	ATT	16673	553	126	63
23.65314	37.94207	ATT	18534	674	175	87
23.81007	38.07659	ATT	14561	878	295	147
23.76461	37.88315	ATT	16561	591	97	48
23.71178	37.97872	ATT	11854	1096	290	145
23.68544	37.9435	ATT	17674	785	211	106
23.76534	37.99299	ATT	11523	1120	406	203
23.66657	37.97476	ATT	18233	815	240	120
23.70523	40.66364	CMC	57014	2223	673	337

22.94226	40.63101	CMC	54623	1792	587	294
23.00504	40.72466	CMC	57200	1585	450	225
22.24465	40.49624	CMC	59100	667	124	62
22.8745	40.99389	CMC	61100	979	289	144
23.82878	41.03015	CMC	62042	457	97	49
23.28228	41.18158	CMC	62400	849	202	101
22.93031	40.67187	CMC	56431	1899	673	336
23.86562	40.17014	CMC	63078	733	187	93
22.95452	40.58541	CMC	55133	1468	461	231
23.0323	40.58168	CMC	55236	1787	436	218
22.9226	40.6438	CMC	54627	2011	604	302
22.90862	40.66949	CMC	56224	2269	609	304
22.94631	40.66158	CMC	56533	1877	558	279
23.54281	41.08831	CMC	62125	516	127	64
22.50866	40.27148	CMC	60132	486	83	42
23.10803	40.29715	CMC	63080	591	142	71
23.44028	40.3778	CMC	63100	763	199	100
22.05069	40.8024	CMC	58200	627	134	67
22.94309	40.63478	CMC	54624	1813	603	301
23.8633	40.89849	CMC	62041	510	101	50
22.41598	39.63878	THE	41222	448	59	30
21.76825	39.55562	THE	42132	668	146	73
21.92381	39.36755	THE	43131	2406	701	351
22.94659	39.3615	THE	38221	1856	537	268
22.74508	39.38092	THE	37500	2739	793	396
21.118	39.04306	WGR	30500	972	205	102
22.0811	38.25058	WGR	25100	1949	549	274
21.73507	38.24621	WGR	26221	2327	597	298

21.44249	37.67254	WGR	27131	4261	1338	669
24.14679	41.15005	EMT	66133	310	60	30
24.40948	40.9365	EMT	65302	209	36	18
24.88367	41.1356	EMT	67133	165	30	15
25.13475	35.33122	CRE	71305	1905	530	265
24.01881	35.51624	CRE	73132	768	178	89
24.47945	35.36557	CRE	74131	803	161	81
22.80866	37.56818	PEL	21100	1008	288	144
22.93203	37.93858	PEL	20131	2173	743	371
22.10912	37.02737	PEL	24131	904	227	114
21.78596	40.30103	WMC	50131	923	265	133
21.42584	40.08355	WMC	51100	899	204	102
20.84021	39.67453	EPI	45333	1432	310	155
20.98617	39.15907	EPI	47132	1201	327	164
22.43368	38.8999	CGR	35132	1403	533	267
23.3229	38.31986	CGR	32200	1091	231	115
20.43696	38.20129	ION	28200	3580	884	442
25.16147	37.53807	SAE	84200	183	30	15

HJM theory and the relation with the Vasicek model

The following hold for the forward rate $f(t,T)$:

$$df(t,T) = \sigma(t,T)dW_t + a(t,T)dt.$$

Integrating df:

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T)dW_s + \int_0^t a(s,T)ds.$$

Then, the short rate

$$r(t) = \lim_{t \to T} F(t,T) = f(0,t) + \int_0^t \sigma(s,t)dW_s + \int_0^t a(s,t)ds.$$

The accumulation factor

$$B(t) = \text{exp}[\int_0^t r(s)ds]$$

and let

$$B(t,T) = \text{exp}[-\int_t^T f(t,u)du].$$

Then, the present value is

$$Z(t,T) = B^{-1}(t) + B(t,T) = \text{exp}(-X_t)$$

where:

$$X_t = \int_0^T f(0,u)du + \int_0^t \int_s^T \sigma(s,u)dudW_s + \int_0^t \int_s^T a(s,u)dudt.$$

Applying Ito's lemma to $Z(t,T)$, we obtain:

$$dZ(t,T) = Z(t,T)\left[-v(t,T)dW_t - \int_t^T a(t,u)dudt + \frac{v^2(t,T)}{2}dt\right],$$

where $v(t,T) = \int_t^T \sigma(t,u)du$. The two necessary conditions in order to make Z(t,T) a martingale and suppress arbitrage opportunities are respectively:

1. $d\tilde{W}_t = dW_t + \gamma_t, \;\; \gamma_t = \frac{1}{v(t,T)} \int_t^T a(t,u)dt - \frac{v(t,T)}{2},$
2. $a(t,T) = \sigma(t,T)[v(t,T) + \gamma_t.$

For the case of Vasicek, assuming $\gamma_t = 0$ and $\sigma(t,T) = \sigma \exp[-\alpha(T-t)]$, where α is the speed reversion, we obtain the well-known equation of the Vasicek SDE.

Simulation based short-rate pricing

The discount process is described by $B(t) = E^Q[\exp(-\int_0^t r(u)du)]$ (Privault, 2012), where $r(t)$ is the short-rate stochastic process denoting the daily Libor rate. and Q is the risk neutral (or martingale) measure. The SDE under the initial natural measure P is:

$$dr_t = a(b - r_t)dt + \sigma dW_t.$$

By the theorem of Girsanov (1960), the change of measure by

$$dW_t = d\tilde{W}_t - \lambda d_t$$

leads to a standard Brownian motion $\tilde{W}_t$, where λ is the market price of risk which gives the extra increase in the expected instantaneous rate of return on the bond per an additional unit of risk.

The SDE under the new measure Q becomes:

$$dr_t = a\left(b - \frac{\lambda\sigma}{a} - r(t)\right)dt + \sigma d\tilde{W}_t.$$

A closed form solution for the Vasicek interest rate model is presented in Nowak (2013). Nowak considered the parameters as estimated by Episcopos (2000). However, we present a simulation methodology which fits to every SDE form. We assume that the 3-month IBR obeys the Vasicek interest rate law, where $r(0)$ is the last observable rate in our data. Following Mikosch (1998), we denote $\tilde{W}_t = W(t) + \lambda_t$ the Brownian motion which is a martingale under the equivalent to P measure Q and $dt = 1$ day. We produce many (M=5000) paths. Each path j has a ten-year maturity for the stochastic processes, and it is constructed as follows:

$$W(0) = 0$$

$$W(t) = W(t-1) + dW_t, \;\; dW(t) \sim N(0,1)$$

$$\tilde{W}(t) = W(t) + \lambda t$$

$$r(t) = r(t-1) + [a(b^* - r(t-1))]dt + \sigma(\tilde{W}(t) - \tilde{W}(t-1)),$$

where $b^* = b - \frac{\lambda\sigma}{a}$. Then,

$$B(k) = E^Q \left[\text{exp} \left(- \sum_{t=0}^{4k} r(t) \right) \right]$$

is the discounting factor for a payment that occurs k years after today (t=0), $k = 1, \ldots, 10$, estimated using the Monte-Carlo method by

$$\hat{B}_k = \frac{\sum_{j=1} M \, \text{exp}(- \sum_{t=0}^{4k} r(t))}{M}.$$

Results using the Short-rate Vasicek model for overnight libor

For example, we first consider the overnight libor rate (in euros) as the discounting short rate. With the use of the Vasicek model, the pricing of long maturity bonds is highly dependent on the choice of the historical data period, resulting in prices of approximately 450 euros for 100 euros of face value in the time period of 2020-2021 or even approximately 22 euros for 100 euros of face value in the time period 2009-2011 for a zero-coupon bond.

From 2000 up to 2009 estimates of the Vasicek model and of B(261 days)=B(1 year): We chose 261 days as the number that LIBOR is announced per year.

$$a_d = 0.017, b_d = 2.7\%, \sigma = 0.2\%$$

and $B(261) = B(1\text{year}) = 0.015$ euros at T=0 for 1 euro at time T=1 year.

From 2009 up to 2011 estimates of the Vasicek model and of B(261 days)=B(1 year):

$$a_d = 0.08, b_d = 0.6\%, \sigma = 0.15\%$$

and $B(261) = B(1\text{year}) = 0.22$ euros at T=0 for 1 euro at time T=1 year.

From 2020 up to 2021 estimates of the Vasicek model and of B(261 days)=B(1 year):

$$a_d = 0.09, b_d = -0.6\%, \sigma = 0.004\%$$

and $B(261) = B(1\text{year}) = 4.5$ euros at T=0 for 1 euro at time T=1 year.

Insurance pricing algorithm using the ETAS model and fault sources

1. Fill a matrix $\mathbf{F_N}$ containing simulations of annual number of earthquakes for each fault source described by a mix of a Poisson and a time-dependent process with lognormal inter-event times as in chapter 2.

2. Fill a matrix $\mathbf{A_N}$ containing simulations of annual number of earthquakes for each area source based on the background seismicity described by a homogeneous in time Poisson Process as in chapter 2.

3. Fill a matrix $\mathbf{F_M}$ containing simulations of magnitudes of earthquakes for each fault source described by a truncated Gaussian distribution as in chapter 2.

4. Fill a matrix $\mathbf{A_M}$ containing simulations of maximum magnitudes over the whole cloud generated by a background event for each area source as described in chapter 2.

5. Create an empty zero matrix $\mathbf{G_4}$ containing the maximum over-a-year loss due to all seismic sources

6. For each area source create an empty zero matrix $\mathbf{G_1}$ containing several simulations of the maximum over-a-year losses of each building due to each source

7. Generate earthquakes based on $\mathbf{A_N}$ and the maximum magnitude $\mathbf{A_M}$ for each area source, attenuate it to all buildings using the attenuation equation (2.7) and transform it to damage using the fragility curves of Kappos et al. (2006) and store it to $\mathbf{G_1}$.

8. Take the element-by-element maximum damage between the previous and next matrix $\mathbf{G_1}$ before moving to the next area source. Repeat for all area

sources to result to a matrix $\mathbf{G_2}$ containing simulations of the maximum over-a-year damage from all area sources to each building of the investigated portfolio.

9. For each fault source create an empty zero matrix $\mathbf{G_3}$ containing several simulations of the maximum over-a-year losses of each building due to each source

10. Generate earthquakes based on $\mathbf{F_N}$ and keep the maximum magnitude based on $\mathbf{F_M}$ for each fault source, attenuate it to all buildings using the attenuation equation (2.8) and transform it to damage using the fragility curves of Kappos et al. (2006) and store it to $\mathbf{G_3}$.

11. Take the element-by-element maximum between the previous $\mathbf{G_2}$ and next matrix $\mathbf{G_3}$ before moving to the next fault source. Repeat for all fault sources to result to a matrix $\mathbf{G_4}$ containing simulations of the maximum over-a-year damage from all sources to each building of the investigated portfolio.

12. The expected loss of each building of $\mathbf{G_4}$ is its risk premium, while risk measures are applied to the total loss to evaluate the SCR.

STAN algorithm for the derivation of the 12-month LIBOR Vasicek estimates

```
vas.stan = ”
data {
int<lower=0> N;
vector[N] libor12;
}
parameters {
real alpha;
real beta;
real<lower=0> sigma;
}
model {
libor12[2:N] ~ normal(alpha*beta + (1-alpha) * libor12[1:N-1], sigma);
}
” fit = stan(model_ code=vas.stan, data=list(libor12=libor12, N=length(libor12)))
theta_ draws = extract(fit)
(alpha2<-mean(theta_ draws$alpha))
(beta2<-mean(theta_ draws$beta))
(sigma2<-mean(theta_ draws$sigma))
```

Dynamic Credit Risk CAT bond pricing for time independent cases

1. For the 3-month IBR of Germany

   (a) 2000-2009

      i. Issuer 1

Table A3: CAT bond pricing using the 3-month IBR of Germany for issuer 1 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	92.50	89.39	87.32	85.03	83.23	81.45	79.48	78.46
2%	94.90	92.92	90.91	89.29	87.77	86.62	85.76	85.17
3%	97.68	96.46	95.12	94.30	93.35	92.59	92.14	91.79
4%	100.12	99.86	99.56	99.46	98.78	98.95	98.79	99.09
5%	103.34	103.80	103.81	104.38	105.01	105.36	105.95	106.63

      ii. Issuer 2

Table A4: CAT bond pricing using the 3-month IBR of Germany for issuer 2 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	54.64	50.71	48.21	46.43	45.08	43.98	43.00	42.18
2%	57.52	54.08	51.75	50.49	49.76	48.80	48.37	47.76
3%	58.33	55.29	53.82	52.92	52.25	51.93	51.77	51.59
4%	60.14	57.52	56.34	55.94	55.73	55.63	55.71	55.75
5%	62.40	60.41	59.50	59.02	59.11	59.02	59.28	59.56

(b) 2010-2015

i. Issuer 1

Table A5: CAT bond pricing using the 3-month IBR of Germany for issuer 1 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	95.53	94.05	91.83	90.10	88.74	87.33	85.83	84.87
2%	98.66	97.48	96.61	95.79	95.25	94.04	93.52	92.97
3%	101.14	100.95	100.64	100.32	100.28	99.75	99.58	99.32
4%	104.40	105.21	106.36	107.16	108.23	108.72	109.72	110.82
5%	107.77	109.51	111.03	113.18	114.30	115.96	117.63	118.95

ii. Issuer 2

Table A6: CAT bond pricing using the 3-month IBR of Germany for issuer 2 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	56.74	53.39	51.26	49.99	49.03	48.08	47.10	46.54
2%	58.36	55.04	53.17	51.87	51.02	50.49	50.13	49.67
3%	60.23	57.67	56.43	55.83	55.52	55.23	55.21	55.41
4%	62.33	60.37	60.01	59.89	60.17	60.16	60.66	61.06
5%	64.58	62.70	62.61	62.92	63.43	64.35	65.16	65.98

2. ECB yield curves modeled with Nelson-Siegel time independent model

(a) 2004-2009

i. Issuer 1

Table A7: CAT bond pricing using the ECB yield curves for issuer 1 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	91.03	87.33	83.97	80.39	77.31	73.95	70.86	68.11
2%	93.93	90.84	87.76	84.64	81.90	78.88	76.23	73.79
3%	96.56	94.30	92.29	89.80	87.23	84.74	82.72	80.60
4%	99.35	97.74	96.33	94.77	92.84	91.18	89.45	87.64
5%	101.94	101.37	100.48	99.10	97.99	96.68	95.55	94.20

ii. Issuer 2

Table A8: CAT bond pricing using the ECB yield curves for issuer 2 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	53.40	49.18	46.07	43.63	41.92	40.08	38.37	36.82
2%	56.18	52.25	49.41	47.08	45.38	43.94	42.58	41.21
3%	56.78	53.42	51.34	49.59	48.12	46.80	45.63	44.56
4%	59.56	56.53	54.67	53.35	52.28	51.34	50.35	49.49
5%	61.02	58.56	57.04	55.82	55.00	54.19	53.42	52.70

(b) 2010-2015

i. Issuer 1

Table A9: CAT bond pricing using the ECB yield curves for issuer 1 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	89.79	84.76	79.67	74.92	70.62	66.04	61.75	57.74
2%	92.58	88.67	84.72	80.43	76.54	72.39	68.44	64.68
3%	95.25	91.87	88.18	84.77	81.25	77.57	74.27	70.98
4%	97.87	95.59	92.88	89.92	86.82	83.91	81.06	78.59
5%	101.01	99.14	96.92	94.46	92.08	89.86	87.61	85.19

ii. Issuer 2

Table A10: CAT bond pricing using the ECB yield curves for issuer 2 under dynamic credit risk

Spread|maturity	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
1%	52.40	47.58	43.76	40.71	38.10	35.60	33.29	31.19
2%	54.21	49.47	46.01	43.45	41.18	39.09	36.88	34.95
3%	57.77	53.21	50.24	47.72	45.34	43.37	41.61	39.92
4%	57.72	53.80	51.33	49.36	47.32	45.66	44.25	42.77
5%	60.10	56.86	54.66	52.95	51.33	50.10	48.76	47.51